Doctoral theses at NTNU, 2021:415

Rabia Naseem

Cross-modality Guided Image Enhancement

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

Rabia Naseem

Cross-modality Guided Image Enhancement

Thesis for the Degree of Philosophiae Doctor

Gjøvik, December 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Department of Computer Science

ISBN 978-82-326-6129-9 (printed ver.)
ISBN 978-82-326-5251-8 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:415

Printed by NTNU Grafisk senter

Abstract

The quality of medical images is a crucial factor that affects the performance of several image analysis tasks. Low contrast and noise are among the widely investigated distortions in medical image enhancement problems. In this thesis, the approaches to improve the contrast of medical images and reduce the noise have been proposed by particularly investigating how the cross-modal guidance from another medical image impacts the enhancement. We are particularly interested in enhancing Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) which are widely used in both diagnosis and therapy planning. The first section of the thesis focuses on contrast enhancement and the second section focuses on denoising. This dissertation presents our research work supported by six original publications (including five published papers and one accepted for publication).

First, in the context of cross-modality guided contrast enhancement, two traditional global enhancement approaches are proposed to improve the contrast of CT images of the human liver using corresponding MR images. The first approach uses context-aware two-dimensional histogram specification (HS) and morphological operations. The objective of this scheme is to improve the visibility of the organ's anatomy to facilitate the tasks of surgeons and radiologists. The second uses 2D-HS followed by an optimization scheme to minimize the artifacts associated with histogram-based methods and simultaneously preserves the structure of the image during enhancement. In this approach, the enhanced images are analyzed from two perspectives (contrast enhancement and improvement in tumor segmentation). Both techniques have been validated on multi-modal data acquired from a hospital in Norway. Furthermore, an acceleration scheme was proposed by parallelizing the steps involved in the proposed CE approach which drastically reduced the execution time of the algorithm. The third method uses deep learning to improve the contrast of medical images using guidance from multi-modal MR images. Cycle-GAN (Generative Adversarial Network) was applied for this purpose where the corresponding high-contrast image from another modality was used as ground truth as opposed to using manually enhanced ground truth/ reference image.

Secondly, noise is another artifact that affects the visual quality of medical images. It not only hampers the visibility of structures for clinicians who inspect these images to thoroughly understand the organ's morphology; but it also affects the subsequent image analysis tasks. It is therefore imperative to remove noise and improve the perceptual quality of medical images. Different kinds of noise contaminate medical images. In this thesis, we proposed a method to denoise T1-weighted (T1-w) MR images contaminated with Rician noise. We exploited the

complementarity-aware information in better perceptual quality multi-modal medical images for denoising purpose. In particular, the role of deep learning approach was investigated in this regard. The features from dual images were combined in a hierarchical manner to extract rich features, which are later combined in a systematic way as opposed to simple feature concatenation. The performance was validated on two public datasets both from a qualitative and quantitative perspective. Moreover, the comparison was done with single image denoising schemes on varying levels of noise.

Acknowledgements

I would like to express my gratitude to all the people who supported me in completing this Ph.D. This research work was accomplished under the auspices of H2020 project 'High-Performance Soft Tissue Navigation (HiPerNav)'. HiPerNav provided me with the opportunity to broaden my knowledge in the area of medical image processing and opened avenues for international collaboration.

I am extremely appreciative of my supervisors who inspired me in several ways. I wish to especially thank Prof. Faouzi Alaya Cheikh for enabling me to join NTNU for my Ph.D., his patience, guidance, and tremendous support to tackle all the challenges I encountered. My Ph.D. would not have been completed without the guidance of Prof. Azeddine Beghdadi, who helped me a lot in refining technical ideas and scientific writing. My special appreciation goes to Prof. Ole Jakob Elle, who administered HiPerNav with competence and provided data for conducting my Ph.D. research. I am also very thankful to Prof. Frank Lindseth for his valuable feedback. All my fellow Ph.Ds in HiPerNav and collaborators are also appreciated for the valuable discussions.

I wish to express my sincere respect to Prof. Marius Pedersen, who ensured to provide all the facilities at NTNU timely and efficiently to support the PhDs. A big thanks to my teachers, friends, and colleagues at NTNU for their support, suggestions, and social meetings during the Ph.D.: Congcong Wang, Irina-Mihaela Ciortan, Anuja, Mekides, Ali Shariq, Vlado, Gabriella, Mohib, Bilal, Sankini, Ying, Majid Ansari, Rafael Palomar, Ahmed Kedir, Helene, Kumiko, Anshul, Gregory, Marina Shalaginova, Urszula, Anne Hilde, Stine Terese, Ingrid von Schantz Bakka and many others. My friends outside NTNU also motivated me a lot. Thanks to Farzana, Ayesha, Saira, Ammara, Nida, Sadaf, Shaista.

Last but not the least, my deepest gratitude goes to my mother for her unconditional love and my siblings for their support throughout my Ph.D. journey. I dedicate this thesis to my beloved father, a friend I relied on for everything, and who supported me tremendously in everything I opted for in my life. I wish he would have seen me completing my Ph.D.

Contents

List of Abbreviations

2D	Two dimensional 15
3D	Three dimensional 13
BRISQUE	Blind/Referenceless Image Spatial Quality Evaluator 43
CDF	Cumulative Distribution Function 20
CE	Contrast Enhancement 5
CLAHE	Contrast-limited adaptive histogram equalization 21
CMAR	Cross-modal Assisted Reconstruction 44
CMGDNet	Cross-modality Guided Denoising Network 44
CMGE	Cross-modality Guided Enhancement 38
CNN	Convolutional Neural Network 31
CT	Computed Tomography 4
FSIM	Feature Similarity Index 45
GAN	Generative Adversarial Network 32
HCC	Hepatocellular Carcinoma 3
HE	Histogram Equalization 20
HS	Histogram Specification 20
MCCEE	Multi-Criteria Contrast Enhancement Evaluation 39

MRI Magnetic Resonance Imaging 4

MSE Mean Squared Error 32

OPTGCE OPTimized Guided Contrast Enhancement 38

PHL Paired Hierarchical Learning 44

SNN Siamese Neural Network 35

SSIM Structural Similarity Index 22

T1-w T1-weighted 5

List of Tables

List of Figures

Part I

Introduction

Chapter 1

Introduction

In this chapter, we present an overview of our research project describing the motivation and context, research questions addressed, list of the published papers and the organization of the thesis.

1.1 Motivation and Context

The incidence of liver cancer is rising globally, making it fifth most frequent cancer and second deadliest in the world [11]. The most common type of primary liver cancer is Hepatocellular carcinoma (Hepatocellular Carcinoma (HCC)) and the potential curative procedure is its surgical resection. During surgical resection, the cancerous liver area need to be completely removed while simultaneously preserving maximum residual healthy tissue. Due to limited hepatic regenerative capability, cancer prognosis is generally poor; therefore, surgical resection of tumor can only be performed for nearly 15% patients even in the initial stages [12].

With advancements in medical imaging and laparoscopic devices, liver tumor resection is progressively being done via laparoscopy. Laparoscopic surgery offers several benefits over open surgery thanks to its minimally invasive nature [13], [14], [15]. During laparoscopic surgery, few small incisions are made into patient's abdomen to pass the surgical instruments and camera. Reduced blood loss, less pain and faster post operative recovery are some of the benefits associated with laparoscopic surgeries. A study [16] compared open surgery versus laparoscopic surgery and reported significantly less post-operative complications in case of latter. Despite several benefits, few challenges also need to be addressed while performing these surgeries, which include poor visualization of the field of view, absence of tactile sense and lack of depth information. In this context, using image

Figure 1.1: Illustration of a typical Image Guided Surgery, Image: courtesy of HiPerNav partners from Oslo University Hospital

guidance techniques can significantly streamline the smoothness of laparoscopic surgeries.

Image guided surgery employs computer based techniques that enable surgeons to correlate the operative field with pre-operative images. This configuration improves the overall surgical safety and precision when incorporated with the laparoscopic interventions. The setup of image guided surgery is demonstrated in figure 1.1. Image guided surgery has witnessed several developments owing to development of efficient imaging systems and computing technologies [17]. Initially developed for neurosurgery, this setup was later incorporated with the spine surgery [18]; however, its adaption to soft tissue surgery is in the early stage. The challenges posed to soft tissue surgery navigation include efficient pre-operative modeling and planning coupled with intra-operative navigation guidance [19].

Pre-operative planning is heavily dependent on the quality of diagnostic images acquired before surgery based on which intra-operative planning is done [20]. Pre-operative diagnostic imaging encompasses various non-invasive techniques that allow clinicians to analyze the anatomy of certain organs and determine the underlying cause of abnormality. It also aids in determining the response of patients to certain treatments. In the context of cancer treatment, Computed Tomography (Computed Tomography (CT)) imaging remains widely accepted as a primary pre-operative evaluation tool. Magnetic Resonance Imaging (Magnetic Resonance Imaging (MRI)) is influential in characterizing indeterministic lesions due to its higher sensitivity in comparison with CT [21]. It is a usual practice to acquire both CT and MRI during therapeutic procedures. During image acquisition, noise and artifacts are introduced making image analysis difficult both for doctors as well as

for the computer vision algorithms. While CT image acquisition system is inexpensive and has faster acquisition, it is not safe for specific patients. MRI is prone to contain motion artefacts and has acquisition time as long as 1.5 hours. Nevertheless, the concordance of MRI and CT findings also becomes essential in case of suspicion.

This thesis aims at investigating and developing computational methods to improve the quality of medical images such as CT and MRI. This enhancement improves the visibility of the internal anatomy of the organs; furthermore, it makes the medical images suitable for subsequent image processing tasks and analysis such as segmentation, registration, feature extraction and classification. This work is implemented under the auspices of European Union funded project High Performance Soft Tissue Navigation (HiPerNav). The objective of HiPerNav is to address the existing bottlenecks associated with laparoscopic surgery of liver tumors. In order to achieve this substantial goal, the project consists of several procedures such as image enhancement, segmentation of structures in the images followed by bio-mechanical modeling, quality assessment of the intermediate procedures and clinical validation of results. The first milestone in the project is medical image enhancement. This thesis is an attempt to particularly investigate the role of incorporating cross-modal guidance information in the enhancement process. The applied research work proposing new and effective solutions to meet this need is presented to fulfill the requirement of the doctoral degree at the Department of Computer Science, NTNU Gjøvik.

1.2 Research Aims

The main objective of this research is to investigate and develop efficient solutions to improve the quality of medical images. Medical images are often degraded by undesirable phenomena such as noise, low contrast and other modality specific artifacts. Therefore, it is imperative to enhance the images before using them in diagnosis and further processing. Techniques that leverage complementary information in the better perceptual quality multi-modal images can play significant role in achieving this objective. Multi-modal images such as CT, T1-weighted (T1-weighted (T1-w)) and T2-weighted (T2-w) MRI have been used in this work. We have investigated the following kinds of enhancement in this thesis:

- **Contrast Enhancement**: In the context of image guided surgeries, the role of image enhancement is even more valuable since it is the first operation before any subsequent process is applied. Therefore, all the subsequent tasks rely on the outcomes of enhancement. CT images are often the first choice to detect existence of any structural abnormality in most cases. Usually,

MR images are also acquired in the later stages of diagnosis and therapy planning. This work particularly focuses on exploiting the better perceptual quality of MR images to improve the contrast of CT images. In case of medical images, contrast enhancement (Contrast Enhancement (CE)) is usually done to achieve different objectives from natural images. The main objective of this project is to develop contrast enhancement solutions and to prove its importance in different tasks in the context of liver cancer diagnosis and treatment. The tasks directly dependent on image quality and targeted through this study are:

1. Visibility of organ's anatomy
2. Segmentation of tumors

- **Denoising:** Medical images are often contaminated with various kinds of noise during acquisition. This includes Gaussian noise, Poisson noise, Rician noise etc. The noise affects the visibility of important structures in medical images and also pose challenge for the computer vision algorithms. Indeed, any denoising operation using conventional low-pass filtering methods tends to affect other important fine structures in the image as well. The denoising problem is therefore very delicate and it is necessary to think of solutions that can guide this process and thus avoid negative effects on the useful signal. Image denoising then becomes indispensable to get rid of the noise while simultaneously preserving the significant details in medical images. We will investigate image denoising in this thesis by particularly using the concept of guided filtering given the availability of complementary imaging modalities such as CT and MRI.

1.3 Research Questions

The above mentioned objectives raise the following research questions stated below. This project was undertaken to address these challenging research problems and presented through the papers listed below..

Part 1: Research Questions related to contrast enhancement

Q1.1 Does incorporating guidance information from a different modality image improve the contrast of a medical image for improved visibility ? (addressed in **Papers A, E**).

Q1.2 Investigate if tumor segmentation methods work better on images enhanced using proposed CE method ? (addressed in **Paper B**)

Q1.3 Can the proposed contrast enhancement method be optimized ? (addressed in **Papers C, D**)

Part 2: Research Questions related to denoising

Q2.1 How multi-modal guidance information could be employed for denoising medical images using deep learning ? (addressed in **Paper F**)

1.4 List of Published Papers

This research work produced the following articles. Five of these articles are published and one is accepted for publication.

Paper A Rabia Naseem, Faouzi Alaya Cheikh, Azeddine Beghdadi, Ole Jakob Elle, and Frank Lindseth. "Cross modality guided liver image enhancement of CT using MRI." In 2019 8th European Workshop on Visual Information Processing (EUVIP), pp. 46-51. IEEE, 2019.

Paper B Rabia Naseem, Zohaib Amjad Khan, Nitin Satpute, Azeddine Beghdadi, Faouzi Alaya Cheikh, Joaquín Olivares, "Cross-modality guided contrast enhancement for improved liver tumor image segmentation", IEEE Access, 9(2021), 118154-118167.

Paper C Nitin Satpute, Rabia Naseem, Rafael Palomar, Orestis Zachariadis, Juan Gómez-Luna, Faouzi Alaya Cheikh, and Joaquín Olivares. "Fast parallel vessel segmentation." Computer methods and programs in biomedicine 192 (2020): 105430.

Paper D Nitin Satpute, Rabia Naseem, Egidijus Pelanis, Juan Gómez-Luna, Faouzi Alaya Cheikh, Ole Jakob Elle, and Joaquín Olivares. "GPU acceleration of liver enhancement for tumor segmentation." Computer methods and programs in biomedicine 184 (2020): 105285.

Paper E Rabia Naseem, Akib Jayed Islam, Faouzi Alaya Cheikh, Azeddine Beghdadi. "Contrast Enhancement: Cross-modal Learning Approach for Medical Images", Electronic Imaging 2022 (Accepted as Oral)

Paper F Rabia Naseem, Faouzi Alaya Cheikh, Azeddine Beghdadi, Khan Muhammad, Muhammad Sajjad, "Cross-Modal Guidance assisted Hierarchical Learning based Siamese Network for MR Image Denoising", Electronics, 10(10):2855, 2021

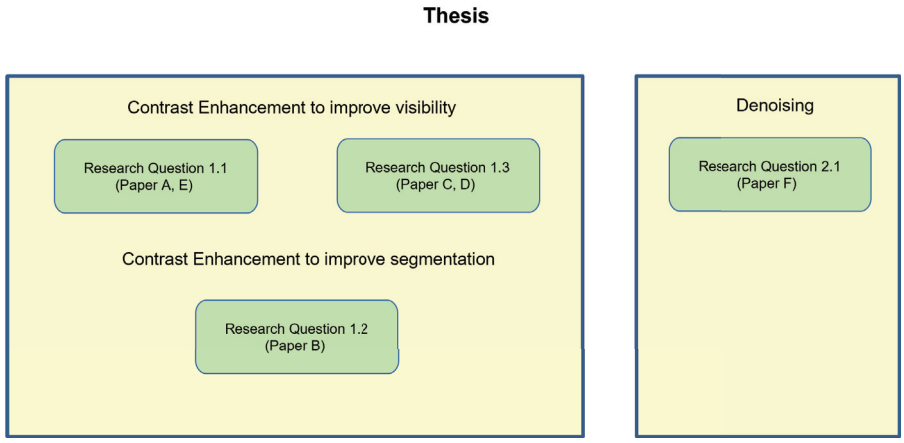


Figure 1.2: Overview of research articles and their relationship with research questions

1.5 Dissertation Structure

This dissertation is organized as a collection of papers combined with the description of how each paper addresses the research question listed in Sec. 1.3. The relationship between research questions and the corresponding papers is demonstrated in figure 1.2. The dissertation is divided into two parts: **Part I** and **Part II**. **Part I** consists of five chapters. The organization of **Part I** of the thesis is as follows:

- Chapter 1 (current chapter) introduces the thesis, describes the research aims and questions addressed in the thesis and lists the original published/ submitted articles.
- Chapter 2 elaborates the background of the research work that will help the readers in understanding the overall context and motivation of carrying out this work. Information regarding laparoscopic liver resection is presented followed by navigation of laparoscopic resection.
- Chapter 3 provides summary of all the articles that constitute the thesis. The objective, proposed methodology, results and contributions are briefly documented. **Paper A** proposes a method to enhance contrast of medical images using cross-modal guidance to emphasize the visibility of the organ details using traditional method. **Paper E** does this task by making use of deep learning approaches. **Paper B** proposes a method to improve contrast of CT images; furthermore, we investigate if tumor segmentation algorithm works

better on contrast enhanced images and addresses research question **Q1.2**. **Papers C and D** address research question **Q1.3**. Finally, **Paper F** proposes deep learning based denoising approach using cross-modal guidance and addresses **Q2.1**.

- Chapter 4 highlights the contributions and limitations of the thesis, few suggestions for future work are given.
- Conclusion is presented in chapter 5.

Original articles produced during the course of this project are appended in the end of dissertation and constitute **Part II** of the thesis.

Chapter 2

Background and Context

This chapter presents the broader context of the problem addressed in this thesis. Brief background knowledge regarding the topic is elaborated. The articles included in this dissertation are closely related to the topics elaborated in this chapter and the link is pointed explicitly. The chapter is structured as follows.

First, introduction to laparoscopic tumor resection is presented in Sec. 2.1 followed by the description of the general workflow of surgical navigation in Sec. 2.2; it consists of pre-operative and intra-operative phases. The most widely used medical imaging modalities are described in Sec. 2.3. Few traditional approaches for contrast enhancement and denoising are explained respectively in Sec. 2.4 and 2.5. The motivation of using cross-modal guidance in image processing is discussed in Sec. 2.6.2. Few neural network architectures used in this work are briefly described (Sec. 2.7.1) which concludes the chapter.

2.1 Laparoscopic Liver Tumor Resection

Liver cancer ranks fifth frequent cancer globally among men with a high death rate [11]. It can be either primary (developed in the liver) or secondary (originated in another organ and later spread to the liver). Liver resection is a high-risk procedure to remove part of the liver containing tumor by adept surgeons and is deemed as an effective remedy for liver cancer patients [22]. The resection is recommended for patients even in the later stages of liver cancer to prevent further progress of the disease. With improved apprehension of liver anatomy, technological advances in curative surgery, the survival rate among liver cancer patients continue to improve during the last few years [23]. Liver resection, also called hepatic resection can be performed as either open surgery or laparoscopic surgery. Figure 2.1 demonstrates

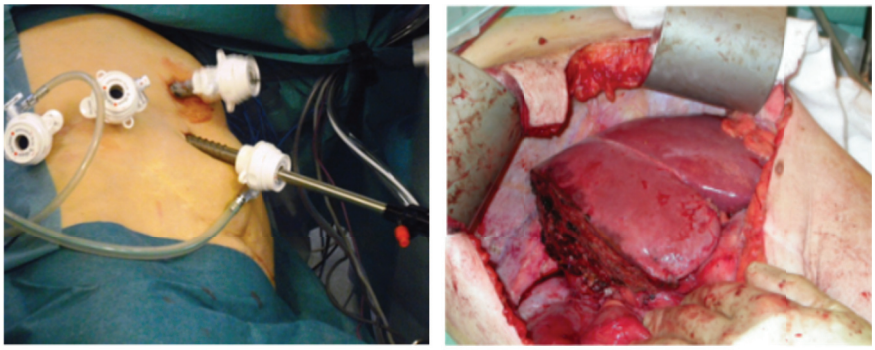


Figure 2.1: Demonstration of (a) laparoscopic surgery and (b) open surgery, Images: courtesy of HiPerNav partners from Oslo University Hospital

the generic setup of open surgery and laparoscopic surgery. The traditional open surgery is undertaken by making a wide incision into the patient's body; the whole surgery is then performed with direct access of the surgeon to the surgical field. In the latter approach, also regarded as minimally invasive surgery, the surgery takes place by using surgical instruments that are inserted into the patient's abdomen via small incisions. While the traditional option for resection has been open surgery, laparoscopy surgery is replacing its open surgery counterpart.

Laparoscopic surgery begins with the creation of pneumoperitoneum when carbon dioxide gas is insufflated into the abdomen. Pneumoperitoneum establishes adequate space to visualize the target organ using a laparoscopic camera. The surgical instruments are then passed into the abdomen. Laparoscopic surgery has been substantially applied to various abdominal organs including the liver and colon during recent years [24] accompanied by reduced blood loss, shorter hospital stay, and less post-operative care [25]. However, these benefits also bring some challenges associated with laparoscopic surgery. As opposed to open surgery, laparoscopic surgery prevents a direct view of the organ. The camera inserted through the abdomen captures a smaller field of view of the organ being operated. Moreover, tissue manipulation using instruments without tactile feedback and strenuous hand-eye coordination are some challenges that necessitate substantial experience for surgeons performing these surgeries [26].

Commercial systems for liver surgery navigation also exist. IQQA ®Liver (EDDA Technology, Inc.) provides comprehensive toolset to evaluate the automatic volumetric quantification of liver, liver segments, hepatic lesions, and vascular structures from CT data. CAS-ONE Surgery and CAS-ONE IR (CAScination, Bern,

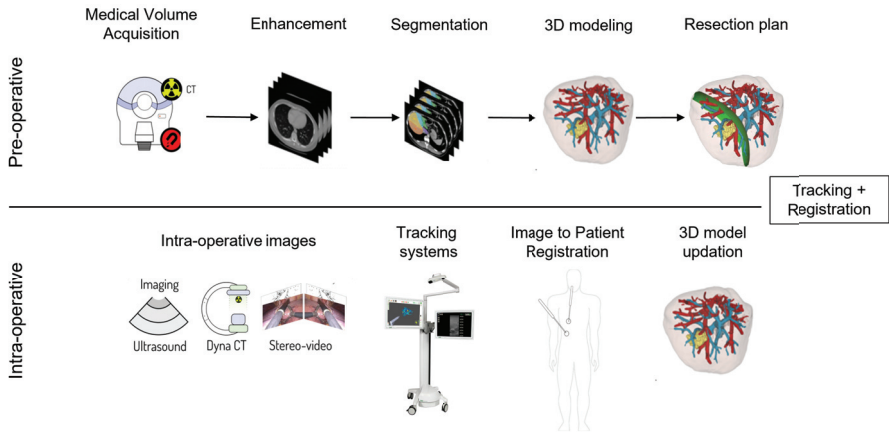


Figure 2.2: Generic workflow of surgical navigation: pre-opearative and intra-operative planning, Images: courtesy of HiPerNav partners from Oslo University Hospital

Switzerland), the navigation systems designed for ablation of tumors found in liver, lung etc. Furthermore, many open-source platforms have been developed such as CustusX [27]; it can localize tumor via navigated Ultrasound during surgery. NorMIT (a collaboration between two hospitals of Norway) offers another platform for image-guided interventions.

The above-stated challenges posed to laparoscopic surgery prompt the development and improvement of computer-assisted systems to facilitate surgical planning and operating environment. In the subsequent section, we present an overview of the surgical navigation systems.

2.2 Treatment Planning and Surgical Navigation

Surgical navigation is a means of using diagnostic images together with positional tracking devices to guide surgeons in carrying out surgical interventions effectively. This setup tracks the location of the surgical equipment through cross-sectional medical images or Three dimensional (3D) anatomical models reconstructed from these images [28].

The first attempt of exploiting imaging techniques in the guidance of surgical interventions dates back to 1985 when x-ray imaging was utilized to identify the location of a sewing needle in a woman's hand [29]. Further advancements in diagnostic imaging methods and allied computer-based technologies are paving the way for the rapid development of surgical navigation systems. This development in navigation has also benefited laparoscopic liver resection. The generic work-

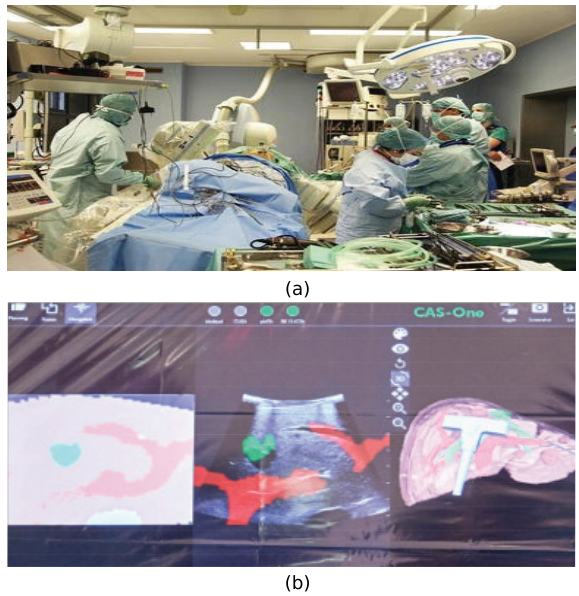


Figure 2.3: Demonstration of intra-operative navigation (a) intra-operative imaging data acquisition using robotic cone-beam CT [1], (b) alignment of pre-operative CT data (left), intra-operative US (middle) and 3D model (right) (CAScination, Bern, Switzerland)

flow of a surgical navigation system can be divided into two phases: pre-operative phase and intra-operative phase. The details of each phase are shown in figure 2.2.

The **pre-operative** phase consists of all the procedures done before the commencement of surgery. This phase starts with image acquisition; CT and MRI are the two modalities commonly acquired during this stage. Different anatomical structures in the liver including liver parenchyma, hepatic and portal veins, and tumors are segmented afterward. 3D modeling techniques combined with segmented structures contain all the information required to do surgical planning where a virtual resection is defined [30]. The virtual resection helps surgeons in visualizing the path that separates the cancerous tissues from the residual healthy liver parenchyma. Intuitively, the geometric information contained in the 3D models and resection plan serves as crucial information helping surgeons prior to and during surgery. During the **intra-operative** stage, surgeons use the pre-operative models as reference during surgery. In this framework, surgeons have to cognitively align the intra-operative surgical reality with the pre-operative models. However, the liver is a complicated organ comparatively prone to tissue deformations and organ shift once surgery starts. Pneumoperitoneum (insufflation of the abdomen) and sur-

gical treatment can cause the organ shift of up to 28 mm [31] and is, therefore, liable to create inconsistency between pre-operative and intro-operative fields. This challenging scenario necessitates intra-operative imaging. Teatini et al. [32] conducted a study to analyze the impact of intra-operative imaging in surgical navigation when displayed as Augmented Reality. It was experimentally validated that intra-operative imaging plays a significant role in compensating the deformations. Therefore, navigation during the intra-operative stage is crucial and several studies have been conducted to further advance in this field.

Intra-operative navigation relies on intra-operative data to update the pre-operative models; furthermore, this data guides surgeons throughout the surgery. This guidance can possibly be a) information regarding tracking instruments and overlaying them on the 3D model. b) delivering accurate localization information of the organ's anatomy obtained by intra-operative imaging corresponding to the 3D model and c) making use of augmented reality during the intervention.

It is pertinent to note that the role of imaging techniques becomes critically important at this stage since intra-operative imaging data is an important source of information for the subsequent processes of navigation. Figure 2.3 shows the pre-operative CT registered with intra-operative ultrasound to locate the tumor being resected. It should be noted that during all these phases the quality of the images and videos is of paramount importance and should be given full attention for the success of the whole workflow. This thesis focuses on the initial phase of the navigation, that is pre-operative planning phases. Next, we describe most commonly used medical imaging modalities.

2.3 Cross-modal Imaging for Treatment Planning

Diagnostic imaging can be generally categorized into two branches: structural and functional imaging. The former is used to determine the morphology of organ and to analyze changes in the structure. Computed Tomography, Magnetic Resonance Imaging and Ultrasound are among the most widely used structural imaging techniques, while Positron emission tomography is the most popular functional imaging method, particularly used in cancer treatment.

2.3.1 Computed Tomography

Computed Tomography combines the Two dimensional (2D) projection imaging of radiography and fluoroscopy to generate a cross-sectional view of the human body. CT acquires multiple projections from arbitrary orientations to reconstruct the object of interest, which facilitates volumetric measurements. CT enables discrimination among tissues having similar density. The first CT scanner was used at a hospital in London in 1971 developed by Hounsfield and Cormack [33]. Ex-

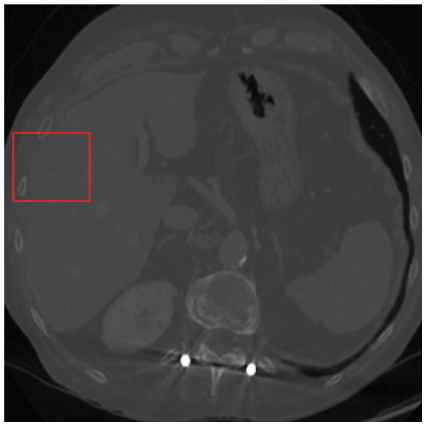


Figure 2.4: CT - liver

tensive research by the medical community has led to the development of variants of CT scanners such as helical CT, cone-beam CT. C-arm Cone-Beam CT (CBCT) is an innovative technology that does not require relocation of a patient and can be integrated into the intervention suite. It is capable of acquiring the whole volumetric data covering a large anatomical area of interest in a single gantry rotation exploiting a 2D detector system, unlike conventional CT. The acquisition time of CT scanners is fairly short; however, since CT scans use x-rays and carry ionizing radiation, they can be potentially harmful to certain patients such as pregnant women. Moreover, there are some limitations such as low contrast which necessitates the use of other modalities such as MRI in certain cases. CT scan of liver is shown in figure 2.4; it can be noticed that the tumor inside liver cannot be seen clearly due to low contrast of the image.

2.3.2 Magnetic Resonance Imaging

MRI is widely used in diagnosis concerned with the organs such as heart, brain, and liver. Several imaging protocols render MRI to highlight and therefore visualize a variety of tissue types. By varying image sequence parameters, soft tissue contrast can be altered, which eventually aids in examining the margin of tumors.

MRI is based on Nuclear Magnetic Resonance. It uses magnetization properties of certain atomic nuclei in the human body (the most common is hydrogen) [34]. These nuclei spin in a randomly aligned axis. MRI scanner creates a strong external magnetic field, which forces nuclei in the human body to align themselves with this field. Radio Frequency (RF) pulse is then brought orthogonal with the magnetic field to excite the nuclei, which induces another magnetic field. The nuclei start oscillating longitudinally to the external magnetic field, which is termed

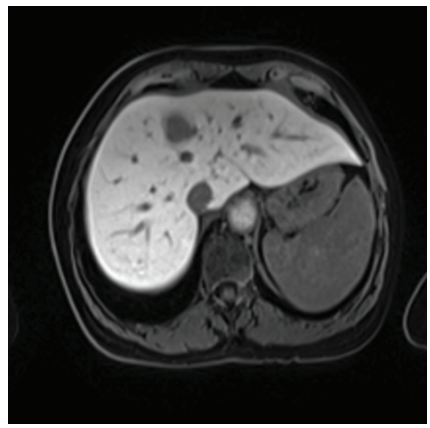


Figure 2.5: MRI - liver

'relaxation'. During this realignment, RF waves are emitted that can be detected by sensors. The amount of energy released and the time it takes for nuclei to re-align with magnetic field depends on the static magnetic field, applied excitation, and nature of molecules. Faster proton realignment will create a brighter image. Magnetic properties can be used to tell difference between various types of tissues in the human organ, which establish contrast in MR images. Liver MRI is shown in figure 2.5. Two kinds of relaxation can be distinguished: T1-Relaxation and T2-Relaxation, which are briefly explained below:

Magnetic Resonance Imaging-T1

T1-Relaxation also called longitudinal relaxation is the procedure of restoring magnetization longitudinal to the external magnetic field after RF pulse is applied [34]. This restoration is described by an exponential function that is subject to time constant T1. This constant quantifies the time needed to recover 63% of initial magnetization; besides, this time is distinct for each kind of tissue. This process is termed 'spin lattice relaxation'. Figure 2.6 shows T1-w MR images of the human brain (axial view); blood and fat appear bright whereas Cerebral Spinal Fluid (CSF) is dark.

Magnetic Resonance Imaging-T2

T2 Relaxation, also called transverse relaxation or spin-spin relaxation quantifies the decay of transversal magnetization since RF pulse is applied. Quantified by time constant T2, it represents the time until the initial signal caused by the transverse magnetization drops to 37% of its primary magnitude [34]. T2 values also depend on the tissue type akin to T1. The clinical use of T2 relaxation is in T2-

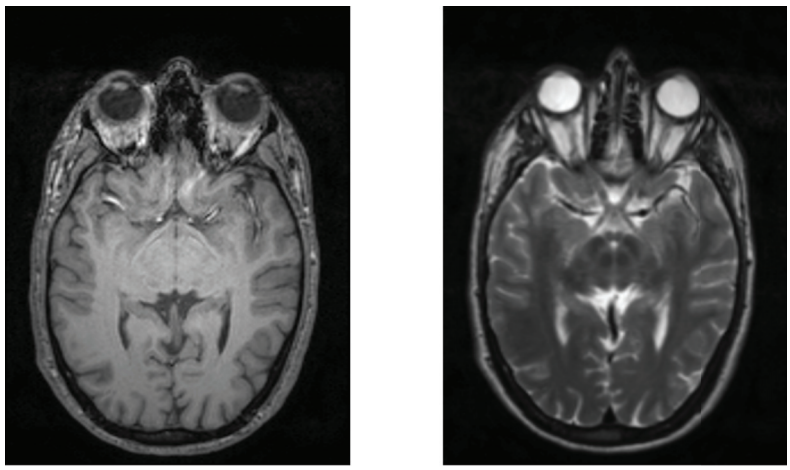


Figure 2.6: Brain MRI a) T1-w b) T2-w

weighted MRI. Body tissues higher in water content appear bright on T2-w MRI. T2-w MRI of human brain is shown in figure 2.6.

2.3.3 Ultrasound Imaging

Ultrasound (US) is used in the clinical suite for several decades. Dr. Karl Theo Dussik applied it for the first time in diagnosis of brain related disorder [35]. US is preferred over other imaging modalities primarily due to its portable and non-ionizing nature, rendering it favorable as an effective intra-operative imaging tool. Furthermore, US also enables doctors to examine the cross-section of an organ. The ultrasound operating frequency in the context of medical imaging ranges between 2 MHz and 40 MHz [36]. The basic principle of medical US image acquisition is based on the pulse-echo technique. The transducer transmits pulses of ultrasound echo to the human body. Meanwhile, few pulses penetrate farther into the body tissues of various acoustic impedances and few are reflected from the tissues back to the transducer. The combination of signals returned from the series of pulses forms an image. US is frequently used in the examination of soft tissues, vessels, and fluids accumulated in organs. Figure 2.7 shows ultrasound of human liver.

After describing the various types of commonly used imaging modalities in surgical navigation, we describe in Sec. 2.6 the basics of guided filtering and how it is related to this thesis. However, we restricted our focus on CT and MRI in this thesis; the role of ultrasound in CE and denoising was not explored due to the

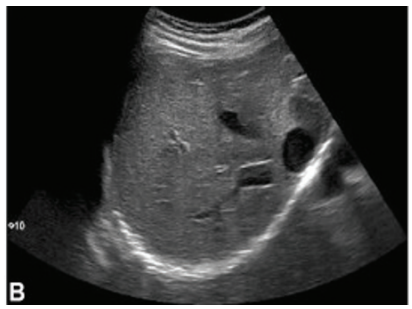


Figure 2.7: Ultrasound - liver

unavailability of US data.

2.4 Contrast Enhancement

Images are subject to several kinds of degradation during acquisition due to varying lighting conditions, sensor resolution, the noise inherent to specific acquisition systems, and some specific physical factors that impact the performance of the available imaging modalities. Considering medical images, other factors such as patient movement could also lead to artifacts that necessitate image enhancement. Generally, enhancement is done to improve contrast and brightness, minimize noise, and improve resolution. The ultimate goal of this process is to make these degraded images visually better for perception. Image enhancement is a broad topic; we summarize in the figure 2.8 the widely researched subdivisions of image enhancement. Among different areas of image enhancement, contrast enhancement is the process of amplifying the intensity difference among pixels in an image to make its interpretation easier and meaningful for humans and computer algorithms. Human eyes are more sensitive in detecting large intensity variations in an image than small variations and thus well perceive the details corresponding to greater intensity differences. The visibility of details also depends on the context in which the structures are observed and in particular their relative spatial frequency sensitivity. Indeed, the visual system acts roughly as a directional bandpass filter. It is therefore important to take this into account implicitly or explicitly in the development of image processing tools and particularly in contrast enhancement.

Contrast enhancement is a fundamental problem in image processing and is a subjective process since the quality of enhanced images is determined distinctly by the human visual system. In medical images, the objective of CE is generally to improve the visibility of low contrast images to facilitate the radiologists in the interpretation of the diagnostic images. Secondly, the medical images are used as

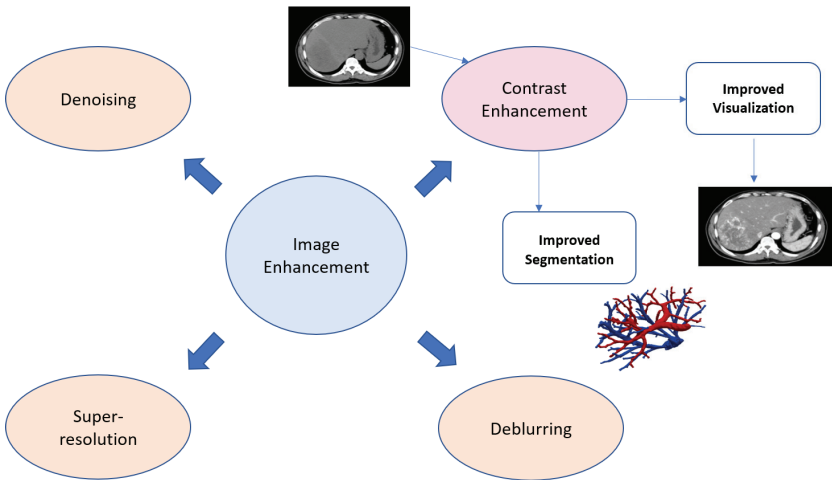


Figure 2.8: Subdivision of Image Enhancement

input to several tasks such as segmentation, classification, registration; therefore, contrast enhancement is an important pre-processing step applied to improve the performance of the underlying algorithms. In this thesis, the contrast of CT images is improved to primarily achieve these twofold objectives.

2.4.1 Histogram-based Approaches

Numerous techniques have been presented to enhance the contrast of images [6, 37, 38, 39]. Contrast enhancement can be done globally or locally, where global techniques consider the whole image (statistics) as opposed to local approaches that operate in a small neighborhood of an underlying pixel [6]. Histogram Equalization (Histogram Equalization (HE)) is one of the most extensively used approaches for image enhancement. It remaps intensity values in the image so the enhanced image has a uniform pixel distribution. Similarly, Histogram Specification (Histogram Specification (HS)) is another approach widely used for contrast enhancement that takes an image of better perceptual quality/ rich contrast where the goal is to remap the pixel values in the input image to that of target image using its histogram as reference [40].

HE suffers from an intrinsic limitation, it tends to over-enhance the image, leading to unnatural appearance and artifacts. These visual flaws were investigated by various researchers and numerous methods were presented to rectify these artifacts [41, 42]. Brightness preserving bi-histogram equalization (BBHE) is one such method that applies HE independently on two sub-images after dividing the original image based on its mean intensity value, one in the range of minimum gray

value to mean and the other from mean to maximum gray value [43]. The proposed method preserves the brightness of the image during enhancement. This idea was followed by another approach Equal Area Dualistic Sub-Image Histogram Equalization [44], that bifurcates the histogram based on the gray level with cumulative distribution function (Cumulative Distribution Function (CDF)) equal to 0.5. Both the approaches produce visually pleasing images compared to HE when applied to very bright or dark images. However, the issue of artifacts somehow still prevails when spikes exist in the histogram.

Adaptive Histogram Equalization (AHE) divides an image into small distinct blocks and then histogram equalization is applied to each block separately. However, it is prone to amplify noise particularly in rather homogeneous areas of the image. Contrast-limited adaptive histogram equalization (Contrast-limited adaptive histogram equalization (CLAHE)) [45] is a variant of AHE and a method to counter the over-amplification phenomenon in the AHE. It does so by limiting the amplification. Bilinear interpolation is applied at the edges of tiles to combine the results.

2.4.2 Context-Aware Histogram-based Approaches

The inherent saturation and over-enhancement phenomenon associated with histogram-based approaches have been improved by several schemes including the ones mentioned above. However, one aspect that should be highlighted here is that mere use of first-order statistics, namely, one-dimensional histogram does not provide adequate information regarding the spatial structural information in the image. This essential issue prevents these approaches to eradicate artifacts. One potential solution in this perspective is to incorporate second-order statistics, i.e., the spatial information of an image, in devising a more robust image enhancement technique. The 2D histogram is one of the ways to incorporate contextual information as well while applying histogram-based operations. Let us consider an input image $[Z] = \{Z(m,n)|1 \leq m \leq M, 1 \leq n \leq N\}$, where $Z(m,n)$ is the grey-level of pixel (m,n) taking its values within the dynamic range $[Z_{min}, Z_{max}]$ and M×N is its size in pixels. One of the ways to express 2D histogram is in terms of Gray-level Co-occurrence Matrix. Its mathematical expression is given as:

$$C_Z(i,j) = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} \delta_{ij}(Z(m,n), Z(p,q)), \tag{2.1}$$

Here, i and j represent the pixel values and (m,n) and (p,q) represent the image coordinates, K is the total number of grey levels, and $0 \leq i,j \leq K-1$,

$$\delta_{ij}(a,b) = \begin{cases} 1, & \text{if } i = a \text{ and } j = b \\ 0, & \text{otherwise} \end{cases}$$

The transition probability of grey-levels, i.e. the 2D normalized histogram, is derived from the GLCM as follows:

$$h_Z(i,j) = \frac{C_Z(i,j)}{\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} C_Z(i,j)} \tag{2.2}$$

The 2D-histogram is then used in the pixel grey-level mapping process using the histogram specification method (**Paper A**). This mapping process is based on the two-dimensional CDF of the input and reference images (**Papers A and B**). Since Two-dimensional Histogram Specification (2D-HS) works in context-aware fashion, it performs better than its context-free (1D) counterpart.

Applying 2D-HS enhances the image contrast but the resultant image loses structural similarity with the original image. This limitation phenomenon can be controlled by including some criteria to maintain this structural similarity with the original image. The strategy proposed in this thesis to counter this effect is the use of structural affinity related measure, namely Structural Similarity Index (Structural Similarity Index (SSIM)) gradient [46] (the mathematical expression is given in **Paper B**). SSIM is a well established index to compute the degree of similarity between two images [47]. Considering one image as reference, the index provides the quality of underlying image in comparison with a reference. SSIM index is calculated between corresponding local blocks in images $[Z]$ and its enhanced variant $[Z']$, after which the average of the values is taken to obtain a single value of SSIM as the overall similarity index. Let us assume that z_x and z'_x represent corresponding block x in both images; the SSIM between the two blocks z_x and z'_x is then expressed as:

$$\mathrm{SSIM}(z_x, z'_x) = \frac{\left(2\mu_{z_x}\mu_{z'_x} + C_1\right)\left(2\sigma_{z_x z'_x} + C_2\right)}{\left(\mu^2_{z_x} + \mu^2_{z'_x} + C_1\right)\left(\sigma^2_{z_x} + \sigma^2_{z'_x} + C_2\right)} \tag{2.3}$$

μ_{z_x} and $\mu_{z'_x}$ represent the mean intensity values of z_x and z'_x and σ_{z_x} and $\sigma_{z'_x}$ are the corresponding standard deviations. C_1 and C_2 are small numbers greater than 0 to ensure denominator is not zero. The key terms in equation 2.3 are described mathematically as:

$$\begin{aligned} \mu_{z_x} &= w * z_x, \\ \sigma_{z_x z'_x} &= w * (z_x z'_x) - \mu_{z_x}\mu_{z'_x}, \\ \sigma^2_{z_x} &= w * z_x^2 - \mu^2_{z_x} \end{aligned} \tag{2.4}$$

where w is 11×11 Gaussian kernel and $*$ indicates convolution.

Contrast enhancement is a highly subjective process. However, it is important to determine the extent of desired enhancement. A goal-oriented contrast enhancement scheme is proposed in the **Paper B**, where the goal is to enhance the CT image so that tumors in the image are discriminated from the rest of the organ, and segmentation algorithms, therefore, work better. An approach proposed in **Paper B** formulates one such criteria.

2.5 Noise in Medical Images

Noise can be conceived as a random variation of the actual pixel value. Noise is introduced in the images as a consequence of several physical procedures occurring during image acquisition. It not only affects the image quality but also the results of several analysis tasks such as feature extraction, segmentation, detection [48, 49]. Several studies report that the results of segmentation and edge detection in medical images show noticeable improvements when the algorithm is applied to denoised images as opposed to noisy image [50, 51].

Image denoising is therefore an essential operation that make images suitable for analysis and further processing. Image denoising entails estimating an unknown noise-free image *Z* given a noisy observation *Y*. Using the classic additive image degradation model:

$$Y = Z + N \tag{2.5}$$

where image *Z* is contaminated by additive white gaussian noise *N* having a variance σ^2. Gaussian is considered as one of the common noise models, where the noise is normally distributed over a range of values. Modeling gaussian noise is important since it is a nice approximation to various kinds of noise. Gaussian additive noise can be described through normal distribution as:

$$p(z) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(z-\mu)^2}{2\sigma^2}} \tag{2.6}$$

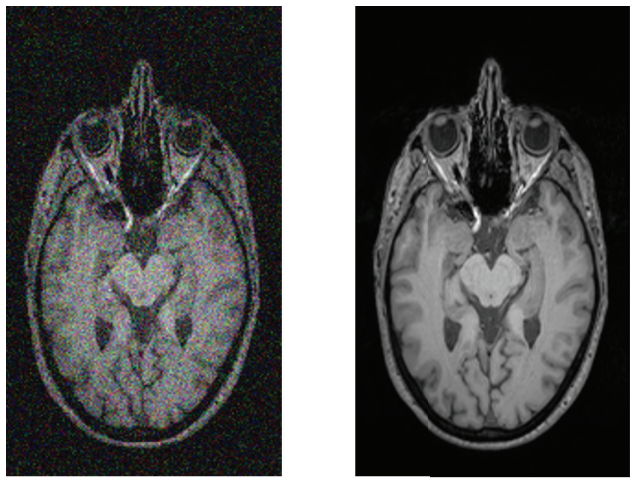


Figure 2.9: Noise in MRI: (a) Noisy (b) Non-noisy

where *p(z)* is a probability density function, z indicates image's gray level, μ indicates mean, while σ and σ^2 are the standard deviation and variance respectively.

Note that MR images are subject to Rician noise [52, 53]. This noise associated with the motion of charged particles in the radio-frequency coils is introduced during MRI acquisition. Assuming noise in Z is Rician distributed, we can express its PDF as:

$$p\left(N \mid Z, \sigma\right) = \frac{N}{\sigma^2} \exp^{-\frac{\left(N^2+Z^2\right)}{2\sigma^2}} J_0\left(\frac{ZN}{\sigma^2}\right) \epsilon(N) \tag{2.7}$$

In the above equation, J_0 represents the 0^{th} order Bessel function, N is a Rician distributed random variable. $\epsilon(.)$ is the unit step Heaviside step function implying that the PDF expression is valid for non-negative values of N. The Rician noise is signal-dependent and shows Gaussian distribution if Signal-to-Noise Ratio (SNR) is high and Rayleigh distribution if SNR is low [54]. Figure 2.9 shows T1-w brain MRI contaminated by Rician noise and the corresponding image without noise. It is certain that noise impacts the visibility of important information in the image.

Confiscating Rician noise is a difficult task because of its signal-dependent characteristic. However, there exist several denoising approaches in the literature [55, 56]. Few denoising approaches are briefly summarized below.

2.5.1 Denoising

A considerable number of spatial filters have been presented in the past to denoise images, which can roughly be categorized as linear and non-linear filters. Primarily, linear filters such as mean filter [57] based on local averaging were designed to eradicate noise, however, they over-smooth the image. Several non-linear filters were proposed to better preserve the edges while denoising [58, 59]. Anisotropic filter [60, 61], for instance, was able to bypass the blur implicit to mean filter; it smoothes the image in a direction orthogonal to the gradient direction. Bilateral filter [62] is another edge-preserving smoothing filter that replaces each pixel by the weighted average of nearby pixels but based on their geometric proximity and variation of intensities to preserve edges. Unlike local smoothing filters, non-local means filter [7, 63] performs particularly well in denoising images when the image contains various regions of different types of textures. They compute the weighted average of not only the local neighborhood but also all the pixels in an image; this weightage depends on the similarity between the target pixel and other pixels in the image. Other approaches to noise filtering by analysing the signal in a multi-dimensional space have been proposed in [64, 65].

Wavelet denoising schemes generally decompose the image into various wavelet coefficients using the discrete wavelet transform followed by the application of hard thresholding and soft thresholding [66] on the detail coefficients. Hard thresholding methods shrink the wavelet coefficient to zero lying in a certain range, while those beyond the specified range remain unaltered. Contrarily, in soft thresholding, the coefficients in the specified range are contracted to zero like hard thresholding but those beyond the specified range are contracted down by a constant value. The wavelet-based denoising methods preserve sharp edges better compared to spatial domain methods. The latest trend in denoising approaches is based on deep learning methods which are briefly summarized in Sec. 2.7.

2.6 Guided Filtering

Guided filter was introduced a decade ago as an edge-preserving smoothing filter [67]. The filter was proposed to address some of the limitations such as over-smoothing of edges associated with filters such as bilateral filter [64]. The filter uses the same image or another image to transfer the characteristics of guidance image to the input image. In many approaches presented later, the guidance image is an image of better perceptual quality while the input image is either a noisy, low-contrast or an image containing some artefacts. Let us consider an input image Z and a guidance image G, image F is the filtered image. The pixel value at index i in image F is then calculated as:

$$F_i = \sum_j W_{ij}(G) Z_j \tag{2.8}$$

i and *j* in above equation represent pixel indices. The filter kernel W_{ij} is a function of *G*. It is assumed that *F* is a linear transform of *G* in a window ω_c centered at pixel c. The value of output pixel at location *i* depends on the value of guidance image's pixel at *i* and two constants *p* and *q* (whose value is assumed to be constant over the entire window ω_c) and is expressed as:

$$F_i = p_c G_i + q_c, \forall i \in \omega_c \tag{2.9}$$

A constraint is introduced to force the filtered output image to be close to the input image *Z*. Another regularization term is included to enforce the degree to which the characteristics of guidance image *G* should be embedded into the output image *F*. The following equation represents this constraint:

$$E\left(p_c, q_c\right) = \sum_{i \in \omega_c} \left(\left(p_c G_i + q_c - Z_i\right)^2 + \epsilon p_c^2 \right) \tag{2.10}$$

The value of constants *p* and *q* is calculated based on both *G* and *Z* as follows:

$$p_c = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_c} G_i Z_i - \mu_c \bar{Z}_c}{\sigma_c^2 + \epsilon} \tag{2.11}$$

where μ_c and σ_c^2 represent the mean and variance of G enclosed in ω_c, and ϵ is the parameter selected for regularization. $\bar{Z}_c$ is mean of Z in ω_c). The constant q is simple to calculate and essentially depends on the mean of input image values and guidance image values contained in ω_c.

$$q_c = \bar{Z}_c - p_c \mu_c \tag{2.12}$$

The basic assumption of a guided filter is a local linear model between the guidance image and the filtered image. The model assumes that the input and guidance images are accurately registered (aligned). Guided filter became very popular and has been applied to problems such as denoising, dehazing, super-resolution [68] and contrast enhancement [69].

Figure 2.10: Cross-modal denoising [2]: (a) RGB (b) NIR (c) Scale map

2.6.1 Cross-modality guided enhancement - Natural Images

Inspired by the idea of exploiting supplementary information in another image to enhance the image with some sort of artifacts (similar to guided filtering), image enhancement has been applied to a variety of sub-tasks such as denoising [2], contrast enhancement [70], super-resolution [71] and de-blurring [72] to name a few. The idea was also extended to multi-modal images. Information from a cross-modal image carrying complementary information (and or better perceptual quality) was integrated in the cross-modality guided denoising and enhancement approaches [2, 73]. These approaches in general outperformed similar approaches that do not include the cross-modal information.

The cross-modality guided enhancement techniques applied to natural images include spatial domain, wavelet domain [70], and variational methods [2]. One such work exploited the rich details in the Near Infra-Red (NIR) image to enhance the corresponding RGB images by transferring contrast and detail using Haar wavelet [74]. Similarly, using a pair of NIR flash images and noisy RGB image, denoising and detail enhancement of the RGB images was done using an edge-preserving based weighted least square optimization approach [70]. Yan et al. [2] presented a denoising approach to improve some limitations of the traditional cross-modal denoising methods [70, 74], such as the failure to handle gradient divergence between the two modalities and the inability to retain sharp edges during restoration. They constructed the scale-map, obtained from the corresponding gradients in guidance and input images. An objective function was then formulated to update the initial scale map which was then used to denoise the RGB image until the convergence is achieved (in about 4 to 6 iterations). NIR image was used as guidance image in their approach. The input and guidance image along with the scale-map produced during restoration in the approach of [2] are shown in figure 2.10.

Conventionally, the cross-modal denoising or enhancement schemes apply registration to align the input and guidance images [2]. This step significantly eases further processing making the application of sophisticated local operations possible. Since these approaches incorporate additional information via cross-modal guidance in the enhancement process, they can retain smaller details in the restored image and better handle the gradient reversal effects. For instance, Yan et al. [2] report the comparison between a well-known single image denoising algorithm BM3D [9] and their proposed cross-modal denoising scheme, where it was demonstrated that [2] yields sharp edges and improved restoration in comparison with BM3D. However, these methods sometimes fail when the guidance image contains shadow or highlight effects that leads to transfer of wrong structures in the restored image. In this scenario, deep learning provides much more robust solutions due to efficient feature learning. The deep learning-based methods are discussed in Sec. 2.7.

2.6.2 Cross modality guided Medical Image Enhancement

Multimodal medical images are being heavily acquired in the clinical suite since each imaging modality brings distinct value to medical image analysis. Computed Tomography (CT) imaging is preferred by radiologists for analyzing hard structures in the human body, for instance, bone fractures and bone tumors [75]. MRI is considered ideal for showing different kinds of soft structures such as tumors and lesions [76]. Ultrasound is widely preferred as an intra-operative imaging modality because of its non-ionizing nature and quick acquisition [77]. Motivated by the performance of cross-modality guided techniques in the context of natural images, similar approaches can be applied to medical image analysis as well. However, very limited research has been done in this area such as the approaches presented by [56] and [78]. This leaves a huge gap for further exploration of this area.

In the next subsections, we elaborate the motivation of employing cross-modal guidance approaches in medical image analysis as well as challenges associated with this area.

Motivation

The trend of using multiple medical imaging modalities for the same target problem is growing, thanks to the availability of multiple imaging methods and technical progress. CT-PET and MRI-PET scans are acquired simultaneously during oncology procedures [79, 80]. PET scans show metabolic changes in tissues, while CT and MRI carry anatomical information; both play a significant role in analyzing tumor profile and therapy planning. Similarly, CT and MRI are acquired for the detection and prognosis of certain disorders. Multiple pulse sequences allow

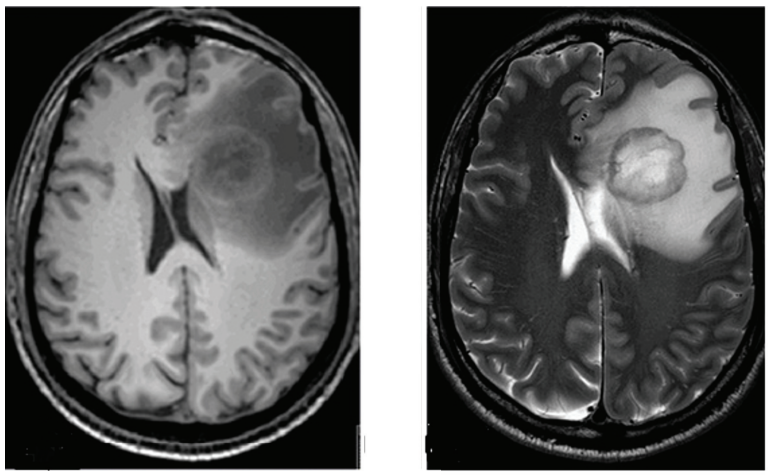


Figure 2.11: Brain MRI a) T1-w b) T2-w

MRI to show several tissues in varying contrasts. T1-w MRI is preferred in the quantification of atrophy [81], an irrevocable loss of neurons linked with multiple sclerosis. T2-w MRI manifests 'substantia nigra', (a portion of the brain affected due to Parkinson's disease) clearly compared to T1-w MRI [82]. The existence of specific pathologies characterized by miscellaneous topography needs to be confirmed using dual modalities. The accurate characterization of 'Craniopharyngiomas (CPs)', a certain type of brain tumor aids surgeons in analyzing surgical risks linked with a specific type. A cohort study including 200 surgically treated CPs inferred that the radiological variables found on both T1 and T2-w MRI anticipated the topography accurately in 86% instances [83]. This correlation was confirmed between the radiological findings and the actual CP topography discerned during surgery. Another study endorses analysis of both modalities to avoid ambiguous hypotheses regarding brain structures [84]. Comparison of T1-w and T2-w brain MRI is shown in figure 2.11 which shows several common structures.

It is important to point out that diverse imaging modalities encompass complementary information that can be particularly influential in computer-assisted diagnosis. This complementarity-aware redundant information has proven its efficiency in tasks such as segmentation, classification, and denoising [85]. The concept of 'weak learnability' also encourages its application in medical imaging. In the context of ensemble learning, the notion of 'weak learnability' states that a learner can be incorporated into the learning system to augment its performance if it can perform slightly better than random guessing [86]. Based on this notion,

a study compared the segmentation outcomes achieved using a single MR image against multi-modal MRI. The segmentation accuracy increased significantly when multi-modal MR images were used. Another study combined PET, MRI, and CT in medical image segmentation [86], it was concluded that combining multiple modalities yielded better performance compared to that obtained using a single modality. A similar deep learning-based approach segmented tumor volume in multi-modal PET/ CT images for head and neck cancer patients [87]. This method outperformed methods that use either PET or CT but not both.

Challenges

Health care sector is a highly sensitive sector where the services and solutions are expected to be accurate irrespective of the cost incurred. The medical images are analyzed and interpreted by medical specialists; this interpretation is subject to subjectivity. Inspired by the success of deep learning in several sectors, the medical sector is also benefiting from the strength of this field. However, deep learning faces certain challenges when applied to medical applications [88], some of which are listed:

- **Dataset:** Deep learning algorithms need a huge amount of data to train the network in order to achieve accurate results. Developing a large number of medical imaging datasets is challenging, especially keeping in view that tasks such as classification and segmentation need annotations by experts. The involvement of multiple experts also becomes critical to minimize the likelihood of human error. Besides, unbalanced data is also an obstacle in applying deep learning to medicine, since it is not easy to acquire data of rare diseases which ensues class imbalance issues. The availability of multi-modal medical imaging data is even challenging. There are very few publicly available databases containing multi-modal medical data [89, 90]; they contain data of few organs such as the brain, while data of several organs is difficult to find.

- **Privacy and Legality:** Data privacy is a significant concern in the medical imaging domain, which is not associated with other kinds of real-world data. Health Insurance Probability and Accountability Act of 1996 legalize patients to protect their personally identifiable data [88]. It restricts health care professionals to disclose patient-related information. Restricted access to data and discarding valuable information could also impact the data utility.

- **Data Interoperability and Data Standards:** The data acquired using different sensors and hardware sufficiently varies. Moreover, deep learning

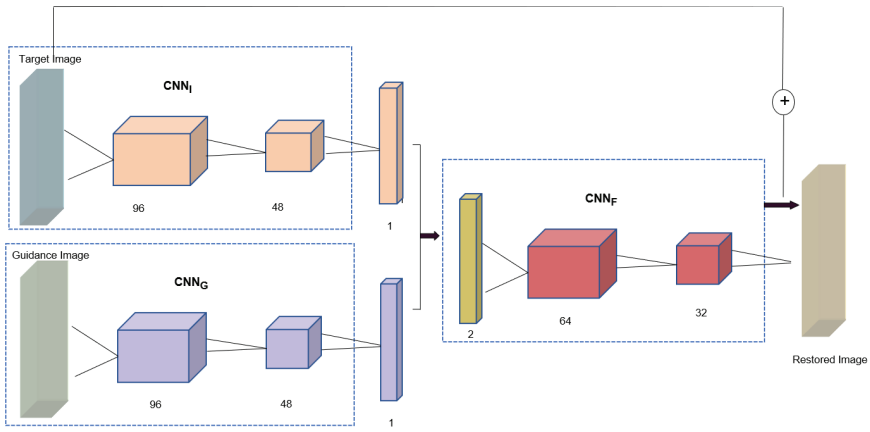


Figure 2.12: Network architecture of Joint Image Filter [3]

methods require a drastic amount of data to yield reliable outcomes. Interoperability of data then becomes indispensable.

- **Deep Learning as Black Box:** Deep Learning opened new avenues of applied research and has solved several challenging problems that were considered unsolvable before. Although, neural networks are based on mathematical models of the functioning of some stages of the cortex of the brain of certain mammals, where the algorithms accepts data, finds patterns and makes the decision, it is still deemed as a black-box and it is difficult to comprehend how a model actually works or performs better [91]. Furthermore, the architectures describing the interconnections between the different layers are not convincingly justified and are often evaluated using more or less blind test approaches based on heuristics.

2.7 Deep Learning applied to Image Enhancement

Similar to other areas of image processing such as object recognition [92], object tracking [93] and semantic segmentation [94], image enhancement [95] has also benefited from deep learning. These approaches include single image enhancement and cross-modal enhancement approaches. Most approaches in this regard formulate image enhancement as a supervised learning problem, where the network is trained end-to-end using low contrast image as input and high contrast image as ground truth [96]. Cross-modal enhancement approaches use a supplementary modality to augment the deep learning model with additional learning capability. One such primary approach denoises depth images using corresponding RGB images [97]. The proposed method was simple, which consisted of ex-

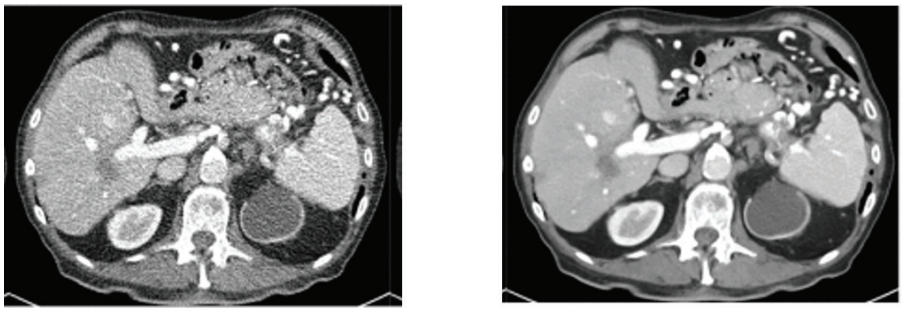


Figure 2.13: Image Denoising: a) Noisy CT (b) Denoised CT

tracting features from the dual images using two separate networks. The networks are three-layer CNNs, CNN_I and CNN_G. These features are concatenated to be fed to the third Convolutional Neural Network (CNN), CNN_F which outputs restored image. This work was further extended by inserting a skip connection between the noisy input image and the network output [3]. It enables the network to predict residuals between the noisy input image and the output. Therefore, the network learns to selectively transfer structures common in images and suppresses the inconsistent structures. Figure 2.12 shows its network architecture.

Next, we explain the deep learning approaches applied to medical image enhancement.

2.7.1 Deep Learning based Medical Image Enhancement

Medical images are susceptible to various kinds of noises introduced during acquisition. CT images, for instance, are subject to a combination of Poisson and Gaussian noise [51]. MRI is usually known to contain Rician noise, while speckle noise is dominant in ultrasound images. Denoising methods aid in removing this undesired noise from the medical images. Denoising techniques discussed in this thesis are the post-processing techniques applied in the image space. Medical imaging benefits from denoising for several reasons. In addition to improving the image quality to facilitate diagnosis for clinicians, denoised images also facilitate feature extraction task for machine learning algorithms. A noisy CT image and its corresponding denoised image are shown in figure 2.13.

Deep learning-based denoising methods can be broadly grouped as single image denoising methods and cross-modal denoising methods. In the context of supervised single image denoising methods, various methods have been proposed. One of the earliest methods proposed for medical images was applied to correct the uneven illumination in dermoscopy images [98]; a fully convolutional neural net-

work(FCN) was used for this purpose. A method applied to CT image denoising includes a Residual Encoder-Decoder CNN (REN-CNN) [99]. The authors used low-dose CT images as input to the network and normal-dose CT as ground truth to demonstrate the learning capability of their network. The majority denoising networks including REN-CNN only use Mean Squared Error (MSE) as a loss function, which makes the resultant images appear blurred. Generative Adversarial Network (GAN) based on generative modeling was introduced to minimize the undesired smoothness introduced due to sole usage of MSE [100]. Furthermore, MSE was combined with structural similarity index loss for better results [101]. Several variants of GANs emerged offering improvements over their previous counterparts in retaining the quality of images during denoising [51]. The generic structure of GANs is explained in the next subsection. The deep learning models are trained to learn the mapping from noisy image *Y* to the noise-free image *Z*. In cross-modality guided denoising approaches, the models are trained to consider another image *G* from a different modality as well during this learning process.

$$\hat{Z} = f(G, Y) \tag{2.13}$$

The deep learning model is then trained to minimize the loss function *L*:

$$f^* = \arg\min L(f(G, Y), Z) \tag{2.14}$$

Very few cross-modality guided denoising schemes targeted to medical images exist in the literature [56, 102]. One such work consolidated information from PET and MRI (T1 and T2 FLAIR) to denoise ultra low-dose (very low-dose) PET images [78]. The proposed method was a fully convolutional network implemented as a multi-scale encoder-decoder structure (U-Net [102] with residual learning). The proposed denoising scheme offers the prospect to curtail the radiation dose given to certain patients without compromising on the quality of acquired images. Another similar approach used T1-w and T2-w brain images for denoising. The proposed method [56] combined a guided filter with the guidance map generator, which was realized as U-Net [102] with two branches in the encoding path. Each branch encodes features of each modality individually which are concatenated in the end. The guidance filter combines the output of the network with the input image to yield a denoised image.

The above-mentioned methods do not fully realize the potential of cross-modal guidance, as they extract features individually from each modality and concatenate in the last layer of encoding path. This approach does not fully exploit the

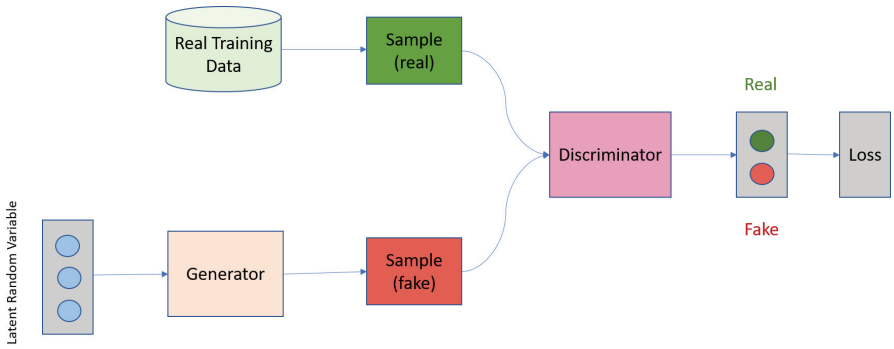


Figure 2.14: Generative Adversarial Network

essence of cross-modal guidance, which leaves a huge gap for further research. A method incorporating hierarchical cross-modal feature manipulation using Siamese network is proposed for denoising brain MRI using cross-modal information (**Paper F**). An overview of Siamese networks is provided in Sec. 2.7.3.

2.7.2 Generative Adversarial Networks

Goodfellow [100] introduced a novel architecture, GAN, to estimate generative models in an adversarial manner. GAN consists of two components: generator and discriminator, that are competing against each other to solve the underlying problem. Both components are neural networks, where the generator learns to generate data that resembles the real data; these generated samples are deemed as negative training instances for the discriminator. Conversely, the discriminator is trained to determine what is the probability of the generated data belonging to a training class, i.e., being real or generated (fake). The generic architecture of GAN is demonstrated in figure 2.14.

When training of GAN begins, the generator starts creating fake samples, that are not good enough to fool the discriminator. However, when the generator is well-trained, the discriminator classifies fake samples as real. The discriminator in this scenario is trained in a supervised manner. It acts as a binary classifier trained on the fake data received from the generator and the real training data. It outputs the probability of the data fed to it based on which the training continues. The training resumes until the discriminator can no longer distinguish the fake data from the real one. The discriminator-generator role in GAN can be explained using the min-max game analogy. In a two-person zero-sum game, a person is a winner only if his competitor loses. Both compete against each other and improve their performance by learning from mistakes until one person outplays.

The loss function in GAN is defined for discriminator and generator both. GAN attempts to mimic the probability distribution of real samples and thus generates fake samples. The loss function is therefore formulated to capture the distance between the data distribution of generated data and that of the genuine training instances. The minimax loss does this. Discriminator minimizes negative log-likelihood, whereas generator maximizes it. Formally, the GAN objective function can be expressed as follows:

$$\min_G \max_D E_{x\sim P_r}[\log(D(x))] + E_{\tilde{x}\sim P_g}[\log(1 - D(\tilde{x}))] \tag{2.15}$$

where P_g and P_r represent model distribution and actual data distribution respectively. In the original GAN paper [100], the objective function is stated as minimization of Jensen Shannon Divergence. GANs have been used in a variety of applications, including image enhancement [103, 104], image-to-image translation [105], super-resolution [106], segmentation [107], denoising [108]. Wolterink et al. [108] used GAN to denoise the low-dose cardiac CT images using normal dose CT as ground truth. Although the proposed method outperforms other similar deep learning approaches, they used Mean Squared Error (MSE) [109] as a loss function in their framework which leads to blurred images and loss of detail. This effect was improved by [51]; a Wasserstein distance-based GAN that calculates the difference between the generated image and ground truth in feature space instead of image space as done when MSE is used.

Several variants of GANs have been proposed for medical images including cycle-GAN [110] and conditional GAN [111] to improve their performance. Besides, most approaches mentioned above use paired input-ground truth data to train the network, which is even more challenging in the context of medical images. Cycle-GANs bypass this barrier and allow the use of unpaired data. They incorporate several losses in the architecture such as cycle consistency loss and identity loss, which enables the network to learn embedding of input data and generate samples in the same space [103]. In our thesis, we used cycle-GAN to enable the proposed method in **Paper E** learn global contrast characteristics from high contrast T1-w images in order to improve the contrast of T2-w images.

2.7.3 Siamese Neural Networks

Siamese Neural Networks (SNNs) are the neural networks that consist of two or more similar subnetworks, also called twin networks. The subnetworks not only have the same architecture but also identical parameters and weights. SNNs compare the feature vectors of the paired input data to identify similarities between them. A simple Siamese Neural Network (SNN) is shown in the figure 2.15. Sia-

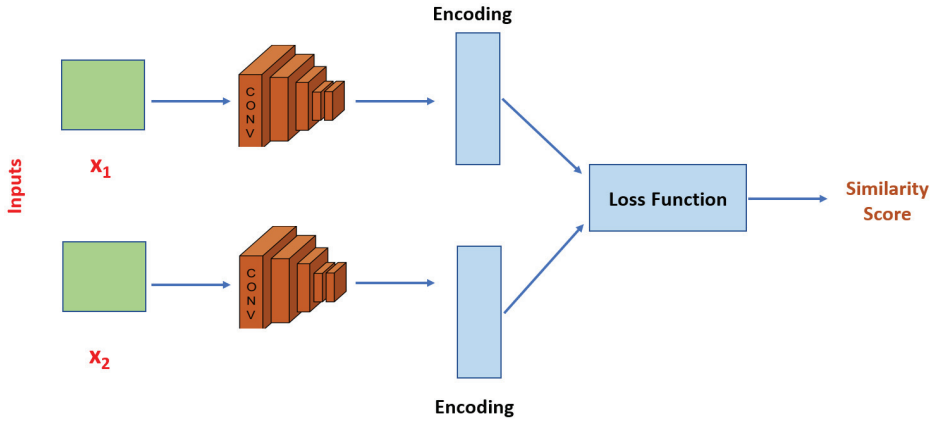


Figure 2.15: Siamese Network

mese networks have been applied to a multitude of applications including classification, segmentation [94], object tracking [93] etc. In medical images, they have been employed to real-time tracking of knee cartilage in temporal and spatio-temporal sequences of 2D ultrasound images [112]. The proposed method 'Siam-U-Net' combines U-Net with Siamese network to give cartilage segmentation in US images.

Multi-modal MRI constituting T2-w and Diffusion-Weighted Imaging (DWI) along with their multi-view representation (Axial and sagittal) were used to retrieve similarly diagnosed lesions in a content-based image retrieval system [113]. The method was presented to aid the radiological interpretation of images in ultimately diagnosing prostate cancer. Utilizing Siamese networks for multi-modal medical image processing is motivated by the notion that these images acquired from the same patient possess a similar structure, therefore, the correlated information they contain can be manipulated in a sophisticated way to accomplish the underlying tasks. SNNs were applied recently to image restoration where dense siamese networks remove the undesired flash effect in images via flash/ no-flash image pairs [114]. A denoising method proposed in **Paper F** combines siamese network with hierarchical feature manipulation strategy.

Chapter 3

Papers Summary

Image contrast enhancement and denoising are the two pre-cursors applied to improve the perceptual quality of images. In this thesis, traditional methods and deep learning methods have been proposed and applied to multi-modal medial images of different organs including the liver and brain using complementary information in the corresponding image. This chapter provides an overview of all the papers produced during this tenure of research.

Overall, the papers provided here can be grouped into two types: traditional methods and deep learning-based methods. **Papers A** - **D** are based on traditional methods, while **Papers E, F** are deep learning-based. All the papers cover two broad areas of image enhancement, that is, contrast enhancement and denoising that are the focus of this work. A summary of all the papers is presented highlighting the objective, main idea and contributions.

Papers A and **B** (Sec. 3.1 and 3.2) respectively suggest the methods for contrast enhancement of liver CT images to improve visibility of structures and enable segmentation algorithms to work better. **Paper D** (Sec. 3.3) proposed a parallelization strategy to first accelerate the approach presented in **Paper A**, and then accelerate the gradient-driven seeded Region growing algorithm. Consequently, the execution time for both tasks is drastically reduced. **Paper C** (Sec. 3.4) proposes an efficient parallelization strategy to improve the execution time of the parallel gradient driven SRG algorithm. The core idea behind the technique is to accelerate the segmentation implementation and was tested on vessel segmentation; however, the input CT images were pre-processed using the approach presented in **Paper A** (Sec. 3.1) to improve the contrast of images and particularly highlight vessels. The focus of **Papers E** and **F** is deep learning; **Paper E** applies deep learning to im-

prove the contrast of T2-w brain images using information from the corresponding T1-w MR images. **Paper F** exploits cross-modal guidance by combining features from multi-modal images to denoise the medical images.

3.1 Summary of Paper A

Paper A proposed a contrast enhancement method applied to liver CT images using cross-modal guidance from the corresponding MR images [115]. The proposed method is the first application of cross-modal guided enhancement applied to medical images to the best of our knowledge. CT images suffer from low contrast, whereas MR images have good contrast. In this scenario, the information in MR image can be exploited to enhance the contrast of CT image.

The proposed method improves the contrast to facilitate the visibility of important structures in the CT images, such as tumors and vessels. (Cross-modality Guided Enhancement (CMGE)) approach thus presented was a combination of two-dimensional histogram specification (2D-HS) and morphological operations. 2D-HS is a traditional contrast enhancement approach that not only considers the individual pixel values while calculating the PDF, rather it is a context-aware technique that also considers the spatial pixel correlation while applying the HS operation. It exploits the neighboring pixels as well while computing the PDF and calculates 2D CDF instead. Morphological operations such as top hat and bottom hat transform are also used in the medical images due to their capability of highlighting dark and brighter areas than their surrounding.

The data for the proposed study was acquired from Intervention Center, Oslo University Hospital, Norway. The flowchart of the proposed method is shown in figure 3.1. The method was compared with other single-image contrast enhancement methods; quantitative assessment of the results validated the better performance of the proposed technique in improving the contrast of liver CT images. The main contributions of the proposed research work are as follows:

1. The proposed technique CMGE is the first attempt to apply the idea of cross-modal guidance to medical images for contrast enhancement to our knowledge.
2. The method was validated for improvement in contrast on real human liver data (CT and corresponding MR images).

3.2 Summary of Paper B

A goal-oriented contrast enhancement approach was proposed for CT images in **Paper B**. It uses a combination of cross-modal guidance information and optim-

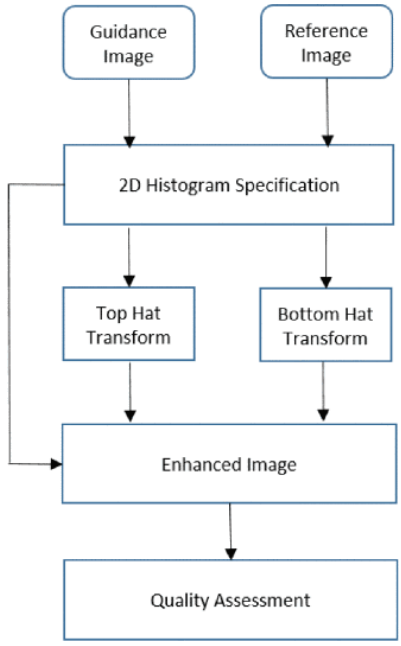


Figure 3.1: Flowchart of proposed approach (CMGE)

ization schemes. The main steps involved in the proposed method OPTimized Guided Contrast Enhancement (OPTGCE) are shown in figure 3.2. The underlying objective was to improve the contrast of images so the tumor edges in CT images are discriminated enabling segmentation algorithms to work well. The motivation of incorporating SSIM based metric was to retain structural similarity between the enhanced image and original CT image during enhancement. A segmentation algorithm was later applied to determine if the algorithm works better on images enhanced using the proposed method in comparison with those enhanced using similar histogram-based optimization schemes.

The data for the study constituting CT-MRI pairs (99 samples) along with tumor segmentation ground truth was provided by Intervention Center, Oslo University Hospital, Norway. The assessment of the proposed method was done using contrast assessment metrics including mutual information-based metric MIGLCM [116], Multi-Criteria Contrast Enhancement Evaluation Metric (Multi-Criteria Contrast Enhancement Evaluation (MCCEE)) [117], and entropy as well as segmentation assessment metrics such as dice score, positive predictive value, and Hausdorff distance. It was concluded that the proposed method performs better than similar optimization-based approaches that do not incorporate cross-modal guidance information both in enhancing the image contrast as well as in terms

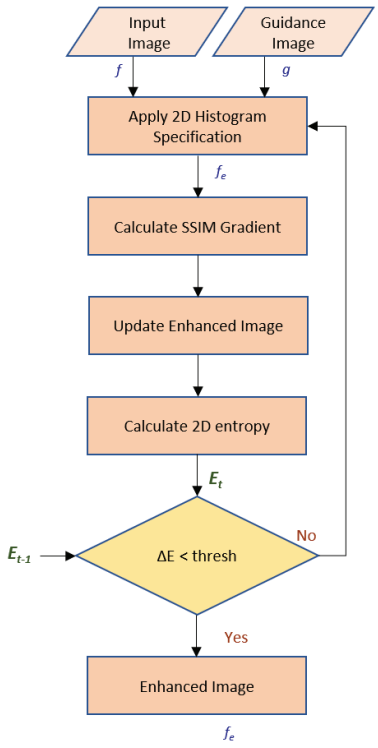


Figure 3.2: Flowchart of proposed approach (OPTGCE)

of segmentation applied post-enhancement. The main contributions of the paper [118] are listed below:

1. The two-dimensional histogram specification-based contrast enhancement process is formulated as an optimization problem in the context of medical images.
2. The image contrast enhancement is controlled using an optimization scheme based on SSIM gradient.
3. Goal-oriented performance evaluation of the proposed method is done through objective quality metrics and through segmentation results.

3.3 Summary of Paper C

Paper C proposed an efficient parallel implementation of gradient-based seeded region growing algorithm applied to vessel segmentation [119]. The fastest algorithms executed on are those where blocks responsible for carrying out each

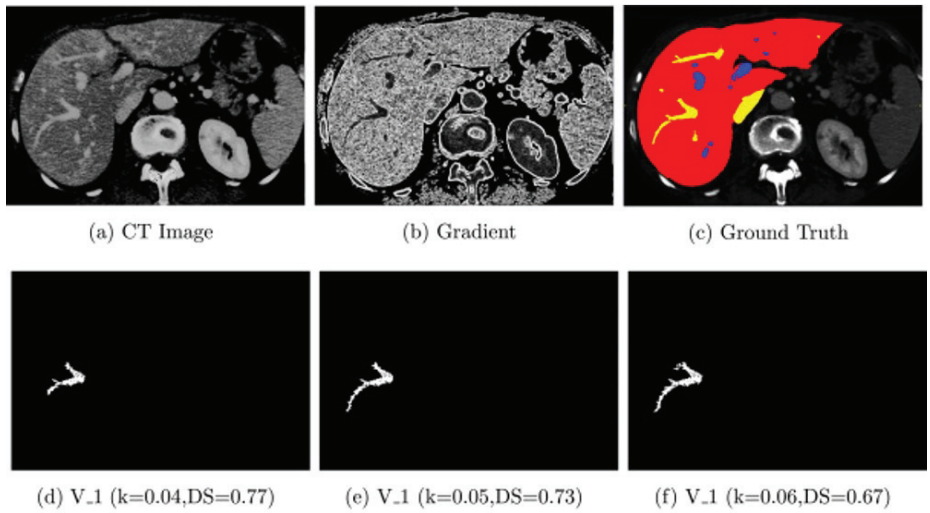


Figure 3.3: Vessel segmentation using parallel SRG

task in the algorithm are not dependent on each other and synchronization among them is minimal. Because of the iterative nature of the seeded region growing algorithm, it is implemented as kernel termination and relaunch (KTRL) on GPU, which involves intermediate memory transfers between CPU and GPU. This process is time-consuming indeed. Therefore, a persistence and grid-stride loop-based parallel approach was proposed in the paper that is executed entirely on GPU. This mechanism prevents intermediate data transfers between GPU and CPU.

The ground truth for the proposed work was provided by clinicians in the Intervention Centre, Oslo University Hospital, Norway. The liver CT images used in this work were enhanced using the technique proposed in **Paper A** to enhance important structures in the CT images such as vessels. Some results of the parallel SRG are shown in figure 3.3. Figure 3.3 a - c show the input CT, gradient of image a and ground truth respectively, whereas figures d-f show the vessels segmented using proposed approach. Intel (R) Core(TM) i7-7700HQ CPU @ 2.80 GHz RAM 24 GB, NVIDIA GPU 1050 (RAM 4 GB), OpenCL1.2 and CUDA Toolkit 10.1 was used to carry out this work.

The proposed method is able to accelerate the parallel mechanism by attaining a speedup of 1.9 times compared to KTRL. Moreover, the parallel seeded region growing yields more accurate results than Chan-Vese [120] and Snake model [121] in segmenting vessels in the liver CT images when assessed using quantitative measures such as dice score.

3.4 Summary of Paper D

Paper D proposes a parallel implementation of **Paper A**, that is cross-modality guided image enhancement [122]. Furthermore, the dynamic region of interest (RoI) based SRG method is presented for tumor segmentation in the enhanced liver CT images. The accelerated performance on GPU is achieved by dividing the tasks into several active threads. Furthermore, experiments were conducted to determine if SRG works better on enhanced images compared to original images. With the help of quantitative measures like sensitivity and accuracy, it was concluded that the algorithm works better on enhanced images. Data and segmentation ground truth for the work was provided by Intervention Center, Oslo University Hospital, Norway.

Figure 3.4 shows all the steps involved in parallelized contrast enhancement and tumor segmentation. After CT and MR images are transferred from CPU to GPU, GPU performs cross-modal contrast enhancement on the low contrast CT image. For the next task, that is, tumor segmentation, the gradient of contrast-enhanced image is calculated and is transferred to the subsequent module in GPU through Interblock GPU synchronization (IBS). SRG is applied on the gradient of the enhanced image afterwards. Based on the value of the initial seed, region growing starts. New seeds are formed from initial seeds depending on the threshold criterion selected. The region grows iteratively until the criteria are fulfilled. The region does not grow further when new seeds cannot be formed. The threshold criterion adopted here is inspired by the work of Rai [123]. It is based on the homogeneity of the region and region aggregation considering pixel values and their gradient direction and magnitude.

The proposed GPU-based implementation was able to achieve a drastic speed-up of around 100 times compared to its corresponding sequential implementation (on CPU) of contrast enhancement and tumor segmentation.

The following contributions in this work are listed:

1. This is the first work addressing high-performance multi-modality guided liver contrast enhancement applied to tumor segmentation to the best of our knowledge.
2. Speed up achieved in case of parallel implementation (GPU-based) of CE and segmentation is more than 100 times compared to its CPU equivalent implementation.
3. It is experimentally validated that the segmentation works better on enhanced images compared to original unenhanced CT images of liver.

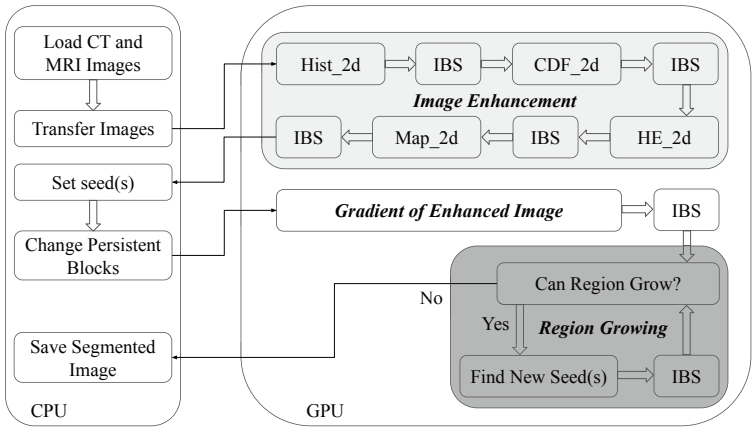


Figure 3.4: GPU implementation of SRG based tumor segmentation

3.5 Summary of Paper E

A deep learning-based contrast enhancement approach is proposed in **Paper E**. In this approach, contrast enhancement is accomplished by using a Generative Adversarial Network (GAN). Since contrast enhancement is highly subjective, it is challenging to find ground truth for it, particularly for medical images. In this scenario, corresponding multi-modal medical images of better perceptual quality could be employed in the enhancement process, thanks to the complementary redundant medical imaging data acquired during certain therapies. The structural similarity between multi-modal MR images can be particularly advantageous in the feature extraction and learning process. Cycle-GANs [110] are considered ideal for learning in the presence of unpaired data. We employed cycle-GANs to solve our cross-modal contrast enhancement problem. Cycle-GANs were applied by Chen et al. [124] for contrast improvement of natural images in case of paired high-contrast ground truth as well as unpaired ground truth (high contrast images with entirely different contents). Inspired by this idea, the proposed method uses U-Net with global features augmented to learn the global contrast from corresponding multi-modal images.

Multi-modal MR data of human brains from IXI database (Hammersmith Hospital, United Kingdom) was used in this work [89]. The contrast of actual T2-w images (input images here) was degraded by applying a combination of morphological operations. The original T1-w images of good contrast were used as reference or ground truth. Comparison of the proposed method was done with existing CE approaches using well-accepted quality assessment metrics Feature Similarity Index (FSIM) and Blind/Referenceless Image Spatial Quality Evaluator

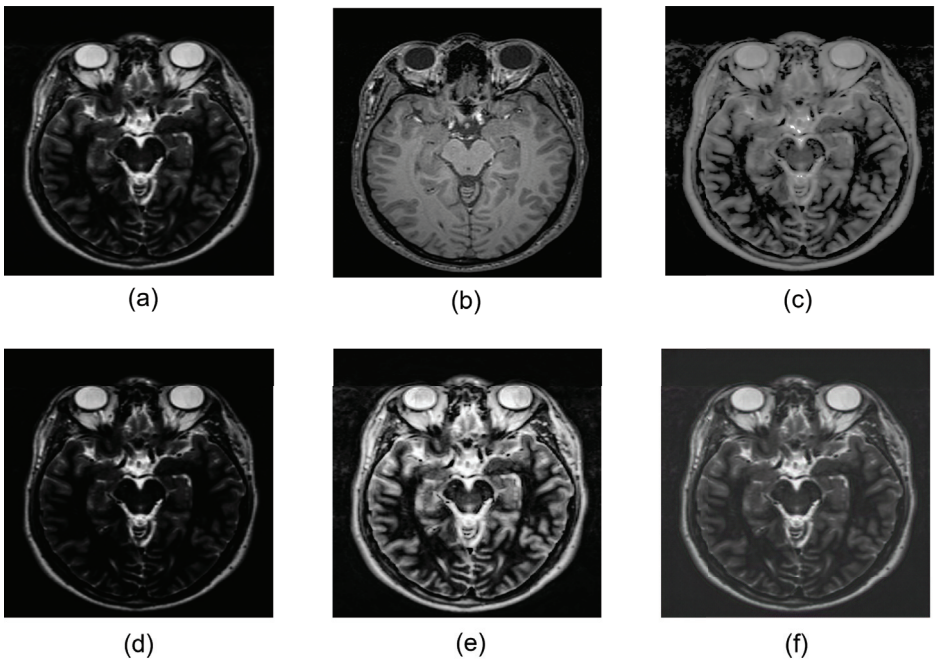


Figure 3.5: Comparison of proposed method with recent CE methods (a) Input (b) Ground Truth (c) CMGE [4] (d) Zohair et al. [5] (e) CLAHE [6] (f) proposed

(Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)). Results from **Paper E** are shown in figure 3.5.

The following contributions of the paper are listed:

1. The CE method proposed in this paper is the first deep learning assisted cross-modality guided CE approach applied to MR images, to our knowledge.
2. Comparison of the proposed method with existing CE approaches showed that the proposed method improves contrast without introducing artifacts in the enhanced image.

3.6 Summary of Paper F

Paper F proposed a deep learning-based denoising approach that exploits the better perceptual quality of T2-weighted images (noise-free) to restore its corresponding noisy T1-w MR images. The proposed method Cross-modality Guided

Denoising Network (CMGDNet) combines Siamese networks with hierarchical feature learning in a supervised manner. CMGDNet consists of two main modules which are further decomposed into two parts. The first module Paired Hierarchical Learning (Paired Hierarchical Learning (PHL)) is responsible for feature extraction in a hierarchical fashion; it acts as an encoder in our model. ResNet-101 was used to extract features [125]. The second module Cross-modal Assisted Reconstruction (Cross-modal Assisted Reconstruction (CMAR)) combines these features in a systematic way from dual modalities. This operation entails combining features from both images using a complementarity-aware mechanism followed by upsampling. CMAR acts as a decoder and is implemented as a dense network using the Inception module. The block diagram of the proposed technique is shown in the figure 2.15. The complete structure of the modules is explained in paper [126].

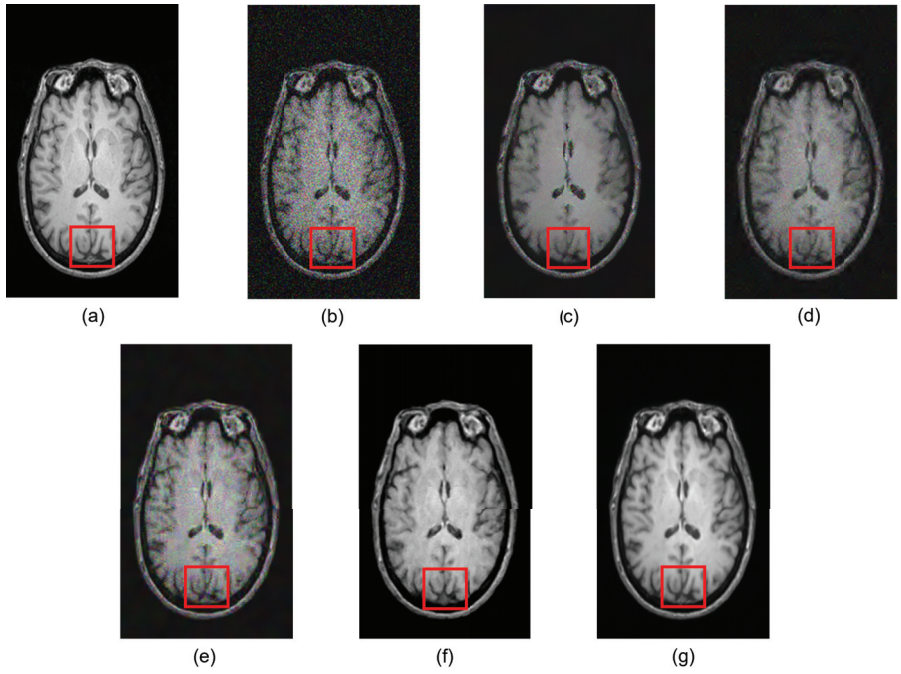


Figure 3.6: Denoising Results: (a) Ground Truth (b) Input (c) NLM [7] (d) SURE [8] (e) BM3D [9] (f) MCDN [10] (g) proposed

Two public datasets constituting T1-w and T2-w MRI of healthy patients (from Hammersmith Hospital and Guy's Hospital), a subset of [89] were used in the study. All the volumes were resampled to have a dimension of $256 \times 256 \times 150$. Rician noise was added to T1-w images that were fed to the proposed model as

input along with noiseless T2-w images while original T1-w images were taken as ground truth. Several experiments were conducted to analyze the performance of the proposed method in denoising the images contaminated by varying levels of Ricain noise. Besides, the role of image registration in improving denoising was also inspected. The proposed method was compared with state-of-the-art methods such as BM3D [9] and MCDN [10]. The denoising results were qualitatively and quantitatively evaluated using Peak Signal-to-Noise Ratio (), SSIM and Feature Similarity Index (Feature Similarity Index (FSIM)). The qualitative comparison is shown in figure 3.6.

Following contributions are made in the paper:

1. A new architecture based on cross-modal guidance is proposed to denoise T1-w brain MR images. Siamese network combined with hierarchical feature learning was used for effective feature learning. The proposed framework employs supplementary information from the better perceptual quality T2-w image in the denoising process.

2. Cross-modal feature synthesis is not well explored in the context of cross-modal denoising. An effective cross-modal information synthesis policy is devised. The results demonstrate that this strategy of feature combination works well.

3. CMGDNet exhibits an average gain of 4.7% and 2.3% respectively in terms of SSIM and FSIM compared to state-of-the-art denoising methods.

4. Our experiments show that registration between multi-modal medical images aids in retaining structural similarity with the original image when both images were fed to the denoising framework, however, improvement in MSE values was not observed.

Chapter 4

Papers Discussion

In this chapter, contributions to cross-modality-guided contrast enhancement and denoising in the context of medical images are presented. Afterward, we discuss the limitations of the research work reported here along with future work suggestions.

4.1 Contributions of the Thesis

This dissertation addresses two essential components of medical image enhancement, that is, contrast enhancement and denoising. Subsection 4.1.1 discusses the contributions in contrast enhancement using traditional methods and deep learning approaches, whereas section 4.1.2 discusses the contributions in denoising task using deep learning.

4.1.1 Contributions to contrast enhancement

Papers A - **B** and **Paper E** included in this dissertation focus on cross-modality guided medical image contrast enhancement. **Papers A** - **B** contribute to the enhancement of liver CT images which are usually preferred for initial diagnosis of liver cancer. MRI is another modality used for liver cancer screening. Both are clinically significant and can be found predominantly in the clinical surgery suite.

CT intrinsically has low contrast, while MRI offers better perceptual quality. We used this redundant information of superior quality in the cross-modal CE framework. Existing medical image contrast enhancement methods do not take advantage of this redundant information to our knowledge, and our approaches in **Paper A** - **B** are an attempt to investigate this notion. The approaches proposed in both papers enhance the contrast of CT images; however, the objective of **Paper A** is to

improve visibility, whereas approach in paper **Paper B** aims to improve contrast to aid tumor segmentation algorithms. Since accurate registration is a challenging problem for multi-modal medical images [127], particularly liver images, a global contrast enhancement approach was adopted. It was proved that extracting complementary information from another image via a context-sensitive approach produces superior results to single image enhancement methods. Our proposed approach investigates the following research question:

Q1.1 *"Does incorporating guidance information from a different modality image improve the contrast of a medical image for improved visibility ?"*

In **Paper B**, contrast enhancement is accomplished using cross-modal guidance information to provide better-contrast CT images as input to the segmentation algorithm. Although histogram-based techniques are prone to introduce artifacts that can possibly affect the appearance of critical structures in medical images, we incorporate an effective optimization strategy to retain the structural affinity of the enhanced image with the original image. Consequently, the proposed technique retains the structure of tumors as well while simultaneously enhancing the contrast. The enhanced images thus obtained possess better-highlighted edges, which leads to improved performance of segmentation. Therefore, **Paper B** addresses the following research question:

Q1.2: *"Investigate if tumor segmentation methods work better on images enhanced using proposed CE method ?"*

As far as quality assessment (QA) of the medical images is concerned, it is not as straightforward as it is for natural images. It is due to the absence of dedicated QA metrics specifically developed for medical images [128]. Therefore, to thoroughly evaluate our proposed goal-oriented CE method, the quality assessment was done from two perspectives. First, the main task that is contrast enhancement was evaluated using quality assessment metrics. One of the comprehensive metrics, which is Multi-Criteria Contrast Enhancement Evaluation (MCCEE) considers the improvement in contrast along with over-enhancement to systemically evaluate the improvement achieved in contrast. Moreover, it considers the correlation between CE criteria and subjective quality scores in a learning-based quality assessment approach. Dice scores calculated from the results of subsequent segmentation were therefore used in the MCCEE score computation. Besides, the results of tumor segmentation applied to enhanced images were also quantitatively evaluated. The experiments were conducted on real liver CT-MR data obtained from the Intervention Centre, Oslo University Hospital, Norway. The quantitative assessment of the performance of proposed method along with a comparison with other single image enhancement approaches is summarized in table 4.1.

Table 4.1: Quantitative assessment of different enhancement methods

Dataset #	Entropy				MIGLCM			
	HEMIC [129]	**AVHEQ** [130]	**CMGE**[4]	**OPTGCE**	**HEMIC** [129]	**AVHEQ** [130]	**CMGE** [4]	**OPTGCE**
1	2.32	2.21	2.76	**3.13**	**1.1**	1.07	0.95	**1.1**
2	1.58	1.7	1.74	**2.4**	0.93	0.9	0.86	**1.12**
3	1.7	1.6	1.7	**1.9**	0.93	0.92	0.91	**1.06**
4	2.03	1.95	2.72	**3.00**	1.22	1.21	1.27	**1.4**
5	2.51	2.45	3.43	**3.62**	1.16	1.12	1.07	**1.31**
6	1.54	1.51	1.92	**2.12**	0.86	0.85	0.82	**0.98**
7	2.61	2.52	3.23	**3.64**	1.28	1.24	1.14	**1.42**
8	1.52	1.46	1.94	**2.1**	0.82	0.81	0.75	**0.94**
9	1.43	1.36	1.67	**1.88**	0.78	0.77	0.75	**0.81**
10	1.11	1.11	1.32	**1.41**	0.64	0.62	0.63	**0.76**

The quick acquisition time associated with CT images makes them favorable to be integrated with the intra-operative navigation. Ultrasound is also used in the intra-operative setting, but, it is difficult to find the multi-modal liver data where one of the dual modalities is ultrasound; therefore, we resorted to the problem of CT image enhancement using cross-modal guidance from MRI. By using contrast-enhanced CT as an intra-operative imaging modality, the performance of subsequent phases of surgical navigation can be streamlined. These phases include segmentation of critical structures such as liver, tumors, hepatic, and portal veins in the liver CT image followed by bio-mechanical modeling of these structures. In this context, the execution time of the contrast enhancement algorithm is one of the factors that determine its feasibility in real-time implementation.

Paper D proposed a strategy to accelerate the contrast enhancement method proposed in **Paper A**, so image enhancement could be realized in the intra-operative phase as well. The acceleration on GPU is achieved by dividing the intermediate tasks of the algorithm into several active threads, where each thread is responsible for carrying out a specific task. Thus, several tasks are accomplished in parallel. Furthermore, experiments were conducted to determine if gradient-driven seeded region growing algorithm [123] works better on enhanced images compared to original images. With the help of quantitative measures like sensitivity and accuracy, it was concluded that the algorithm works better on enhanced images. This method was the first parallel implementation of the cross-modality-guided contrast enhancement method presented in **Paper A**. The method achieved a speed-up of around 100 times on NVIDIA GPU 1050 compared to its corresponding sequential implementation.

Paper C presented an efficient parallel implementation of gradient-based seeded

region growing algorithm, which was applied to vessel segmentation. The focus of this work was GPU-based optimization of segmentation method, however, the input CT images were pre-processed using the method proposed in **Paper A**, before SRG was applied to them. The vessel segmentation ground truth was provided by the Intervention Centre, Oslo University Hospital, Norway.

In this thesis, the potential of deep learning was also analyzed in solving CE and denoising problems. Deep learning has revolutionized the digital image processing domain [131, 132]. Deep learning-based methods predominantly outperformed the existing techniques [133]. Medical imaging also benefited from deep learning. Deep learning was applied to cross-modal medical image contrast enhancement in **Paper E**. It is challenging to obtain ground truth for contrast enhancement particularly in medical images, moreover, it varies subjectively as well. **Paper E** presented an idea to learn the contrast enhancement using the corresponding high contrast multi-modal medical images. To this end, dark (low-contrast) T2-w MR images were enhanced using the corresponding T1-w MR images. Since both share a structural similarity; however, they come from different domains, therefore we used cycle-GAN to learn the global contrast characteristics. Cycle-GANs offer a promising way of learning from unpaired ground truth and generally perform superior to similar DL-based methods. The proposed approach was the first deep learning-based method applied to cross-modal contrast enhancement in the medical imaging domain to the best of our knowledge.

4.1.2 Contributions to Denoising

Cross-modal denoising in the medical imaging context is a relatively less-explored area. Few deep learning approaches proposed recently extract features from the multi-modal input images and merely concatenate them in the later encoding layer of the encoder-decoder network [56, 78]. These approaches do not combine the features extracted from both modalities in an efficient manner and therefore do not fully exploit the complementary information. In **Paper F**, we propose a cross-modal denoising approach that first extracts hierarchical features and combines them, thereby providing a better feature combination strategy. Furthermore, the impact of registration on denoising performance is particularly investigated.

Brain Magnetic Resonance images (T1-w and T2-w) from a publicly available database were used for this study. One of the objectives of particularly experimenting with brain image data was that there are several open-source software providing tools for multi-modal brain image registration [134]. However, it is comparatively difficult to register multi-modal liver images. This paper attempts to investigate the following research question:

Q2.1: *"How multi-modal guidance information could be employed for denoising medical images using deep learning ?"*

4.2 Limitations and Future Perspectives

This section discusses the limitations of our research work followed by some suggested directions in which the work can be extended in the future.

Paper A improves the contrast of CT images using a combination of histogram-based approaches in conjunction with morphological operations. Although the proposed techniques incorporates contextual information, however, approaches such as wavelet domain methods and other sophisticated approaches could be investigated in the future to address some of the limitations associated with histogram approaches. The results in the paper were quantitatively evaluated, however, since the work focuses on medical data, it would be valuable to involve clinicians particularly radiologists in the subjective assessment and viability of such methods.

Paper B also proposed a contrast enhancement approach that uses histogram combined with a structural similarity based measure to retain the similarity of the enhanced image with the original image. Both the methods are traditional (i.e. non DL-based methods); however, in the future, deep learning methods can be explored to solve the underlying problem. Moreover, the 2D histogram-based optimization methods can be applied to the multi-modal medical images of other organs such as lungs, kidneys, etc. Since cross-modality guided CE methods have shown promising results for natural images, we believe that cross-modal CE for medical images is an interesting and promising future work direction that has the potential to provide significant improvements.

The comparison of traditional methods and deep learning methods would enable to analyze the results obtained using both contemporary approaches. Multi-modal medical imaging data of a few organs can be found on public databases [90], which could be utilized for this purpose.

Contrast enhancement is a subjective process. **Paper E** proposed a deep learning-based approach, specifically cycle-GAN that learns the rich contrast characteristics from the corresponding multi-modal MR image (unpaired). However, in the future, unsupervised learning approaches would be an interesting direction to investigate this problem. EnlightenGAN [95] is one of the unsupervised learning methods that eradicates the need for paired ground truth in the problem of low light image enhancement. It proposed to minimize perceptual loss between the feature maps of input images captured under low-light illumination and its enhanced variant. EnlightenGAN employs pre-trained VGG [135] to extract features from network

input and output. This loss was termed 'self feature preserving loss'. The unsupervised approaches in this framework differ from their supervised counterparts where ground truth is a manually enhanced image. A similar unsupervised mechanism can be employed to enhance a low-contrast medical image provided paired multi-modal images are fed to the enhancement model, where the other image is a guidance image of better perceptual quality. Moreover, the proposed network in **Paper E** was trained using T1 T2-w paired MR images of the human brain. Other paired modalities such as CT -MRI and or CT-ultrasound, MRI-ultrasound can be used in an unsupervised learning configuration. Furthermore, other organs can also be experimented with.

The objective assessment of medical image enhancement methods is done using the quality assessment metrics proposed for natural images. Quality assessment in the case of natural images is a widely investigated area; several full-reference, reduced-reference, and no-reference metrics have been presented. The objectives of medical image enhancement differ from those of natural images. For instance, the objective of contrast enhancement in **Paper B** was to discriminate tumor borders so the segmentation methods could better segment based on the intensity value of the tumor pixel and its neighborhood. In this case, the QA metrics need not consider over-enhancement effects as an important quantifiable factor/ parameter. In such a scenario, there is a need for quality assessment metrics specifically designed for medical images specially for CT, MRI, and Ultrasound images. It should be highlighted here that an approach similar to MCCEE (used for evaluation of goal-oriented CE in **Paper B**) that takes into account subjective assessment as well can be further explored [117]. Methodologies proposed for natural image quality evaluation can be adopted to investigate QA methods for medical images as a future research direction.

Cross-modal denoising in the medical imaging context is a relatively less-explored area. Besides, the existing techniques do not fully exploit the complementary information. **Paper F** proposed a method to denoise T1-w MR images using the corresponding better perceptual quality / non-noisy T2-w MR images. Although, the scheme extracted features from the dual input images and combined them by addition and multiplication of the features at respective encoding layers, however, more sophisticated ways of combining features can be explored in the future that does not merely rely on simple mathematical operations. Second, the noise was introduced into images synthetically, both for training and testing. However, in the future, it would be interesting to analyze how the proposed method performs on real noisy MR images. Another interesting observation is that both images were fed to the network and features were extracted from the whole images. Since the proposed method also aimed to investigate the impact of registration on the denois-

ing performance of the network, patch-based paired input could be provided to the network to thoroughly investigate the impact of the image to image registration on cross-modal denoising.

As mentioned earlier, unsupervised learning approaches [136] can also be explored and applied to cross-modal denoising. The deployment of these approaches in the clinical setup is a more practical approach. Similarly, other combinations of multi-modal imaging data can be provided to the supervised and unsupervised learning models to thoroughly interpret their behavior.

The discussion provided in this section points towards few future perspectives regarding the modules specifically presented in this research work. Now, we provide an overview of the future work in the context of vision-based surgical navigation in general.

The artifacts encountered at the medical image acquisition stage can impact the performance of the intermediate steps carried out during the pre-operative planning and surgical navigation, may it be segmentation of critical structures, biomechanical modeling in addition to visibility of organ's pathology.

Intra-operative imaging particularly CT and ultrasound imaging (because of quick acquisition time and provision of adequate information) is deployed nowadays in the surgical suite to update the initially formulated navigation plan as well as to convey precise updated representation of the actual organ anatomy. In this scenario, image enhancement and denoising mechanisms need to be incorporated in the real-time setting. Although there is no study focused on the requirements and constraints in connection with the speed of the navigation system, real-time implementation is the need of the hour and would be beneficial indeed.

The above-stated challenges are difficult to address. However, with the advancements in technologies, hardware and imaging techniques, intra-operative information can be acquired efficiently. New CT imaging techniques such as Dyna-CT are being used at some hospitals. Moreover, the potential of combining the information from images such as CT or ultrasound with intra-operative imaging captured in real-time such as laparoscopic video can be explored. The success of computer vision and deep learning is already being witnessed in applications such as tracking, augmented reality. Inspired by this promising performance, deep learning techniques can be introduced into surgery planning and navigation.

With the advancement in hardware such as robotics, HoloLens, imaging techniques, and modern computer vision algorithms, the integration of image enhancement techniques with laparoscopic navigation system can address the existing bottlenecks in the implementation of these systems.

Chapter 5

Conclusion

The goal of this dissertation is cross-modality guided contrast enhancement and denoising of medical images in the context of surgical planning and navigation in image guided laparoscopic liver resection (described in Sec. 1.2 and 1.3 respectively). The six articles published during this research tenure cover our contributions to addressing these tasks.

The first part of our work presented in Sec. 2.4 addressed the research questions related to cross-modal contrast enhancement. The ultimate objective of this task is twofold, i.e., improving the visibility of structures in the organ and improving tumor segmentation, as pointed in figure 2.8.

We proposed a traditional cross-modality guided global contrast enhancement method in **Paper A**. The results showed that the proposed method improves contrast of CT images and visibility of critical structures in the liver. One of the challenges posed to tumor segmentation algorithms is discriminating the vague edges of the tumors owing to meager contrast between tumors and the liver parenchyma. We proposed a CE technique (in **Paper B**) that not only discriminates the tumor border from the rest of the organ but also maintains the uniformity of the tumor area. Consequently, the segmentation algorithm works well on the enhanced images using the proposed approach compared to other similar single image enhancement approaches. Quantitative assessment from both perspectives (enhancement and segmentation) validate the superior performance of our proposed algorithm, this is further supported by qualitative assessment. Moreover, the execution time of the CE technique proposed in **Paper A** was accelerated using GPU (parallelization of the tasks in **Paper D**).

Deep learning was also exploited in this dissertation to achieve contrast enhance-

ment (using cross-modal guidance information) in the absence of paired ground truth (**Paper E**). It is important to mention here that cross-modal information was exploited for the first time in this work to improve the contrast of medical images (using both traditional and deep learning approaches), to the best of our knowledge.

In the second part of the thesis, our research work was directed towards eliminating noise in T1-w MR images using guidance from the corresponding T2-w images. In particular, a sophisticated mechanism of combining information from both images at a hierarchical feature-level was proposed. Our results outperformed some traditional as well as deep learning methods that do not incorporate cross-modal information. The results of this research study were reported in the form of a journal paper (**Paper F**).

To conclude, we implemented contrast enhancement and denoising methods in this thesis; these two pre-processing tasks play a significant role in surgical planning and navigation. As the proposed contrast enhancement exploiting redundancy and complementary information in multi-modal medical images is implemented for the first time, it is anticipated that the idea can be further extended and improved. Several open questions could not be investigated in this work due to constrained time. Some of these unaddressed aspects of this research are discussed in Sec. 4.2 that can be investigated in the future. We believe that this Ph.D. work can provide a good basis for the comprehensive study of cross-modal medical image contrast enhancement and denoising in the framework of pre-operative planning and surgical navigation.

Bibliography

[1] H. G. Kenngott, M. Wagner, M. Gondan, F. Nickel, M. Nolden, A. Fetzer, J. Weitz, L. Fischer, S. Speidel, H.-P. Meinzer *et al.*, "Real-time image guidance in laparoscopic liver surgery: first clinical experience with a guidance system based on intraoperative ct imaging," *Surgical endoscopy*, vol. 28, no. 3, pp. 933–940, 2014.

[2] Q. Yan, X. Shen, L. Xu, S. Zhuo, X. Zhang, L. Shen, and J. Jia, "Cross-field joint image restoration via scale map," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1537–1544.

[3] Y. Li, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Joint image filtering with deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 8, pp. 1909–1923, 2019.

[4] R. Naseem, F. A. Cheikh, A. Beghdadi, O. J. Elle, and F. Lindseth, "Cross modality guided liver image enhancement of ct using mri," in *2019 8th European Workshop on Visual Information Processing (EUVIP)*, 2019, pp. 46–51.

[5] Z. Al-Ameen and G. Sulong, "A new algorithm for improving the low contrast of computed tomography images using tuned brightness controlled single-scale retinex." *Scanning*, vol. 37 2, pp. 116–25, 2015.

[6] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics gems*, pp. 474–485, 1994.

[7] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. IEEE, 2005, pp. 60–65.

[8] F. Luisier, T. Blu, and M. Unser, "A new sure approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Transactions on image processing*, vol. 16, no. 3, pp. 593–606, 2007.

[9] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[10] D. Jiang, W. Dou, L. Vosters, X. Xu, Y. Sun, and T. Tan, "Denoising of 3d magnetic resonance images with multi-channel residual learning of convolutional neural network," *Japanese journal of radiology*, vol. 36, no. 9, pp. 566–574, 2018.

[11] A. Yasmin, D. P. Regan, L. B. Schook, R. C. Gaba, and K. M. Schachtschneider, "Transcriptional regulation of alcohol induced liver fibrosis in a translational porcine hepatocellular carcinoma model," *Biochimie*, vol. 182, pp. 73–84, 2021.

[12] D. Anwanwan, S. K. Singh, S. Singh, V. Saikam, and R. Singh, "Challenges in liver cancer and possible treatment approaches," *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, vol. 1873, no. 1, p. 188314, 2020.

[13] K. T. Nguyen, J. W. Marsh, A. Tsung, J. J. L. Steel, T. C. Gamblin, and D. A. Geller, "Comparative benefits of laparoscopic vs open hepatic resection: a critical appraisal," *Archives of surgery*, vol. 146, no. 3, pp. 348–356, 2011.

[14] X. Cai, "Laparoscopic liver resection: the current status and the future," *Hepatobiliary surgery and nutrition*, vol. 7, no. 2, p. 98, 2018.

[15] Å. A. Fretland, V. J. Dagenborg, G. Waaler Bjørnelv, D. Aghayan, A. M. Kazaryan, L. Barkhatov, R. Kristiansen, M. Fagerland, B. Edwin, and M. H. Andersen, "Quality of life from a randomized trial of laparoscopic or open liver resection for colorectal liver metastases," *British Journal of Surgery*, vol. 106, no. 10, pp. 1372–1380, 2019.

[16] Å. A. Fretland, V. J. Dagenborg, G. M. W. Bjørnelv, A. M. Kazaryan, R. Kristiansen, M. W. Fagerland, J. Hausken, T. I. Tønnessen, A. Abildgaard, L. Barkhatov *et al.*, "Laparoscopic versus open resection for colorectal liver metastases," *Annals of surgery*, vol. 267, no. 2, pp. 199–207, 2018.

[17] K. Cleary and T. M. Peters, "Image-guided interventions: technology review and clinical applications," *Annual review of biomedical engineering*, vol. 12, pp. 119–142, 2010.

[18] U. Mezger, C. Jendrewski, and M. Bartels, "Navigation in surgery," *Langenbeck's archives of surgery*, vol. 398, no. 4, pp. 501–514, 2013.

[19] H. P. S. T. Navigation. (hipernav). [Online]. Available: https://hipernav.eu/

[20] Z. Morise, N. Kawabe, H. Tomishige, H. Nagata, J. Kawase, S. Arakawa, R. Yoshida, and M. Isetani, "Recent advances in liver resection for hepatocellular carcinoma," *Frontiers in surgery*, vol. 1, p. 21, 2014.

[21] L. Bonanni, N. d. Carino, R. Deshpande, B. Ammori, D. Sherlock, J. W. Valle, E. Tam, and D. O'Reilly, "A comparison of diagnostic imaging modalities for colorectal liver metastases," *European Journal of Surgical Oncology*, vol. 40, no. 5, pp. 545–550, 2014.

[22] J.-H. Zhong, A. C. Rodríguez, Y. Ke, Y.-Y. Wang, L. Wang, and L.-Q. Li, "Hepatic resection as a safe and effective treatment for hepatocellular carcinoma involving a single large tumor, multiple tumors, or macrovascular invasion," *Medicine*, vol. 94, no. 3, 2015.

[23] R. J. Aragon and N. L. Solomon, "Techniques of hepatic resection," *Journal of gastrointestinal oncology*, vol. 3, no. 1, p. 28, 2012.

[24] H. Kaneko, S. Takagi, Y. Otsuka, M. Tsuchiya, A. Tamura, T. Katagiri, T. Maeda, and T. Shiba, "Laparoscopic liver resection of hepatocellular carcinoma," *The American journal of surgery*, vol. 189, no. 2, pp. 190–194, 2005.

[25] E. Vibert, T. Perniceni, H. Levard, C. Denet, N. Shahri, and B. Gayet, "Laparoscopic liver resection," *British Journal of Surgery*, vol. 93, no. 1, pp. 67–72, 2006.

[26] M. Kleemann, S. Deichmann, H. Esnaashari, A. Besirevic, O. Shahin, H.-P. Bruch, and T. Laubert, "Laparoscopic navigated liver resection: technical aspects and clinical practice in benign liver tumors," *Case reports in surgery*, vol. 2012, 2012.

[27] C. Askeland, O. V. Solberg, J. B. L. Bakeng, I. Reinertsen, G. A. Tangen, E. F. Hofstad, D. H. Iversen, C. Våpenstad, T. Selbekk, T. Langø *et al.*, "Custusx: an open-source research platform for image-guided therapy," *International journal of computer assisted radiology and surgery*, vol. 11, no. 4, pp. 505–519, 2016.

[28] H. Liao, N. Hata, S. Nakajima, M. Iwahara, I. Sakuma, and T. Dohi, "Surgical navigation by autostereoscopic image overlay of integral videography," *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 2, pp. 114–121, 2004.

[29] E. H. Burrows, "Pioneers and early years. a history of british radiology," 1986.

[30] B. Preim and C. P. Botha, *Visual computing for medicine: theory, algorithms, and applications*. Newnes, 2013.

[31] M. Zijlmans, T. Langø, E. F. Hofstad, C. F. Van Swol, and A. Rethy, "Navigated laparoscopy–liver shift and deformation due to pneumoperitoneum in an animal model," *Minimally Invasive Therapy & Allied Technologies*, vol. 21, no. 3, pp. 241–248, 2012.

[32] A. Teatini, E. Pelanis, D. L. Aghayan, R. P. Kumar, R. Palomar, Å. A. Fretland, B. Edwin, and O. J. Elle, "The effect of intraoperative imaging on surgical navigation for laparoscopic liver resection surgery," *Scientific reports*, vol. 9, no. 1, pp. 1–11, 2019.

[33] E. C. Beckmann, "Ct scanning the early days," *The British journal of radiology*, vol. 79, no. 937, pp. 5–8, 2006.

[34] E. M. Haacke, R. W. Brown, M. R. Thompson, R. Venkatesan *et al.*, *Magnetic resonance imaging: physical principles and sequence design*. Wiley-liss New York:, 1999, vol. 82.

[35] V. Chan and A. Perlas, "Basics of ultrasound imaging," in *Atlas of ultrasound-guided procedures in interventional pain management*. Springer, 2011, pp. 13–19.

[36] A. Maier, S. Steidl, V. Christlein, and J. Hornegger, "Medical imaging systems: An introductory guide," 2018.

[37] S.-W. Jung, "Two-dimensional histogram specification using two-dimensional cumulative distribution function," *Electronics letters*, vol. 50, no. 12, pp. 872–874, 2014.

[38] A. Beghdadi and A. Le Negrate, "Contrast enhancement technique based on local detection of edges," *Computer vision, graphics, and image processing*, vol. 46, no. 2, pp. 162–174, 1989.

[39] A. Saleem, A. Beghdadi, and B. Boashash, "Image fusion-based contrast enhancement," *EURASIP Journal on Image and Video Processing*, vol. 2012, no. 1, pp. 1–17, 2012.

[40] D. Coltuc, P. Bolon, and J.-M. Chassery, "Exact histogram specification," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1143–1152, 2006.

[41] Z. Chen, B. R. Abidi, D. L. Page, and M. A. Abidi, "Gray-level grouping (glg): an automatic method for optimized image contrast enhancement-part i: the basic method," *IEEE transactions on image processing*, vol. 15, no. 8, pp. 2290–2302, 2006.

[42] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Transactions on image processing*, vol. 9, no. 5, pp. 889–896, 2000.

[43] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1–8, 1997.

[44] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Transactions on Consumer Electronics*, vol. 45, no. 1, pp. 68–75, 1999.

[45] E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, and S. M. Pizer, "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," *Journal of Digital imaging*, vol. 11, no. 4, p. 193, 1998.

[46] A. N. Avanaki, "Exact global histogram specification optimized for structural similarity," *Optical review*, vol. 16, no. 6, pp. 613–621, 2009.

[47] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[48] A. Kumar and M. Ansari, "Performance evaluation of de-noised medical images after removing speckled noise by wavelet transform," *International Journal of Biomedical Engineering and Technology*, vol. 36, no. 4, pp. 318–330, 2021.

[49] E. K. Wang, C.-M. Chen, M. M. Hassan, and A. Almogren, "A deep learning based medical image segmentation technique in internet-of-medical-things domain," *Future Generation Computer Systems*, vol. 108, pp. 135–144, 2020.

[50] X. Yi and P. Babyn, "Sharpness-aware low-dose ct denoising using conditional generative adversarial network," *Journal of digital imaging*, vol. 31, no. 5, pp. 655–669, 2018.

[51] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1348–1357, 2018.

[52] S. O. Rice, "Mathematical analysis of random noise," *The Bell System Technical Journal*, vol. 23, no. 3, pp. 282–332, 1944.

[53] H. Gudbjartsson and S. Patz, "The rician distribution of noisy mri data," *Magnetic resonance in medicine*, vol. 34, no. 6, pp. 910–914, 1995.

[54] L. He and I. R. Greenshields, "A nonlocal maximum likelihood estimation method for rician noise reduction in mr images," *IEEE transactions on medical imaging*, vol. 28, no. 2, pp. 165–172, 2008.

[55] C. Pal, P. Das, A. Chakrabarti, and R. Ghosh, "Rician noise removal in magnitude mri images using efficient anisotropic diffusion filtering," *International Journal of Imaging Systems and Technology*, vol. 27, no. 3, pp. 248–264, 2017.

[56] B. Stimpel, C. Syben, F. Schirrmacher, P. Hoelter, A. Dörfler, and A. Maier, "Multimodal deep guided filtering for comprehensible medical image processing," *IEEE transactions on medical imaging*, vol. 39, no. 5, pp. 1703–1711, 2019.

[57] W. R. Gonzalez RC, *Digital image processing, 3rd edn.* Prentice- Hall, Inc, Upper Saddle River, 2006.

[58] A. Beghdadi and A. Khellaf, "A noise-filtering method using a local information measure," *IEEE Transactions on Image Processing*, vol. 6, no. 6, pp. 879–882, 1997.

[59] K. Belkacem-Boussaid and A. Beghdadi, "A new image smoothing method based on a simple model of spatial processing in the early stages of human vision," *IEEE Transactions on Image Processing*, vol. 9, no. 2, pp. 220–226, 2000.

[60] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[61] J. Monteil and A. Beghdadi, "A new interpretation and improvement of the nonlinear anisotropic diffusion for image enhancement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 940–946, 1999.

[62] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*. IEEE, 1998, pp. 839–846.

[63] W. Souidene, A. Beghdadi, and K. Abed-Meraim, "Image denoising in the transformed domain using non local neighborhoods," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 2. IEEE, 2006, pp. II–II.

[64] Q. B. Do, A. Beghdadi, and M. Luong, "Image denoising using bilateral filter in high dimensional pca-space," in *International Conference on Computer Analysis of Images and Patterns*. Springer, 2011, pp. 372–379.

[65] Q. B. Do, A. Beghdadi, and M. Luong, "Combination of closest space and closest structure to ameliorate non-local means method," in *2011 IEEE Symposium On Computational Intelligence For Multimedia, Signal And Vision Processing*. IEEE, 2011, pp. 134–141.

[66] L. David and J. Donoho, "Denoising by soft-thresholding," *IEEE Transactions on information theory*, vol. 41, no. 3, pp. 613–627, 1995.

[67] K. He, J. Sun, and X. Tang, "Guided image filtering," in *European conference on computer vision*. Springer, 2010, pp. 1–14.

[68] D.-t. Huang, W.-q. Huang, P.-t. Gu, P.-z. Liu, and Y.-m. Luo, "Image super-resolution reconstruction based on regularization technique and guided filter," *Infrared Physics & Technology*, vol. 83, pp. 103–113, 2017.

[69] Z. Lu, B. Long, K. Li, and F. Lu, "Effective guided image filtering for contrast enhancement," *IEEE Signal Processing Letters*, vol. 25, no. 10, pp. 1585–1589, 2018.

[70] S. Zhuo, X. Zhang, X. Miao, and T. Sim, "Enhancing low light images using near infrared flash images," in *2010 IEEE International Conference on Image Processing*. IEEE, 2010, pp. 2537–2540.

[71] X. Deng and P. L. Dragotti, "Deep coupled ista network for multi-modal image super-resolution," *IEEE Transactions on Image Processing*, vol. 29, pp. 1683–1698, 2019.

[72] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," in *ACM SIGGRAPH 2007 papers*, 2007, pp. 1–es.

[73] X. Shen, C. Zhou, L. Xu, and J. Jia, "Mutual-structure for joint filtering," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 3406–3414.

[74] X. Zhang, T. Sim, and X. Miao, "Enhancing photographs with near infra-red images," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.

[75] F. Johannesdottir, B. Allaire, D. Kopperdahl, T. Keaveny, S. Sigurdsson, M. Bredella, D. Anderson, E. Samelson, D. Kiel, V. Gudnason *et al.*, "Bone density and strength from thoracic and lumbar ct scans both predict incident vertebral fractures independently of fracture location," *Osteoporosis International*, vol. 32, no. 2, pp. 261–269, 2021.

[76] K. E. Fasmer, E. Hodneland, J. A. Dybvik, K. Wagner-Larsen, J. Trovik, Ø. Salvesen, C. Krakstad, and I. H. Haldorsen, "Whole-volume tumor mri radiomics for prognostic modeling in endometrial cancer," *Journal of Magnetic Resonance Imaging*, vol. 53, no. 3, pp. 928–937, 2021.

[77] I. J. Gerard, M. Kersten-Oertel, J. A. Hall, D. Sirhan, and D. L. Collins, "Brain shift in neuronavigation of brain tumors: An updated review of intra-operative ultrasound applications," *Frontiers in Oncology*, vol. 10, p. 3390, 2021.

[78] J. Xu, E. Gong, J. Ouyang, J. Pauly, and G. Zaharchuk, "Ultra-low-dose 18f-fdg brain pet/mr denoising using deep learning and multi-contrast information," in *Medical Imaging 2020: Image Processing*, vol. 11313. International Society for Optics and Photonics, 2020, p. 113131P.

[79] D. A. Torigian, H. Zaidi, T. C. Kwee, B. Saboury, J. K. Udupa, Z.-H. Cho, and A. Alavi, "Pet/mr imaging: technical aspects and potential clinical applications," *Radiology*, vol. 267, no. 1, pp. 26–44, 2013.

[80] H. Zaidi and A. Del Guerra, "An outlook on future design of hybrid pet/mri systems," *Medical physics*, vol. 38, no. 10, pp. 5667–5689, 2011.

[81] M. A. Rocca, M. Battaglini, R. H. Benedict, N. De Stefano, J. J. Geurts, R. G. Henry, M. A. Horsfield, M. Jenkinson, E. Pagani, and M. Filippi, "Brain mri atrophy quantification in ms: from methods to clinical application," *Neurology*, vol. 88, no. 4, pp. 403–413, 2017.

[82] F. Agosta, S. Galantucci, and M. Filippi, "Advanced magnetic resonance imaging of neurodegenerative diseases," *Neurological sciences*, vol. 38, no. 1, pp. 41–51, 2017.

[83] R. Prieto, J. Pascual, and L. Barrios, "Topographic diagnosis of craniopharyngiomas: the accuracy of mri findings observed on conventional t1 and t2 images," *American Journal of Neuroradiology*, vol. 38, no. 11, pp. 2073–2080, 2017.

[84] M. Ganzetti, N. Wenderoth, and D. Mantini, "Mapping pathological changes in brain structure by combining t1-and t2-weighted mr imaging data," *Neuroradiology*, vol. 57, no. 9, pp. 917–928, 2015.

[85] A. Teramoto, H. Fujita, O. Yamamuro, and T. Tamaki, "Automated detection of pulmonary nodules in pet/ct images: Ensemble false-positive reduction using a convolutional neural network technique," *Medical physics*, vol. 43, no. 6Part1, pp. 2821–2827, 2016.

[86] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, "Deep learning-based image segmentation on multimodal medical imaging," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 162–169, 2019.

[87] Z. Guo, N. Guo, K. Gong, Q. Li *et al.*, "Gross tumor volume segmentation for head and neck cancer radiotherapy using deep dense multi-modality network," *Physics in Medicine & Biology*, vol. 64, no. 20, p. 205015, 2019.

[88] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Classification in BioApps*, pp. 323–350, 2018.

[89] B. IXI. Brain ixi database. [Online]. Available: https://brain-development.org/team/

[90] C. I. Archive. Cancer imaging archive. [Online]. Available: https://www.cancerimagingarchive.net/

[91] M. Gaur, K. Faldu, and A. Sheth, "Semantics of the black-box: Can knowledge graphs help make deep learning systems more interpretable and explainable?" *IEEE Internet Computing*, vol. 25, no. 1, pp. 51–59, 2021.

[92] J. Tan, C. Wang, B. Li, Q. Li, W. Ouyang, C. Yin, and J. Yan, "Equalization loss for long-tailed object recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11 662–11 671.

[93] G. Brasó and L. Leal-Taixé, "Learning a neural solver for multiple object tracking," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 6247–6257.

[94] N. Araslanov and S. Roth, "Single-stage semantic segmentation from image labels," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 4253–4262.

[95] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "Enlightengan: Deep light enhancement without paired supervision," *IEEE Transactions on Image Processing*, vol. 30, pp. 2340–2349, 2021.

[96] L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma, "Msr-net: Low-light image enhancement using deep convolutional network," *arXiv preprint arXiv:1711.02488*, 2017.

[97] Y. Li, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep joint image filtering," in *European Conference on Computer Vision*. Springer, 2016, pp. 154–169.

[98] X.-F. Mei, F.-Y. Xie, and Z.-G. Jiang, "Uneven illumination removal based on fully convolutional network for dermoscopy images," in *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. IEEE, 2016, pp. 243–247.

[99] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose ct with a residual encoder-decoder convolutional neural network," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[100] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *arXiv preprint arXiv:1406.2661*, 2014.

[101] C. You, Q. Yang, H. Shan, L. Gjesteby, G. Li, S. Ju, Z. Zhang, Z. Zhao, Y. Zhang, W. Cong *et al.*, "Structurally-sensitive multi-scale deep neural network for low-dose ct denoising," *IEEE Access*, vol. 6, pp. 41 839–41 855, 2018.

[102] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[103] Y. Ma, Y. Liu, J. Cheng, Y. Zheng, M. Ghahremani, H. Chen, J. Liu, and Y. Zhao, "Cycle structure and illumination constrained gan for medical image enhancement," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 667–677.

[104] O. Sidorov, C. Wang, and F. A. Cheikh, "Generative smoke removal," in *Machine Learning for Health Workshop*. PMLR, 2020, pp. 81–92.

[105] K. Armanious, C. Jiang, M. Fischer, T. Küstner, T. Hepp, K. Nikolaou, S. Gatidis, and B. Yang, "Medgan: Medical image translation using gans," *Computerized medical imaging and graphics*, vol. 79, p. 101684, 2020.

[106] D. Mahapatra, B. Bozorgtabar, and R. Garnavi, "Image super-resolution using progressive generative adversarial networks for medical image analysis," *Computerized Medical Imaging and Graphics*, vol. 71, pp. 30–39, 2019.

[107] D. Peng, X. Yu, W. Peng, and J. Lu, "Dgfau-net: Global feature attention upsampling network for medical image segmentation," *Neural Computing and Applications*, pp. 1–15, 2021.

[108] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose ct," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.

[109] D. Wallach and B. Goffinet, "Mean squared error of prediction as a criterion for evaluating and comparing system models," *Ecological modelling*, vol. 44, no. 3-4, pp. 299–306, 1989.

[110] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[111] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[112] M. Dunnhofer, M. Antico, F. Sasazawa, Y. Takeda, S. Camps, N. Martinel, C. Micheloni, G. Carneiro, and D. Fontanarosa, "Siam-u-net: encoder-decoder siamese network for knee cartilage tracking in ultrasound images," *Medical image analysis*, vol. 60, p. 101631, 2020.

[113] A. Rossi, M. Hosseinzadeh, M. Bianchini, F. Scarselli, and H. Huisman, "Multimodal siamese network for diagnostically similar lesion retrieval in prostate mri," *IEEE Transactions on Medical Imaging*, 2020.

[114] Y. Chang, C. Jung, J. Sun, and F. Wang, "Siamese dense network for reflection removal with flash and no-flash image pairs," *International Journal of Computer Vision*, pp. 1–26, 2020.

[115] R. Naseem, F. A. Cheikh, A. Beghdadi, O. J. Elle, and F. Lindseth, "Cross modality guided liver image enhancement of ct using mri," in *2019 8th European Workshop on Visual Information Processing (EUVIP)*. IEEE, 2019, pp. 46–51.

[116] M. A. Qureshi, M. Deriche, A. Beghdadi, and M. Mohandes, "An information based framework for performance evaluation of image enhancement methods," in *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*. IEEE, 2015, pp. 519–523.

[117] Z. A. Khan, A. Beghdadi, F. A. Cheikh, M. Kaaniche, and M. A. Qureshi, "A multi-criteria contrast enhancement evaluation measure using wavelet decomposition," in *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2020, pp. 1–6.

[118] R. Naseem, Z. A. Khan, N. Satpute, A. Beghdadi, F. A. Cheikh, and J. Olivares, "Cross-modality guided contrast enhancement for improved liver tumor image segmentation," *IEEE Access*, 2021.

[119] N. Satpute, R. Naseem, R. Palomar, O. Zachariadis, J. Gómez-Luna, F. A. Cheikh, and J. Olivares, "Fast parallel vessel segmentation," *Computer methods and programs in biomedicine*, vol. 192, p. 105430, 2020.

[120] X.-F. Wang, D.-S. Huang, and H. Xu, "An efficient local chan–vese model for image segmentation," *Pattern Recognition*, vol. 43, no. 3, pp. 603–618, 2010.

[121] Y. Cheng, X. Hu, J. Wang, Y. Wang, and S. Tamura, "Accurate vessel segmentation with constrained b-snake," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2440–2455, 2015.

[122] N. Satpute, R. Naseem, E. Pelanis, J. Gómez-Luna, F. A. Cheikh, O. J. Elle, and J. Olivares, "Gpu acceleration of liver enhancement for tumor segmentation," *Computer methods and programs in biomedicine*, vol. 184, p. 105285, 2020.

[123] G. Rai and T. Nair, "Gradient based seeded region grow method for ct angiographic image segmentation," *arXiv preprint arXiv:1001.3735*, 2010.

[124] Y.-S. Chen, Y.-C. Wang, M.-H. Kao, and Y.-Y. Chuang, "Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6306–6314.

[125] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[126] R. Naseem, F. A. Cheikh, A. Beghdadi, K. Muhammad, and M. Sajjad, "Cross-modal guidance assisted hierarchical learning based siamese network for mr image denoising," *Electronics*, vol. 10, no. 10, p. 2855, 2021.

[127] C. Qin, B. Shi, R. Liao, T. Mansi, D. Rueckert, and A. Kamen, "Unsupervised deformable registration for multi-modal images via disentangled representations," in *International Conference on Information Processing in Medical Imaging*. Springer, 2019, pp. 249–261.

[128] Z. A. Khan, M. Kaaniche, A. Beghdadi, and F. A. Cheikh, "Joint statistical models for no-reference stereoscopic image quality assessment," in *2018 7th European Workshop on Visual Information Processing (EUVIP)*. IEEE, 2018, pp. 1–5.

[129] C. K. Wong, S. Liu, S. C. Liu, M. B. Rahman, S. Lin, G. Jiang, N. Kwok, and H. Shi, "Image contrast enhancement using histogram equalization with maximum intensity coverage," *Journal of Modern Optics*, vol. 63, pp. 1618 – 1629, 2016.

[130] S. Lin, C. Y. Wong, M. A. Rahman, G. Jiang, S. Liu, N. Kwok, H. Shi, Y. Yu, and T. Wu, "Image enhancement using the averaging histogram equalization (avheq) approach for contrast improvement and brightness preservation," *Comput. Electr. Eng.*, vol. 46, pp. 356–370, 2015.

[131] C. Wang, A. K. Mohammed, F. A. Cheikh, A. Beghdadi, and O. J. Elle, "Multiscale deep desmoking for laparoscopic surgery," in *Medical Imaging 2019: Image Processing*, vol. 10949. International Society for Optics and Photonics, 2019, p. 109491Y.

[132] S. Bolkar, C. Wang, F. A. Cheikh, and S. Yildirim, "Deep smoke removal from minimally invasive surgery videos," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 3403–3407.

[133] C. Wang, F. A. Cheikh, A. Beghdadi, and O. J. Elle, "Adaptive context encoding module for semantic segmentation," *Electronic Imaging*, vol. 2020, no. 10, pp. 27–1, 2020.

[134] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[135] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[136] J. Gu and J. C. Ye, "Adain-based tunable cyclegan for efficient unsupervised low-dose ct denoising," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 73–85, 2021.

Part II

Original Articles

Chapter 6

Paper A: Cross modality guided liver image enhancement of CT using MRI

Chapter 7

Paper B: Cross-modality guided contrast enhancement for improved liver tumor image segmentation

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 4, 2021, accepted August 10, 2021, date of publication August 24, 2021, date of current version August 31, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3107473

Cross-Modality Guided Contrast Enhancement for Improved Liver Tumor Image Segmentation

RABIA NASEEM[1], ZOHAIB AMJAD KHAN[2], (Student Member, IEEE), NITIN SATPUTE[3], AZEDDINE BEGHDADI[2], (Senior Member, IEEE), FAOUZI ALAYA CHEIKH[1], (Senior Member, IEEE), AND JOAQUÍN OLIVARES[4]

[1]Norwegian Colour and Visual Computing Laboratory, Norwegian University of Science and Technology, 7491 Gjøvik, Norway
[2]Laboratory of Information Processing and Transmission (L2TI), Institut Galilée, Université Sorbonne Paris Nord, 93430 Villetaneuse, France
[3]Department of Electrical and Computer Engineering, Aarhus University, 8000 Aarhus, Denmark
[4]Department of Electronic and Computer Engineering, Maimonides Biomedical Research Institute of Cordoba (IMIBIC), Universidad de Córdoba, 14071 Córdoba, Spain

Corresponding author: Rabia Naseem (rabia.naseem@ntnu.no)

This work was supported by the Project High Performance soft tissue Navigation (HiPerNav) through the European Union Horizon 2020 Research and Innovation Program under Grant 722068.

ABSTRACT Tumor segmentation in Computed Tomography (CT) images is a crucial step in image-guided surgery. However, low-contrast CT images impede the performance of subsequent segmentation tasks. Contrast enhancement is then used as a preprocessing step to highlight the relevant structures, thus facilitating not only medical diagnosis but also image segmentation with higher accuracy. In this paper, we propose a goal-oriented contrast enhancement method to improve tumor segmentation performance. The proposed method is based on two concepts, namely guided image enhancement and image quality control through an optimization scheme. The proposed OPTimized Guided Contrast Enhancement (OPTGCE) scheme exploits both contextual information from the guidance image and structural information from the input image in a two-step process. The first step consists of applying a two-dimensional histogram specification exploiting contextual information in the corresponding guidance image, i.e. Magnetic Resonance Image (MRI). In the second step, an optimization scheme using a structural similarity measure to preserve the structural information of the original image is performed. To the best of our knowledge, this kind of contrast enhancement optimization scheme using cross-modal guidance is proposed for the first time in the medical imaging context. The experimental results obtained on real data demonstrate the effectiveness of the method in terms of enhancement and segmentation quality in comparison to some state-of-the-art methods based on the histogram.

INDEX TERMS Guided enhancement, cross-modality, contrast enhancement, 2D histogram specification (HS), SSIM gradient, tumor segmentation.

I. INTRODUCTION

Liver cancer is the fifth most prevalent cancer in the world, carrying a low survival rate [1]. Nevertheless, timely detection of cancerous tumors and effective treatment strategies can improve the overall survival rate. Diagnostic imaging techniques such as CT facilitate timely diagnosis of cancer; however, low contrast and noise limit their utility [2]. Moreover, such low-contrast images make segmentation and tumor detection challenging problems that can be overcome by applying a contrast enhancement beforehand.

It is also worth mentioning here that a single medical imaging modality is unable to capture all the relevant structural information from the organs. For this reason, it is now becoming more common to acquire both CT and MR images periodically during liver cancer diagnosis and treatment [3]. Therefore, it would be interesting to use the additional captured information from one imaging modality (e.g. MRI) to enhance the other (e.g. CT). The concept of enhancing the image from one modality using cross-modal image

The associate editor coordinating the review of this manuscript and approving it for publication was Humaira Nisar.

information is not novel; similar ideas have been successfully applied to natural images [4]–[6]. One such approach for liver CT image enhancement using the corresponding MR images was proposed to improve the visibility of tumors and vessels [7]. In general, the cross-modality guided enhancement methods have shown better performance in comparison with the classic single image enhancement methods [8], [9].

Currently, there are two main challenges related to image enhancement in the medical context. Firstly, most recent enhancement techniques are tailored to only specific types of images. Secondly, it is not easy to find a well-established benchmark for evaluating the existing enhancement methods. For these reasons, the effectiveness of the enhancement approaches is often assessed based on their impact on the underlying application. For medical imaging, the motivation of CE, in general, is to improve the visual appearance of relevant organ structures for better diagnosis and intervention [10]. However, limited research has been done on image quality enhancement to improve the segmentation of such organ structures [4], [11]–[14]. By using CE as preprocessing step, improved segmentation of relevant structures in CT images could be achieved as concluded in [13]. Therefore, there is a dire need for efficient CE algorithms for such images.

Traditional enhancement methods suffer from limitations such as saturation, over-enhancement, and uneven contrast spatial distribution, that may result from the uncontrolled CE process. One way to overcome such limitations is to combine the contrast enhancement approach with a quality control scheme. Inspired by the guided filtering approach and the simplicity of context-aware histogram-based image quality enhancement, we propose in this paper a cross-modality guided histogram specification technique to improve the contrast of liver CT images using MRI images as guiding input data. Furthermore, optimization is incorporated to prevent the saturation artifacts inherent to histogram-based methods. A similar idea was proposed for enhancing natural images in [15]. It consists of mapping the histogram of the input image to that of a reference image combined with an optimization technique to preserve the structures of the input image. In this work, we propose a similar approach for medical images using cross-modal information. The new CE approach is hence based on two concepts, namely, cross-modality-guided medical image enhancement to improve the global contrast, and quality control to preserve the local structures during enhancement. Here, we formulate the cross-modal CE as an optimization problem, where the gradient of structural similarity index measure (SSIM) is used for local structure preservation and minimizing artifacts introduced during enhancement [16]. Later, the role of CE is analyzed in facilitating tumor segmentation. The overall processing scheme is evaluated on a real dataset containing CT and MRI of the human liver with segmentation ground truth. The main contributions of this paper are:

- The two-dimensional histogram specification-based CE process is formulated as an optimization problem and extended to multi-modal medical imaging data for the first time.
- SSIM gradient is incorporated in the optimized cross-modality guided 2D-HS framework to preserve structural fidelity of the enhanced image with the original image while applying enhancement.
- In order to obtain the objective of contrast enhancement without affecting the important structures of the image, the algorithm achieves a nice balance between retaining structural similarity with input image (by integrating SSIM gradient) and enhancing contrast by employing 2D entropy. The suggested combination of cross-modal guidance and quality control enhances the CT image exploiting contextual information, as opposed to context-unaware schemes.
- A new goal-oriented performance evaluation of the proposed approach is done utilizing objective quality metrics and through segmentation results applied on real multi-modal liver data. Comparison with single enhancement techniques validate the superior performance of the proposed method.

The rest of the paper is organized as follows. Section II provides a brief review of relevant contrast enhancement methods. Section III describes the proposed Optimized guided CE method. Experimental results of CE are discussed in section IV. The results of applying segmentation on the enhanced images are described in section V, followed by conclusion in section VI.

II. RELATED WORK

Among the methods for image quality enhancement, CE is the most intuitive and widely used solution in various fields of application, especially in medical imaging. Contrast enhancement methods can be broadly classified into two main categories: direct methods and indirect methods [17]. In the first category, the contrast is first defined and amplified to then deduce the modified value associated with the pixel to be treated. In the second category, the pixel value is transformed by means of an operation defined from global characteristics, such as the distribution of pixel values, or local characteristics such as the intensity gradient of the pixel. These operations can be performed in the spatial or transform domain or even in the joint spatial/spatial-frequency domain. It is well known that methods that operate in the transform domain and more particularly those that exploit multi-scale aspects and directional selectivity are more efficient but at the cost of increased complexity and prohibitive computation time in the case of large volumes of image data. It is therefore natural to turn to global methods and more specifically those that exploit the distribution of pixel values in the case of medical images. Among these methods, those based on histograms developed in general for natural images constitute an interesting alternative to other complex and time-consuming methods. Indeed, histogram-based CE approaches are well investigated both for natural as well as medical images, thanks to their low complexity and acceptable performance [18], [19]. For instance,

histogram equalization (HE) is one of the conventional histogram-based CE methods that map the global Cumulative Distribution Function (CDF) of the input image to that of a uniform distribution. Other histogram-based methods like Adaptive Histogram Equalization (AHE) and Contrast Limited Adaptive Histogram Equalization (CLAHE) operate on a small region around each pixel of the image to improve local contrast [20]. Besides these, another histogram-based approach, Histogram Specification (HS), uses CDF of the reference image with better perceptual quality to improve the visual appearance of a low contrast image [21], [22]. Similarly, another interesting histogram-based approach proposed for natural images uses contextual information through the 2D-CDF of the target image to exploit the inter-pixel correlation [23], [24]. This approach [23] has been shown to outperform other histogram-based approaches including minimum within-class variance multi-histogram equalization [25].

Among these histogram-based methods, some have also been used for medical image enhancement such as AHE and CLAHE for CE of FLuid Attenuated Inversion Recovery (FLAIR) MR images of brain [18]. In this method, the authors first performed contrast stretching before applying AHE/CLAHE. Eventually, to highlight abnormal hyperintense regions they have detected regional maxima followed by performing a local averaging of pixel values. However, their method is strictly application-specific and they have only used PSNR and average gradient for comparison, both of which are not suitable for performance evaluation of CE methods. Similarly, another histogram-based method to enhance CT images has combined normalized gamma correction with CLAHE to reduce the excessive brightness introduced by CLAHE [19]. In this work as well, the similarity metrics like SSIM [16] are inefficient for performance evaluation of CE methods, since there is no reference image in the CE task that can be used for similarity comparison. Few enhancement techniques [2], [26], [27] decompose the low-contrast image into detail and base layer separating high frequency and low-frequency contents in the image and apply histogram modification to enhance the base layer. The histogram is modified to prevent over-enhancement of the image by ensuring that the minimum gray level of the enhanced image approximates zero.

Despite the promising results of histogram-based methods in many applications, they sometimes introduce artifacts in the resulting image. Optimization strategies could be then used to reduce these side effects [28]. For instance, a HE-variant approach [29] finds an optimal threshold gray level value that separates the histogram into four sub-histogram. Histogram clipping is then applied to adjust the threshold according to the distribution of the original histogram. The proposed approach preserves naturalness of the image while minimizing over-enhancement. MedGA [30], another optimization-based technique was proposed to enhance the MR images. It is a combination of histogram-based method and genetic algorithm; however, only specific Region of Interest was enhanced in this way. Similarly, Lin *et al.* [31] achieved a nice balance between image brightness and contrast using a histogram averaging and remapping scheme. CE is conditioned by minimizing the average brightness difference between the input and enhanced images while maximizing the entropy in the processed image. However, this method is not well suited to dark images because of its inherent and fundamental brightness preserving property. Another method that maximizes information content and minimizes the artifacts has been proposed in [32]. The intermediate histogram equalized image and the original image are combined using a weighting factor computed by using a golden section search algorithm to attain uniform distribution. However, the local details are lost in case the image contains unevenly illuminated regions.

A breakthrough in the field of image enhancement was achieved when He *et al.* [33] introduced the idea to exploit information in a similar image to enhance the underlying image. This concept was further extended to cross-modality guided natural image enhancement [9]. The majority of cross-modality guided natural image enhancement methods process input images emanating from both modalities with the same contents and an accurate pixel to pixel correspondence, which largely simplifies the enhancement process [8]. Based on this strategy, Near Infrared (NIR) images were enhanced using photographs [34]. This method used gradient-based histogram matching to embed contrast of NIR images in photos. Moreover, wavelet domain processing was done to improve texture information. A similar idea was applied to medical images recently, where 2D HS was applied to map histogram of liver CT image to that of corresponding MR image [7]. However, optimization to control CE was not done in this work.

III. METHODOLOGY: OPTimized GUIDED CONTRAST ENHANCEMENT (OPTGCE)

Generally, classical Contrast Enhancement (CE) methods do not optimize an objective function or contrast-related measure; instead, they manipulate the pixel values according to a predefined distribution. Besides, these approaches amplify the contrast without objectively controlling the possible artifacts that may arise from the CE process. To the best of our knowledge, there are very few works where the contrast enhancement effect is controlled according to a well-defined framework. The proposed method OPTGCE operates according to this strategy. It applies HS-based CE to the low-contrast CT image based on the second-order distribution of an image of a complementary modality, that is MRI. The motivations behind the use of histogram-based methods are essentially their simplicity, reduced computational load, and the fact of exploiting a global statistical quantity that contains essential information on the distribution of pixel values. This is especially advantageous in the case of a large size of medical imaging data. Therefore, the 2D histogram effectively exploits the inter-pixel interactions, i.e. second-order statistics, in the design of the CE scheme.

This treatment enhances the overall contrast well but may suffer from some side effects. Indeed, the locally relevant structures of the image can be negatively affected, leading the processed image to be divergent from the original. Therefore, it is necessary to control the critical parameters of the CE process to amplify local contrast while simultaneously preserving the intrinsic structures of the image. One strategy to prevent the CE from side effects is to control the enhancement by using a local similarity measure between the input and enhanced image or some stopping criteria. Here, we perform the optimization using a measure that is directly related to the structural information in the image and carries contrast information. Furthermore, the extent of contrast is quantified through the two-dimensional entropy. The flowchart of the proposed technique is shown in Fig. 1. The three essential components of the proposed method, namely the 2D histogram specification-based CE, the structural gradient-based similarity measure, and 2D entropy are described below.

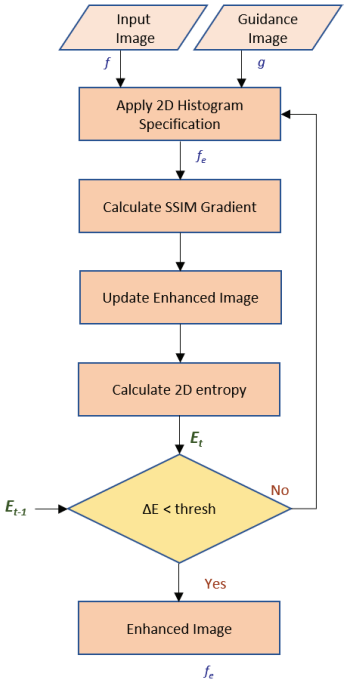


FIGURE 1. Flowchart of the proposed method.

A. 2D HISTOGRAM SPECIFICATION

Most image processing methods based on the distribution of pixel values involve only one-dimensional histogram. This has the disadvantage of not taking into account the strong spatial correlation of pixels and exploiting it in order to avoid side effects associated with histogram approaches. These limitations have led to the use of higher-order statistics of pixel values and characteristics to develop more efficient methods. The two-dimensional grayscale histogram is the simplest higher-order distribution. A two-dimensional histogram was then introduced in order to exploit the pixel inter-correlation in various image processing tasks such as image classification and grey-level thresholding [35], [36]. Later, this idea was also applied to 2D-HS and 2D-HE [23]. Indeed, it has been shown that the 2D HS clearly outperforms its one-dimensional HE and HS counterparts in terms of visual quality [23]. The approach in [23] is driven by the principle that the local and global contrast of the image could be enhanced by amplifying the grey-level transitions of neighboring pixels. One way to accentuate such transitions is to exploit the grey-level transition probability, i.e. 2D grey-level histogram. In recent work, 2D-HS based approach was applied to improve the contrast of liver CT images using MR images [7]. Although this method produces an acceptable enhancement, it is accompanied by a darkening or brightening effect in certain areas of the image due to the use of top hat and bottom hat transforms [37].

For the sake of completeness of the article, we recall here the basic notions and concepts introduced in the methods of 2D histogram specification. Let us consider an input image $[f] = \{f(m,n)|1 \leq m \leq M, 1 \leq n \leq N\}$, where $f(m,n)$ is the grey-level of pixel located at (m,n); the dynamic range of image is $[f_{min}, f_{max}]$ and $M \times N$ is its height and width respectively. The principle of guided contrast enhancement is to transform an input image $[f]$ into an output image $[f_e]$ to improve its contrast. The guide or reference image used in this process is represented by $[g]$, that is an image of better perceptual quality. One way to achieve this objective is to use the traditional Histogram Modification Framework and map the 1D-CDF of $[f]$ to that of $[g]$. However, as pointed out before, it is more efficient to consider pixel context, use higher-order statistics and compute the 2D-CDF instead when using the HS approach. Therefore, two-dimensional histograms of both guidance and input images, h_g and h_f respectively are derived from the Grey Level Co-occurrence Matrix (GLCM) computed from the two images. GLCM is a square matrix containing the number of occurrences of pairwise combinations of grey levels when exploring the whole image using a sliding window and a defined neighborhood. For the sake of simplicity, the neighborhood is generally restricted to the two nearest neighbors of the current pixel, i.e. left and above pixels. Let $f(m,n)$ denote the input image pixel's grey level. The GLCM is then computed as follows:

$$C_f(i,j) = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} \delta_{ij}(f(m,n), f(p,q)), \tag{1}$$

Here, i and j represent the pixel values and (m,n) and (p,q) represent the image coordinates, K is the total number of grey levels, and $0 \leq i,j \leq K-1$,

$$\delta_{ij}(a,b) = \begin{cases} 1, & \text{if } i = a \text{ and } j = b \\ 0, & \text{otherwise} \end{cases}$$

The transition probability of grey-levels, i.e. the 2D normalized histogram, is derived from the GLCM as follows:

$$h_f(i,j) = \frac{C_f(i,j)}{\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} C_f(i,j)} \tag{2}$$

The 2D-histogram is then used in the pixel grey-level mapping process using the histogram specification method as described below. This mapping process is based on the two-dimensional Cumulative Distribution Function (CDF) of the input and guidance images computed as follows.

$$H_f(i,j) = \sum_{i=0}^{K-1}\sum_{j=0}^{K-1} h_f(i,j) \tag{3}$$

The expression of the 2D-CDF of the guidance image is computed similarly and is represented as H_g. Once the 2D-CDF of both images is computed, the transformation T allowing the mapping between the input signal and the desired signal is obtained as follows:

$$T(i,j) = \arg\min_{[k,l]} |H_f(i,j) - H_g(k,l)| + \eta(|i-k| + |j-l|) \tag{4}$$

The mapping is accomplished by searching the target pixel values, $T(i,j)_1$ and $T(i,j)_2$ corresponding to pixel values i and j in $[f]$. The second term in Eq. 4, i.e. $|i-k| + |j-l|$ ensures to select a nearby pixel pair for which the difference between both CDFs among the candidate pixel values is minimized. η represents a very small number and its value $= 10^{-4}$.

The final step in 2D-HS consists of mapping the intensity values in $[f]$ to new values. To this end, each pixel and its immediate neighbor are considered. The intensity values for the enhanced image $[f_e]$ are therefore calculated using the equation below:

$$f_e(m,n) = T(f(m,n), f(m,n+1)) \tag{5}$$

From Eq. 5, it can be inferred that transformation of each value in the original image $[f]$ to a new value in the enhanced image $[f_e]$ also depends on its neighboring element. Therefore, unlike the 1D histogram specification which only considers individual pixel values for calculating the CDFs and ultimately mapping these values, this approach also exploits the contextual information among the pixels. Next, we look at the SSIM gradient approach.

B. GRADIENT BASED STRUCTURAL SIMILARITY MEASURE

As mentioned above, histogram specification is widely applied to enhance image contrast. However, like many transformation-based CE methods relying on global statistical descriptors, it affects local and global image structures. One way to control processing distortions is to integrate into the CE process a stopping criteria or an objective function and formulate the whole problem in a constrained optimization framework. The method proposed by Avanaki [15] belongs to this kind of solution. The idea is to apply global HS to a low-contrast image driven by an SSIM-based measure to control the enhancement through structural similarity changes between the original image and its enhanced variant. SSIM is a well-established measure to calculate the extent of similarity between two images [16]. Considering one image as a reference, the index provides the quality of the image under analysis in comparison with a reference. SSIM index is calculated between corresponding local blocks in images $[A]$ and $[B]$, after which the average of the values is taken to obtain a single value of SSIM as the overall similarity index. Let us assume that a_x and b_x represent corresponding blocks x in both images; μ_{a_x} and μ_{b_x} represent the mean intensity values of a_x and b_x and the standard deviations are given by σ_{a_x} and σ_{b_x}. C_1 and C_2 are small numbers greater than 0 to ensure the denominator is not zero. The SSIM between the two blocks a_x and b_x is then expressed as:

$$\mathrm{SSIM}(a_x,b_x) = \frac{(2\mu_{a_x}\mu_{b_x}+C_1)(2\sigma_{a_xb_x}+C_2)}{(\mu_{a_x}^2+\mu_{b_x}^2+C_1)(\sigma_{a_x}^2+\sigma_{b_x}^2+C_2)} \tag{6}$$

Few terms in Eq. 6 are described mathematically as:

$$\begin{aligned} \mu_{a_x} &= w * a_x, \\ \sigma_{a_xb_x} &= w * (a_x b_x) - \mu_{a_x}\mu_{b_x}, \\ \sigma_{a_x}^2 &= w * a_x^2 - \mu_{a_x}^2 \end{aligned} \tag{7}$$

where w is 11×11 Gaussian kernel and $*$ indicates convolution. Eq. 6 could be regarded as expression for SSIM index map, $\mathrm{SSIM_{map}}$ calculated via element wise addition and multiplication using parameters expressed in Eq. 7. Then, at all points, $\mathrm{SSIM_{map}}$ indicates local similarity between images $[A]$ and $[B]$. The global SSIM index for the overall images can then be expressed as:

$$\mathrm{SSIM}(A,B) = \frac{1}{Z}\sum_{\forall x} \mathrm{SSIM_{map}}(a_x, b_x; x) \tag{8}$$

where Z denotes the number of pixels in either image. $\mathrm{SSIM_{map}}(a_x, b_x; x)$ is SSIM index value corresponding to the window x of size c $\times$ c in images $[A]$ and $[B]$, starting from the upper left corner of images and proceeding to the bottom right. For the local SSIM measures in Eq. 6, we define the following terms for compactness:

$$\begin{aligned} \alpha_1(a_x,b_x) &= 2\mu_{a_x}\mu_{b_x} + C_1, \\ \alpha_2(a_x,b_x) &= 2\sigma_{a_xb_x} + C_2 \end{aligned} \tag{9a}$$

$$\begin{aligned} \beta_1(a_x,b_x) &= \mu_{a_x}^2 + \mu_{b_x}^2 + C_1, \\ \beta_2(a_x,b_x) &= \sigma_{a_x}^2 + \sigma_{b_x}^2 + C_2 \end{aligned} \tag{9b}$$

As discussed earlier, the SSIM gradient-based optimization method [15] is applied to the cross-modal medical image enhancement in this work. Here, 2D-HS is applied to enhance CT images by exploiting the better quality of MR images. When applied in the framework of optimization, the SSIM gradient refines the enhancement process incrementally.

The integration of SSIM ultimately preserves the overall morphology of the original image with minimal information

TABLE 1. Description of important notations used in the paper.

Symbol	Description
$f(m,n)$	Gray level in the input image $[f]$ at location (m,n)
$g(m,n)$	Gray level in the guidance image $[g]$ at location (m,n)
f_e	Enhanced output image
K	Total number of gray levels in images $[f]$ or $[g]$
M,N	Total number of rows and columns respectively in either image $[f]$ or $[g]$
$C_f(i,j)$	Co-occurrence matrix corresponding to image $[f]$
h_f	Probability Density Function (2D) associated with $[f]$
H_f	Cumulative Distribution Function (2D) associated with $[f]$
$T(i,j)$	Transformation Matrix
a_x, b_x	Neighborhood /Blocks of images $[A]$ and B respectively (size $c \times c$) defined for SSIM calculation
x	Window/ Block number in the images $[A]$ and $[B]$ for SSIM
$\text{SSIM}(A,B)$	SSIM value between images $[A]$ and $[B]$
$\partial_{f_e} \text{SSIM}(f, f_e)$	SSIM gradient of image $[f_e]$ with respect to image $[f]$
Z	Total number of pixels in either image $[f]$ or $[g]$ (both are same size)
μ_a	Mean intensity value of window a in image $[A]$
σ_a^2	Variance of block a in image $[A]$
E_t	Two-dimensional entropy of enhanced image $[f_e]$ at iteration t
t	Total number of times the proposed algorithm runs
ΔE	Change in entropy between successive iterations
$\Delta SSIM(t)$	Increase in SSIM value at iteration t
α	Step size used to update image using SSIM gradient

loss during enhancement. Here, we denote the input image as $[f]$ and the image whose structural similarity is being compared with $[f]$ as $[f_e]$; $[f_e]$ is obtained after applying 2D-HS. Now, to adapt the notion of SSIM gradient to our scenario, let us replace $[A]$ by $[f_e]$ and $[B]$ by $[f]$ and rewrite Eq. 8 as:

$$\text{SSIM}(f_e, f) = \frac{1}{Z} \sum_{\forall x} \text{SSIM}_{\text{map}}(f_{e_x}, f_x; x) \tag{10}$$

Calculating the derivative of Eq. 10 with reference to $[f_e]$ gives the SSIM gradient expression as follows:

$$\begin{aligned} &\partial_{f_e} \text{SSIM}(f_e, f) \\ &= \frac{2}{Z}\left[\left(w * \frac{\alpha_1}{\beta_1\beta_2}\right) f\right] \\ &\quad + \left(w * \frac{-\text{SSIM}_{\text{map}}}{\beta_2}\right) f_e \\ &\quad \times \left[+w * \frac{\mu_{f_e}(\alpha_2 - \alpha_1) - \mu_f(\beta_2 - \beta_1)\,\text{SSIM}_{\text{map}}}{\beta_1\beta_2}\right] \end{aligned} \tag{11}$$

where α_1, α_2, β_1 and β_2 have been described in Eq. 9a and 9b. Eq. 11 is a closed form solution and simple expression for SSIM gradient obtained by decomposing SSIM in linear terms. For thorough understanding of the mathematical computations, the reader is referred to [15]. The important notations used in this paper are listed in Table 1.

C. CONTRAST ENHANCEMENT WITH QUALITY CONTROL

After briefly introducing 2D-HS and SSIM gradient methods, the OPTGCE method is described in this subsection. Initially, we set the input CT image $[f]$ equal to $[f']$ and guidance MRI as $[g]$. The CDFs of $[f]$ and $[g]$ are calculated using Eq. 3. Eq. 4 mathematically defines how to calculate the transformation matrix T. The pixel values in $[f']$ are mapped to new values using Eq. 5 to get enhanced image $[f_e]$ in the manner explained in the section III-A. The proposed algorithm then calculates the structural similarity between $[f_e]$ and $[f]$ using Eq. 6 followed by SSIM gradient calculation with respect to image $[f_e]$, represented by ∂_{f_e}. Afterwards, $[f']$ is updated as mentioned in step 5 of the algorithm 1. Images $[f']$ and $[f_e]$ are updated in every iteration, while $[g]$ remains unchanged. Since SSIM computes the quality over the local neighborhood, it is capable of capturing local dissimilarities better than global approaches. Hence, optimizing the enhancement process using SSIM gradient offers better outcomes in terms of retaining the structure of the original image $[f]$. Algorithm 1 describes the steps of our proposed approach:

In the above algorithm, α represents the step size or the factor by which $[f_e]$ is updated in every iteration. Next, we describe a method to calculate the optimal value of α.

1) CALCULATION OF SUITABLE STEP SIZE

In this subsection, we elaborate the empirical approach similar to [15] for calculating an optimal step size α, so the algorithm attains higher SSIM in fewer iterations. The estimated increase in SSIM at iteration t is mathematically described as:

$$\Delta\,\text{SSIM}(t) = \alpha Z \sum_{\forall x} \left(\partial_{f_e} \text{SSIM}(f, f_e(t))\right)^2 \tag{12}$$

Based on the behavior of $SSIM(t)$ at several iterations, $\Delta\,\text{SSIM}(t)$ can be modeled by $\alpha r s^t$ [15]. The final value of SSIM (after several iterations) can be expressed as:

$$\text{SSIM}_{\text{f}} = \text{SSIM}' + \frac{r\alpha Z}{1-s} \tag{13}$$

Algorithm 1 OPTimized Guided Contrast Enhancement Algorithm (OPTGCE)

input CT image $= f$ and guidance image $= g$
Calculate 2D-CDF of guidance image as H_g and that of input image as H_f.
Set $f' = f$, $threshold = 0.05$ and $t = 1$
while $\Delta E >=$threshold **do**
1) Apply 2D histogram specification to f' to match 2D histogram of image g and generate enhanced image f_e.
2) Calculate structural similarity between f_e and f, SSIM(f, f_e) and SSIM gradient ∂_{f_e}SSIM(f, f_e)
3) Calculate E_t and ΔE.
4) Increment t as $t = t + 1$
5) Update f' contents using SSIM gradient driven factor as: $f' = f_e + \alpha Z \partial_{f_e}$SSIM$(f, f_e)$.

end while
Output enhanced image f_e

where $r = \sum_{\forall x} \left(\partial_{f_e} \text{SSIM}(f, f_e(1))\right)^2$, $s = \Delta\text{SSIM}(2)/\Delta\text{SSIM}(1)$, $\Delta\text{SSIM}(2)$ and $\Delta\text{SSIM}(1)$ denote the increase in SSIM values at $t = 2$ and $t = 1$ respectively. SSIM$'$ denotes the initial value of SSIM computed after first iteration. Our experiments show that SSIM value changes faster in earlier iterations, therefore the algorithm is executed three times to calculate the quantities in Eq. 13. Replacing SSIM_f value by 1 (the ideal value) and substituting the above values in Eq. 13, the approximated upper bound on α can be calculated as:

$$\alpha = \frac{1 - s}{rZ}\left(1 - \text{SSIM}'\right) \tag{14}$$

In our experiments, the value of α was calculated for the middle slice of each volume; the same value was used for all the slices in that volume, since SSIM values between original image and corresponding enhanced images among all the slices of specific volume were very close. The range of α found for our dataset was [20, 60]. Furthermore, it is important to mention that for any value of α in the specified range, the SSIM index value improves compared to that obtained in the first iteration, i.e. without incorporating SSIM gradient.

As stated in the previous sections, the objective of this work is to improve contrast while maintaining structural similarity with the input image to facilitate tumor segmentation. Therefore, along with ensuring this structural similarity (via SSIM gradient), we incorporate another criterion in our proposed method to measure the contrast enhanced at each iteration by applying 2D-HS. Therefore, 2D entropy is used to control the level of enhancement. The stopping criterion for the enhancement process is determined by the gain in two-dimensional entropy achieved for the enhanced image. The rationale of our methodology is to exploit inter-pixel correlation; therefore, we have used 2D entropy to formulate this criterion as:

$$E_t = -\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} h_{f_{e_{(t)}}}(i,j)\ln(h_{f_{e_{(t)}}}(i,j)) \tag{15}$$

In Eq. 15, E_t represents the value of the 2D entropy for an image f_e at iteration t, where t varies from 1 to 10. $h_{f_{e_{(t)}}}(i,j)$ is the value of transition probability of gray level pairs. The change in entropy of the enhanced image gained with every iteration is calculated as follows:

$$\Delta E = E_t - E_{t-1} \tag{16}$$

Moreover, the change in entropy values (normalized to lie in the range [0,1]) across all the iterations is shown in Fig. 2 when the algorithm is applied to our dataset. Since the entropy values of all the images in a particular volume were very similar, entropy value of the middle slice from each volume is plotted for the sake of compactness. At a specific point in the optimization process, when ΔE becomes negligible (close to zero) or when the ΔE value starts oscillating, the enhancement process is stopped. Both these scenarios imply that further application of the enhancement process either will not further enhance the image or will likely introduce artifacts in the image. We observe an obvious increase in entropy when applying the proposed method for the first iteration. Subsequent applications bring a slow entropy increase; however, we show in section V that the segmentation accuracy is higher when segmentation is applied on images enhanced using the proposed algorithm. The result of applying the OPTGCE method and comparison with other methods is presented in the following section.

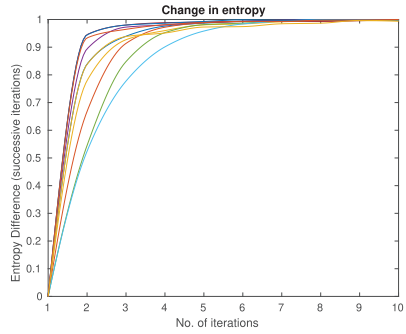


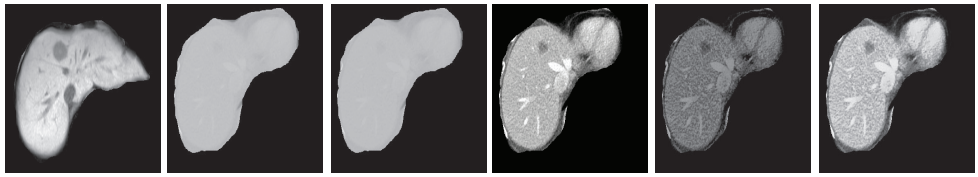
FIGURE 2. Variation in entropy values with iterations.

IV. ASSESSMENT OF CONTRAST ENHANCEMENT

In this section, we describe the dataset used in the experiment and the results obtained using different methods [7], [31], [32]. The qualitative and quantitative assessments are elaborated below.

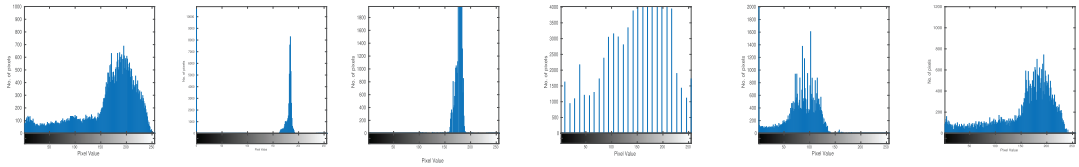
A. DATASET

The data used in this research work is provided by the Intervention Center, Oslo University Hospital in Norway.



(a1) Guidance (a2) Input (a3) HEMIC (a4) AVHEQ (a5) CMGE (a6) OPTGCE

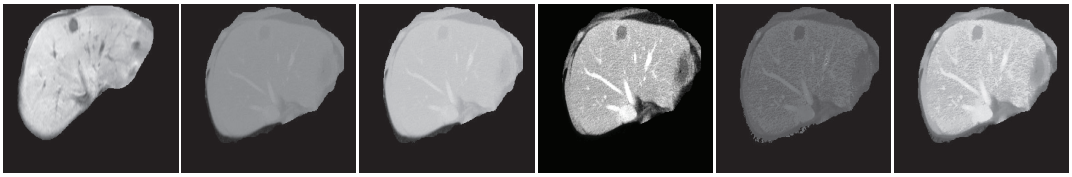
(a) Enhancement Results from state of the art methods



(b1) histogram a1 (b2) histogram a2 (b3) histogram a3 (b4) histogram a4 (b5) histogram a5 (b6) histogram a6

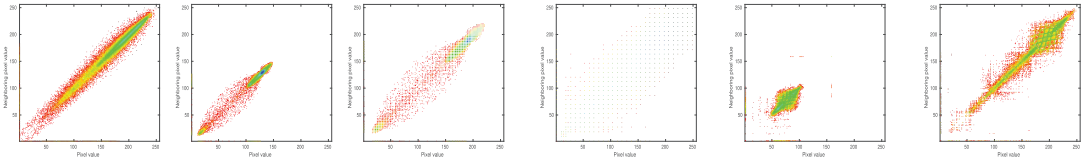
(b) Corresponding histograms of guidance, input and enhanced images

FIGURE 3. Comparison of proposed method with state of the art methods and their corresponding histograms.



(a1) Guidance (a2) Input (a3) HEMIC (a4) AVHEQ (a5) CMGE (a6) OPTGCE

(a) Enhancement Results from state of the art methods



(b1) plot a1 (b2) plot a2 (b3) plot a3 (b4) plot a4 (b5) plot a5 (b6) plot a6

(b) Corresponding GLCM plots of guidance, input and enhanced images

FIGURE 4. Comparison of proposed method with state of the art methods and their corresponding GLCM plots.

Liver CT and MR data of the same patient are used; however, CT-MRI data is not registered since registration is not required for global enhancement methods. We tested our method on 10 patients' data constituting 99 CT-MR image pairs (containing tumors). The images from different volumes are of different spatial sizes (such as 512×512, 360×240) with pixel values in the range [0, 255]. In medical image processing tasks such as segmentation and enhancement, the processing is often restricted to a particular organ and the nearby organs are removed from the medical images [38], [39]. The liver area in the images is therefore separated and processing is applied only to this region.

1) QUALITATIVE ANALYSIS

In this subsection, few enhanced images alongwith their corresponding histograms and GLCM plots are presented in Fig. 3 and 4 respectively. To ensure a fair comparison, we selected all histogram based methods where [31] and [32] employ optimization based histogram processing and [7] applies cross-modality guided HS. We denote these methods as Averaging Histogram Equalization (AVHEQ) [31], Histogram Equalization with Maximum Intensity Coverage (HEMIC) [32] and Cross-Modality Guidance-based enhancement (CMGE) [7].

The input image in Fig. 3a2 has low contrast as validated by its histogram. Similarly, the input image in Fig. 4a2 is

TABLE 2. Quantitative assessment of different enhancement methods.

Dataset #	Entropy				MIGLCM			
	HEMIC [32]	AVHEQ [31]	CMGE [7]	OPTGCE	HEMIC [32]	AVHEQ [31]	CMGE [7]	OPTGCE
1	2.32	2.21	2.76	**3.13**	**1.1**	1.07	0.95	**1.1**
2	1.58	1.7	1.74	**2.4**	0.93	0.9	0.86	**1.12**
3	1.7	1.6	1.7	**1.9**	0.93	0.92	0.91	**1.06**
4	2.03	1.95	2.72	**3.00**	1.22	1.21	1.27	**1.4**
5	2.51	2.45	3.43	**3.62**	1.16	1.12	1.07	**1.31**
6	1.54	1.51	1.92	**2.12**	0.86	0.85	0.82	**0.98**
7	2.61	2.52	3.23	**3.64**	1.28	1.24	1.14	**1.42**
8	1.52	1.46	1.94	**2.1**	0.82	0.81	0.75	**0.94**
9	1.43	1.36	1.67	**1.88**	0.78	0.77	0.75	**0.81**
10	1.11	1.11	1.32	**1.41**	0.64	0.62	0.63	**0.76**

low-contrast CT image. The images enhanced using HEMIC (Fig. 3a3 and 4a3) do not show noticeable contrast improvement. Although CMGE (Fig. 3a5) expands the dynamic range of the image (Fig. 3a5), it darkens the image. AVHEQ (Fig. 4a4) stretches the dynamic range of enhanced images, however, its GLCM plot (Fig. 4b4) shows significant gaps among the pixel pairs and consequently the compactness of the plot is lost. The plot of OPTGCE enhanced image (Fig. 4b6) reflects the uniform distribution of the pixel pairs. Furthermore, it approximates the plot of the guidance image (Fig.4b1) in evenly distributing the pixel pairs; this similarity is also verified in their histograms (Fig. 3b1 and Fig. 3b6).

2) QUANTITATIVE ANALYSIS

Image Quality Assessment (IQA) is a well-investigated research field especially in the case of natural images [40]. However, the use of existing IQA metrics has serious limitations in the medical context [41]. The objectives of CE in the medical context are quite different [42], [43]. While in the case of natural images the objective is to measure the effect of various distortions on the perceptual quality of the image; in the medical context even if some degradation may disturb the radiologists the focus is rather on the diagnosis. Therefore, the existing IQA metrics must be used with special care. Another challenging topic is how to evaluate the performance of a given image quality enhancement algorithm in terms of perceptual quality [44]. In the present study, we focus on some contrast enhancement evaluation (CEE) metrics.

The motivation of the OPTGCE is to emphasize the appearance of specific structures in the image and convey the maximum structural information to facilitate tumor segmentation. To this end, we have chosen three different CEE metrics to evaluate the quality of enhanced images. The first metric is a mutual information-based no reference metric called MIGLCM [45]. This metric offers quantitative criteria that examines the changes in the statistical features, joint entropy, and mutual information, acquired from the GLCM of the original and the enhanced images. Besides MIGLCM, we have used a recent metric Multi-Criteria Contrast Enhancement Evaluation (MCCEE) found to be effective for the evaluation of CE in CT images that have been enhanced to improve tumor segmentation [46]. It is a comprehensive metric as it not only measures improvement in contrast but also considers other evaluation criteria like over-enhancement. For MCCEE, four features are evaluated for each image corresponding to four different criteria. These criteria include contrast enhancement, structure preservation, lightness order preservation, and brightness preservation. Two of the four features corresponding to the structure and lightness order preservation are evaluated from the subband images after wavelet decomposition. MCCEE is finally evaluated using a trained Support Vector Regressor (SVR) with subjective quality scores or DICE from the subsequent segmentation. MCCEE here is applied on data of three patients only since the rest of the data from seven patients is used for training.

TABLE 3. Median MCCEE values for different methods.

Data.#	HEMIC [32]	AVHEQ [31]	CMGE [7]	OPTGCE
1	0.25	0.32	0.31	**0.34**
2	0.23	0.22	0.22	**0.26**
3	0.21	0.26	0.36	**0.56**

For the last metric, we have used entropy, which is often used in QA of medical image enhancement [31]. Table 2 lists the median values of MIGLCM and entropy, whereas Table 3 shows the MCCEE values. It is pertinent to mention that the higher the MCCEE score, the better enhancement result is; the range of MCCEE is [0,1]. Similarly, a higher value of MIGLCM reflects better performance of CE algorithms. Besides, higher entropy values also correspond to superior CE performance; however, there is no specified range for this metric. From the tabular results, we can observe that OPTGCE demonstrates the best performance. For MCCEE and entropy, CMGE, HEMIC, and AVHEQ are ranked low overall by the two QA metrics. In the case of MIGLCM, HEMIC is ranked as the second-best and CMGE gives the

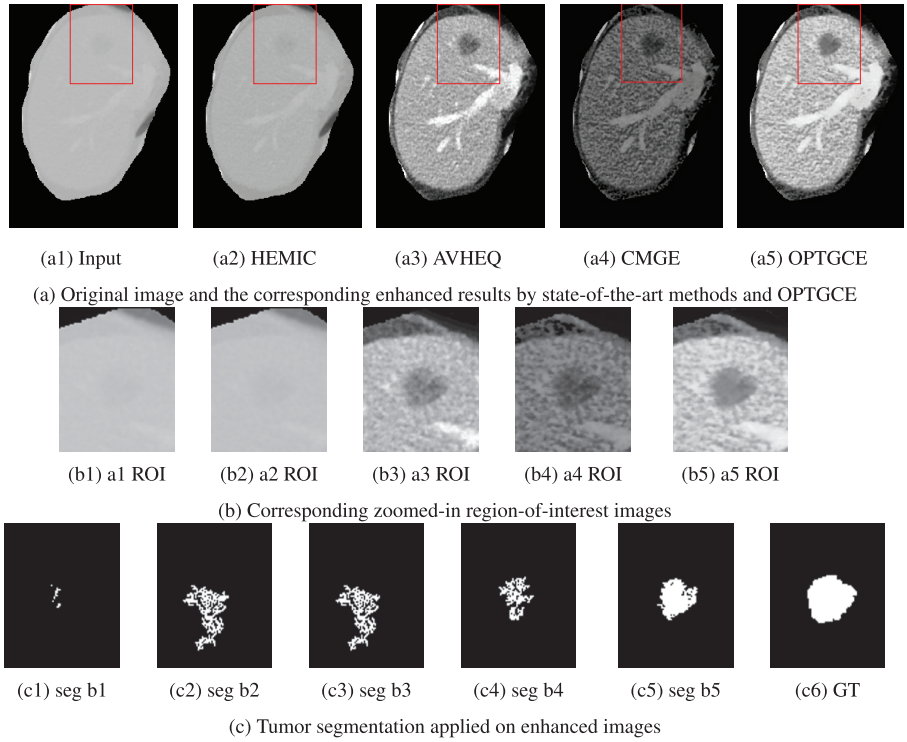


FIGURE 5. Comparison of tumor segmentation results with the ground truth.

poorest results. All in all, we observe that OPTGCE shows the best performance for all the quantitative metrics chosen. In this section, we compared the performance of our proposed method individually without looking at its effect on the subsequent segmentation task. In the following section, the application of the gradient-driven Seeded Region Growing (SRG) method on the enhanced images will be discussed.

V. TOWARDS AN OPTIMAL SEGMENTATION PRESERVING LOCAL STRUCTURES

Segmentation in low contrast medical images, particularly CT images is a delicate operation. The segmentation process is often accompanied by miss-classification errors that negatively impact high-level tasks, for instance, diagnosis in the medical context. Similar to several fields of scientific research, deep learning-based image segmentation approaches seem to dominate the state of the art [47]. However, DL-based techniques require an extensive amount of data to train the networks, which is difficult to acquire in the medical context due to confidentiality and ethical considerations. This work entails the segmentation of liver tumors from a rather limited volume of data consisting of only 99 liver tumor images (out of 10 patients' data). We, therefore, resort to the traditional approach, SRG. Several studies report the use of SRG for segmenting medical images compared to more sophisticated approaches [48]–[50]. However, as has been pointed out, classical segmentation, whether stochastic or deterministic, inevitably induces pixel classification errors, which could be fatal in the case under study here. One solution to minimize these errors is to apply pre-processing so as to amplify the inter-pixel gradient to facilitate the discrimination of local structures. Contrast enhancement is the most intuitive solution. In this work, one of the segmentation methods that seems to us the most adequate is the region growing technique based on the gradient of pixel values [51]. Indeed, the fact of first carrying out CE amplifies the gradient, which results in putting the gradient-based region growing method in the most favorable conditions. A constraint that may limit the utility of this approach is the execution time; therefore, we use its parallel implementation as proposed in [13]. In this section, we present the results of applying segmentation on enhanced images along with its quantitative assessment.

A. QUANTITATIVE ASSESSMENT OF SEGMENTATION

The results of applying gradient-driven SRG algorithm on enhanced as well as input images are demonstrated in Fig. 5 and 6. It can be noticed that the tumor in the

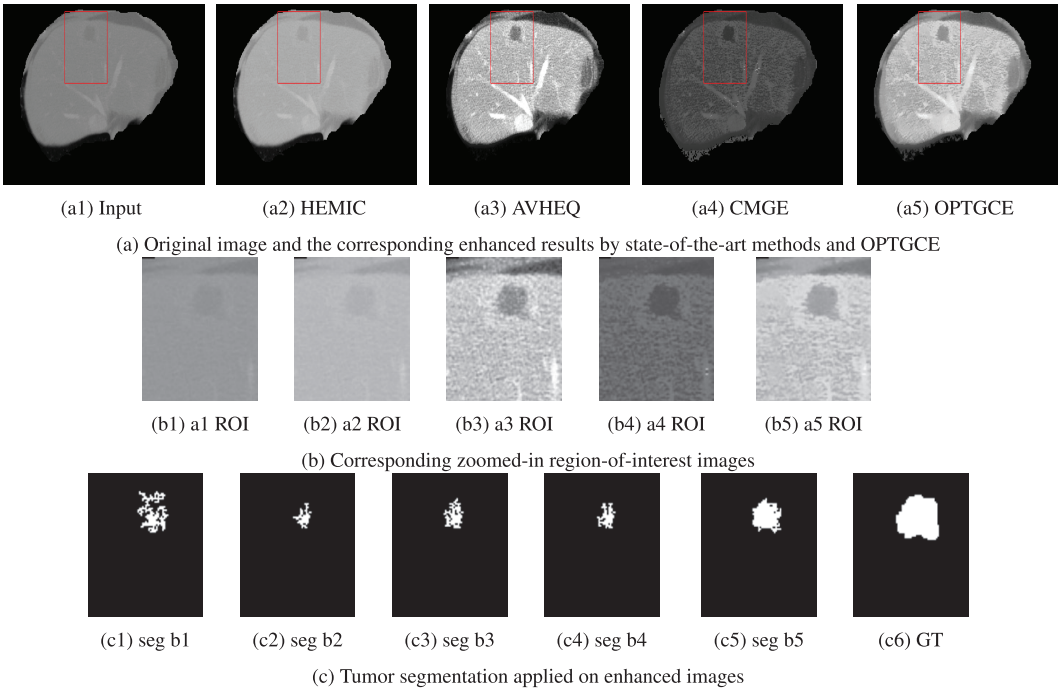


FIGURE 6. Comparison of tumor segmentation results with the ground truth.

TABLE 4. Comparison of different segmentation assessment method for enhancement results.

Data	Input w/o processing			HEMIC [32]			AVHEQ [31]			CMGE [7]			OPTGCE		
	PPV	Dice	Haus.	PPV	Dice	Haus.	PPV	Dice	Haus.	PPV	Dice	Haus.	PPV	Dice	Haus.
vol1	**0.67**	0.16	14.21	0.63	0.26	23.5	0.66	0.173	17	0.56	0.36	14.6	**0.67**	**0.46**	**9.92**
vol2	**0.9**	0.14	8.66	0.83	0.26	8.26	0.86	0.27	7.7	0.88	0.38	7.3	**0.9**	**0.45**	**5.9**
vol3	0.85	0.31	16.18	0.85	0.36	15.78	0.86	0.42	14	0.83	0.48	12.47	**0.91**	**0.57**	**10.5**

input image can hardly be seen in Fig. 5a1 and 6a1 without applying CE. In general, application of the CE methods improve the contrast of the input image, which ultimately enables SRG to locate tumor contours favorably. However, OPTGCE well preserves uniformity in the structure of tumors in the enhanced image together with yielding sharp tumor edges. Therefore, Seeded Region Growing (SRG) algorithm is better able to locate the tumor contours in the OPTGCE-enhanced images. This property enables OPTGCE to outperform other CE methods in facilitating tumor segmentation. The quantitative segmentation assessment as well as qualitative comparison with ground truth also supports our claim.

It should be noted here that the segmentation results are demonstrated without applying any kind of post-processing such as morphological region filling. We believe that better results could be obtained if appropriate post-processing was applied to the segmented images. For further validation, we quantitatively evaluate the segmentation results using three assessment metrics, i.e., Positive Predictive Value (PPV), Dice and Hausdorff distance (with Euclidean distance). The average values of these metrics obtained for each volume are shown in Table 4. In the past, several metrics have been proposed to evaluate the performance of segmentation algorithms such as intensity-based, shape-based, and or distance-based. One of the challenges in medical image segmentation assessment is that the object of interest constitutes a small part of the image, therefore the assessment methods are biased to yield more weightage to specificity compared to sensitivity. Distance-based metrics such as Hausdorff distance are capable of detecting data outliers in such cases where intensity-based approaches may often fail. There is no standard range for the hausdorff distance values, however, the lower value indicates superior segmentation outcomes.

Among the numeric results in Table 4, PPV values in general are greater than 0.8 for all the segmentations. Since PPV computes the ratio between the number of pixels

correctly classified as tumors to the number of pixels correctly classified and the non-tumor pixels wrongly classified as tumors, this metric gives similar values to all the methods. It should be noted that the range of PPV metric is [0,1]; where 1 implies accurate segmentation. It can also be observed that all the segmentations in Fig. 5 and 6 do not include many non tumor pixels in the resultant segmentation when compared to the ground truth. Although Dice yields better scores for segmentation applied on images enhanced using the proposed method, the overall dice scores are low. Dice similarity metric gives higher value to the terms that compute the intersection between true positives in segmentation under test and ground truth. The range of dice score lies between 0 and 1, where 1 corresponds to the perfect segmentation. In the proposed CE approach, the segmented area does not completely overlap with the GT, contributing to lower dice scores; moreover, not applying any kind of post-processing to segmentation also introduces discontinuities and non uniformity in segmented tumors. It is worth mentioning here that the Dice scores are lowest when SRG algorithm is applied on the input images without any kind of enhancement, whereas Hausdorff distance shows highest value. The segmentation in the case of the OPTGCE method consistently achieves lower Hausdorff distance values for all three volumes. CMGE is the second-best while HEMIC ranks lowest in all the three cases tested.

VI. CONCLUSION

This study proposes an optimization-based guided contrast enhancement approach OPTGCE for low contrast CT images. The proposed technique adopts a context-aware 2D histogram-based scheme of exploiting information in the better perceptual quality guidance image for global contrast enhancement, while local image structures are enhanced through SSIM based measure in an optimization framework. This combination effectively improves the contrast while minimizing the artifacts associated with typical histogram-based enhancement methods to preserve the morphological information of the image during enhancement. The qualitative and quantitative analysis using metrics including entropy, MCCEE, and MIGLCM shows the superiority of the proposed method in comparison with the existing methods that do not include guidance mechanism. Finally, a tumor segmentation algorithm is applied on the enhanced images to analyze the performance of the proposed method in facilitating tumor segmentation. The comparison with the ground truth and quantitative assessment using Hausdorff distance, dice, and PPV metrics validate the superior performance of OPTGCE. With the availability of more data, goal-oriented contrast enhancement can be implemented using deep neural networks to facilitate tumor segmentation in different organs.

ACKNOWLEDGMENT

The authors thank the Intervention Centre, Oslo University Hospital, Norway, specially Ole Jakob Elle and Egidijuis Pelanis for providing images with ground truths to clinically validate tumor segmentation.

REFERENCES

[1] Y. Alvi, D. Regan, L. Schook, R. Gaba, and K. Schachtschneider, "Transcriptional regulation of alcohol induced liver fibrosis in a translational porcine hepatocellular carcinoma model," *Biochimie*, vol. 182, pp. 73–84, Mar. 2021.

[2] B. Sdiri, M. Kaaniche, F. A. Cheikh, A. Beghdadi, and O. J. Elle, "Efficient enhancement of stereo endoscopic images based on joint wavelet decomposition and binocular combination," *IEEE Trans. Med. Imag.*, vol. 38, no. 1, pp. 33–45, Jan. 2019.

[3] M. R. Oliva, "Liver cancer imaging: Role of CT, MRI, US and PET," *Cancer Imag.*, vol. 4, pp. S42–S46, Sep. 2004.

[4] B. Li and W. Xie, "Adaptive fractional differential approach and its application to medical image enhancement," *Comput. Electr. Eng.*, vol. 45, pp. 324–335, Jul. 2015.

[5] B. Kim, J. Ponce, and B. Ham, "Deformable kernel networks for joint image filtering," *Int. J. Comput. Vis.*, vol. 129, no. 2, pp. 579–600, Feb. 2021.

[6] Y. Li, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Joint image filtering with deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 1909–1923, Aug. 2019.

[7] R. Naseem, F. A. Cheikh, A. Beghdadi, O. J. Elle, and F. Lindseth, "Cross modality guided liver image enhancement of CT using MRI," in *Proc. 8th Eur. Workshop Vis. Inf. Process. (EUVIP)*, Oct. 2019, pp. 46–51.

[8] X. Guo, Y. Li, J. Ma, and H. Ling, "Mutually guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, pp. 694–707, Mar. 2020.

[9] Q. Yan, X. Shen, L. Xu, S. Zhuo, X. Zhang, L. Shen, and J. Jia, "Cross-field joint image restoration via scale map," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1537–1544.

[10] Z. Al-Ameen and G. Sulong, "A new algorithm for improving the low contrast of computed tomography images using tuned brightness controlled single-scale Retinex," *Scanning*, vol. 37, no. 2, pp. 116–125, 2015.

[11] N. Satpute, R. Naseem, R. Palomar, O. Zachariadis, J. Gómez-Luna, F. A. Cheikh, and J. Olivares, "Fast parallel vessel segmentation," *Comput. Methods Programs Biomed.*, vol. 192, Aug. 2020, Art. no. 105430.

[12] W. Zhu, H. Jiang, E. Wang, Y. Hou, L. Xian, and J. Debnath, "X-ray image global enhancement algorithm in medical image classification," *Discrete Continuous Dyn. Syst.*, vol. 12, nos. 4–5, p. 1297, 2019.

[13] N. Satpute, R. Naseem, E. Pelanis, J. Gómez-Luna, F. A. Cheikh, O. J. Elle, and J. Olivares, "GPU acceleration of liver enhancement for tumor segmentation," *Comput. Methods Programs Biomed.*, vol. 184, Feb. 2020, Art. no. 105285.

[14] S. Survarachakan, E. Pelanis, Z. A. Khan, R. P. Kumar, B. Edwin, and F. Lindseth, "Effects of enhancement on deep learning based hepatic vessel segmentation," *Electronics*, vol. 10, no. 10, p. 1165, May 2021.

[15] A. N. Avanaki, "Exact global histogram specification optimized for structural similarity," *Opt. Rev.*, vol. 16, no. 6, pp. 613–621, Nov. 2009.

[16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[17] A. Beghdadi and A. Le Negrate, "Contrast enhancement technique based on local detection of edges," *Comput. Vis. Graph. Image Process.*, vol. 46, no. 2, pp. 162–174, May 1989.

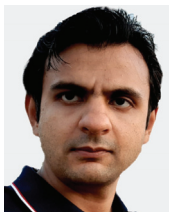
[18] I. S. Isa, S. N. Sulaiman, M. Mustapha, and N. K. A. Karim, "Automatic contrast enhancement of brain MR images using average intensity replacement based on adaptive histogram equalization (AIR-AHE)," *Biocybern. Biomed. Eng.*, vol. 37, no. 1, pp. 24–34, 2017.

[19] Z. Al-Ameen, G. Sulong, A. Rehman, A. Al-Dhelaan, T. Saba, and M. Al-Rodhaan, "An innovative technique for contrast enhancement of computed tomography images using normalized gamma-corrected contrast-limited adaptive histogram equalization," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 1, pp. 1–12, Dec. 2015.

[20] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, and J. B. Zimmerman, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.

[21] C.-C. Sun, S.-J. Ruan, M.-C. Shie, and T.-W. Pai, "Dynamic contrast enhancement based on histogram specification," *IEEE Trans. Consum. Electron.*, vol. 51, no. 4, pp. 1300–1305, Nov. 2005.
[22] B. Xiao, H. Tang, Y. Jiang, W. Li, and G. Wang, "Brightness and contrast controllable image enhancement based on histogram specification," *Neurocomputing*, vol. 275, pp. 2798–2809, Jan. 2018.
[23] T. Celik, "Two-dimensional histogram equalization and contrast enhancement," *Pattern Recognit.*, vol. 45, no. 10, pp. 3810–3824, Oct. 2012.
[24] S. Jung, "Two-dimensional histogram specification using two-dimensional cumulative distribution function," *Electron. Lett.*, vol. 50, no. 12, pp. 872–874, 2014.
[25] D. Menotti, L. Najman, J. Facon, and A. D. A. Araujo, "Multi-histogram equalization methods for contrast enhancement and brightness preserving," *IEEE Trans. Consum. Electron.*, vol. 53, no. 3, pp. 1186–1194, Aug. 2007.
[26] Z. Huang, X. Li, N. Wang, L. Ma, and H. Hong, "Simultaneous denoising and enhancement for X-ray angiograms by employing spatial-frequency filter," *Optik*, vol. 208, Apr. 2020, Art. no. 164287.
[27] Z. Huang, Y. Zhang, Q. Li, T. Zhang, N. Sang, and H. Hong, "Progressive dual-domain filter for enhancing and denoising optical remote-sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 759–763, May 2018.
[28] Z. Huang, H. Fang, Q. Li, Z. Li, T. Zhang, N. Sang, and Y. Li, "Optical remote sensing image enhancement with weak structure preservation via spatially adaptive gamma correction," *Infr. Phys. Technol.*, vol. 94, pp. 38–47, Nov. 2018.
[29] Z. Huang, Z. Wang, J. Zhang, Q. Li, and Y. Shi, "Image enhancement with the preservation of brightness and structures by employing contrast limited dynamic quadri-histogram equalization," *Optik*, vol. 226, Jan. 2021, Art. no. 165877.
[30] L. Rundo, A. Tangherloni, M. S. Nobile, C. Militello, D. Besozzi, G. Mauri, and P. Cazzaniga, "MedGA: A novel evolutionary method for image enhancement in medical imaging systems," *Expert Syst. Appl.*, vol. 119, pp. 387–399, Apr. 2019.
[31] S. C.-F. Lin, C. Y. Wong, M. A. Rahman, G. Jiang, S. Liu, N. Kwok, H. Shi, Y.-H. Yu, and T. Wu, "Image enhancement using the averaging histogram equalization (AVHEQ) approach for contrast improvement and brightness preservation," *Comput. Elect. Eng.*, vol. 46, pp. 356–370, Aug. 2015.
[32] C. Y. Wong, S. Liu, S. C. Liu, M. B. Rahman, S. Lin, G. Jiang, N. Kwok, and H. Shi, "Image contrast enhancement using histogram equalization with maximum intensity coverage," *J. Mod. Opt.*, vol. 63, no. 16, pp. 1618–1629, 2016.
[33] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proc. ECCV*, 2010, pp. 1–14.
[34] S. Zhuo, X. Zhang, X. Miao, and T. Sim, "Enhancing low light images using near infrared flash images," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 2537–2540.
[35] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.
[36] F. Deravi and S. K. Pal, "Grey level thresholding using second-order statistics," *Pattern Recognit. Lett.*, vol. 1, nos. 5–6, pp. 417–422, Jul. 1983.
[37] J. C. Mello Roman, H. Legal-Ayala, and J. L. Vazquez Noguera, "Top-hat transform for enhancement of aerial thermal images," in *Proc. 30th SIBGRAPI Conf. Graph., Patterns Images (SIBGRAPI)*, Oct. 2017, pp. 277–284.
[38] N. Al-Najdawi, M. Biltawi, and S. Tedmori, "Mammogram image visual enhancement, mass segmentation and classification," *Appl. Soft Comput.*, vol. 35, pp. 175–185, Oct. 2015.
[39] T. Hopp, M. Dietzel, P. Baltzer, P. Kreisel, W. A. Kaiser, H. Gemmeke, and N. Ruiter, "Automatic multimodal 2D/3D breast image registration using biomechanical FEM models and intensity-based optimization," *Med. Image Anal.*, vol. 17, 2, pp. 209–218, 2013.
[40] D. Chandler, "Seven challenges in image quality assessment: Past, present, and future research," *Int. Scholarly Res. Notices*, vol. 2013, pp. 1–53, Mar. 2013.
[41] J.-F. Pambrun and R. Noumeir, "Limitations of the SSIM quality metric in the context of diagnostic imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 2960–2963.
[42] M. Outtas, L. Zhang, O. Deforges, W. Hammidouche, A. Serir, and C. Cavaro-Menard, "A study on the usability of opinion-unaware no-reference natural image quality metrics in the context of medical images," in *Proc. Int. Symp. Signal, Image, Video Commun. (ISIVC)*, 2016, pp. 308–313.
[43] C. Cavaro-Menard, L. Zhang, and P. Le Callet, "Diagnostic quality assessment of medical images: Challenges and trends," in *Proc. 2nd Eur. Workshop Vis. Inf. Process. (EUVIP)*, Jul. 2010, pp. 277–284.
[44] M. A. Qureshi, A. Beghdadi, and M. Deriche, "Towards the design of a consistent image contrast enhancement evaluation measure," *Signal Process., Image Commun.*, vol. 58, pp. 212–227, Oct. 2017.
[45] M. A. Qureshi, M. Deriche, A. Beghdadi, and M. Mohandes, "An information based framework for performance evaluation of image enhancement methods," in *Proc. Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Nov. 2015, pp. 519–523.
[46] Z. A. Khan, A. Beghdadi, F. A. Cheikh, M. Kaaniche, and M. Ali Qureshi, "A multi-criteria contrast enhancement evaluation measure using wavelet decomposition," in *Proc. IEEE 22nd Int. Workshop Multimedia Signal Process. (MMSP)*, Sep. 2020, pp. 1–6.
[47] Y. Zhang, B. Jiang, J. Wu, D. Ji, Y. Liu, Y. Chen, E. X. Wu, and X. Tang, "Deep learning initialized and gradient enhanced level-set based segmentation for liver tumor from CT images," *IEEE Access*, vol. 8, pp. 76056–76068, 2020.
[48] G. Lin, W. Wang, C. Kang, and C. Wang, "Multispectral MR images segmentation based on fuzzy knowledge and modified seeded region growing," *Magn. Reson. Imag.*, vol. 30, 2, pp. 230–246, 2012.
[49] D. Dreizin, U. K. Bodanapally, N. Neerchal, N. Tirada, M. Patlas, and E. Herskovits, "Volumetric analysis of pelvic hematomas after blunt trauma using semi-automated seeded region growing segmentation: A method validation study," *Abdominal Radiol.*, vol. 41, no. 11, pp. 2203–2208, Nov. 2016.
[50] J. Lian, Y. Ma, Y. Ma, B. Shi, J. Liu, Z. Yang, and Y. Guo, "Automatic gallbladder and gallstone regions segmentation in ultrasound image," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 4, pp. 553–568, Apr. 2017.
[51] G. N. Harikrishna Rai and T. R. Gopalakrishnan Nair, "Gradient based seeded region grow method for CT angiographic image segmentation," 2010, *arXiv:1001.3735*. [Online]. Available: http://arxiv.org/abs/1001.3735

RABIA NASEEM received the master's degree in software engineering from the University of Engineering and Technology, Taxila, Pakistan. She is currently pursuing the Ph.D. degree with the Department of Computer Science, Norwegian University of Science and Technology, Norway. She is also working as a Marie Curie Fellow under the European Union-Funded Project High-Performance Soft Tissue Navigation and her research interests include medical image processing, particularly medical image enhancement and deep learning.

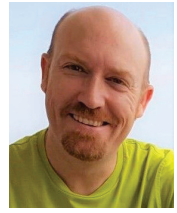


ZOHAIB AMJAD KHAN (Student Member, IEEE) received the dual master's degree in information and communication technologies from the Politecnico di Torino, Italy, and Karlsruhe Institute of Technology, Germany, in 2010. He has been a Ph.D. Researcher working at the Institut Galilée, Université Sorbonne Paris Nord, France, since 2017. His research interests include medical imaging, image and video processing and quality assessment, applied deep learning, and virtual reality simulations.

NITIN SATPUTE received the M.E. degree in embedded systems from BITS Pilani, India, and the Ph.D. degree (Hons.) in computer science from the University of Cordoba, Spain. Before joining AU, he was a Marie Curie Researcher under the HiPerNav Project funded by the European Union (EU) at the University of Cordoba. He is currently a Postdoctoral Researcher in machine learning and computational intelligence with Aarhus University (AU), Denmark. He is also involved in the fast efficient distributed training for deep neural networks (DNN). His research interests include AI, parallel computing, medical image analysis, and automated crowd management systems.

AZEDDINE BEGHDADI (Senior Member, IEEE) received the master's degree in optics and signal processing from Paris-Saclay University, in June 1983, and the Ph.D. degree in physics specialization in optics and signal processing from Sorbonne University, in June 1986. He has been a Full Professor with University Sorbonne Paris Nord (USPN), since 2000. He is currently the Founding Member of the Laboratory of Information Processing and Transmission (L2TI laboratory), and was its Director, from 2010 to 2016. He worked at different places during his Ph.D. thesis, including the Laboratoire d'Optique des Solides (University Pierre et Marie Curie—Sorbonne University) and Groupe d'Analyse d'Images Biomédicales (CNAM Paris). From 1987 to 1989, he held the position of a Lecturer at USPN. During the period 1987 to 1998, he was with the LPMTM-CNRS Laboratory working on scanning electron microscope materials image analysis. He has published over more than 300 international refereed scientific papers. His research interests include image/video quality enhancement and assessment, bio-inspired models for image analysis and processing, and machine learning. He is an Elected Member of IEEE-MMSP TC and EURASIP VIP-TAC. He also served as a session organizer and a member for the organizing and technical committees for many IEEE conferences. He served as the conference chair and the technical chair for many international conferences. He is a EURASIP Member and IEEE-MMTC Member. He is the Founder and the Steering Committee Chair of the European Workshop on Visual Information Processing (EUVIP). He is an Associate Editor of *Signal processing: Image Communication* journal (Elsevier), *European Journal on Image and Video Processing* (Springer, Verlag), and *Mathematical Problems in Engineering* journal (Hindawi).

FAOUZI ALAYA CHEIKH (Senior Member, IEEE) received the Ph.D. degree in information technology from Tampere University of Technology, Tampere, Finland, in April 2004. He had worked as a Researcher with the Signal Processing Algorithm Group, Tampere University of Technology, in 1994. Since 2006, he has been affiliated with the Department of Computer Science and Media Technology, Gjøvik University College, Norway, as an Associate Professor. Since January 2016, he has also been with the Norwegian University of Science and Technology (NTNU). He teaches courses on image and video processing and analysis and media security. He is currently the co-supervisor of five Ph.D. students. He has been involved in several European and national projects, such as ESPRIT, NOBLESS, COST 211Quat, HyPerCept, IQ-Med, and H2020 ITN HiPerNav. His research interests include e-Learning, 3-D imaging, image and video processing and analysis, video-based navigation, biometrics, pattern recognition, embedded systems, and content-based image retrieval. In these areas, he has published over 100 peer-reviewed journal articles and conference papers, and supervised four postdoctoral researchers, five Ph.D., and a number of M.Sc. thesis projects. He is a member of NOBIM and Forskerforbundet [The Norwegian Association of Researchers (NAR)]. He is on the Editorial Board of the *IET Image Processing Journal* and the *Journal of Advanced Robotics & Automation*, and the technical committees of several international conferences. He is an expert reviewer to a number of scientific journals and conferences related to the field of his research.



JOAQUÍN OLIVARES received the B.S. and M.S. degrees in computer sciences and the M.S. degree in electronics engineering from the Universidad de Granada, Spain, in 1997, 1999, and 2003, respectively, and the Ph.D. degree from the Universidad de Córdoba, Spain, in 2008. He has been an Associate Professor with the Department of Electronic and Computer Engineering, Universidad de Córdoba, since 2001. He is currently the Founder and the Head of the Advanced Informatics Research Group. His research interests include the field of the Internet of Things, embedded systems, computer vision, and high performance computing.

Chapter 8

Paper C: Fast parallel vessel segmentation

Computer Methods and Programs in Biomedicine 192 (2020) 105430

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Fast parallel vessel segmentation

Nitin Satpute[a,*], Rabia Naseem[b], Rafael Palomar[c], Orestis Zachariadis[a], Juan Gómez-Luna[d], Faouzi Alaya Cheikh[b], Joaquín Olivares[a]

[a] Department of Electronic and Computer Engineering, Universidad de Córdoba, Spain
[b] Norwegian Colour and Visual Computing Lab, Norwegian University of Science and Technology, Norway
[c] The Intervention Centre, Oslo University Hospital, Norway
[d] Department of Computer Science, ETH Zurich, Switzerland

ARTICLE INFO

Article history:
Received 17 December 2019
Revised 17 February 2020
Accepted 2 March 2020

Keywords:
Seeded region growing
GPU
Kernel termination and relaunch (KTRL)
Persistent
Grid-stride loop

ABSTRACT

Background and Objective: Accurate and fast vessel segmentation from liver slices remain challenging and important tasks for clinicians. The algorithms from the literature are slow and less accurate. We propose fast parallel gradient based seeded region growing for vessel segmentation. Seeded region growing is tedious when the inter connectivity between the elements is unavoidable. Parallelizing region growing algorithms are essential towards achieving real time performance for the overall process of accurate vessel segmentation.

Methods: The parallel implementation of seeded region growing for vessel segmentation is iterative and hence time consuming process. Seeded region growing is implemented as kernel termination and relaunch on GPU due to its iterative mechanism. The iterative or recursive process in region growing is time consuming due to intermediate memory transfers between CPU and GPU. We propose persistent and grid-stride loop based parallel approach for region growing on GPU. We analyze static region of interest of tiles on GPU for the acceleration of seeded region growing.

Results: We aim fast parallel gradient based seeded region growing for vessel segmentation from CT liver slices. The proposed parallel approach is 1.9x faster compared to the state-of-the-art.

Conclusion: We discuss gradient based seeded region growing and its parallel implementation on GPU. The proposed parallel seeded region growing is fast compared to kernel termination and relaunch and accurate in comparison to Chan-Vese and Snake model for vessel segmentation.

1. Introduction

In medical imaging, vessel segmentation from liver slices is one of the challenging tasks. Seeded region growing (SRG) is a widely used approach for semi automatic vessel segmentation [1,2]. Delibasis et. al. [3] have proposed a tool based on a modified version of SRG algorithm, combined with a priori knowledge of the required shape. SRG starts with a set of pixels called seeds and grows a uniform, connected region from each seed. Key steps to SRG are to define seed(s) and a classifying criterion that relies on the image properties and user interaction [4]. SRG starts from a seed and finds the similar neighboring points based on the threshold criteria using 4 or 8 connectivity. Region is grown if the threshold criteria is satisfied. Similar neighbors are new seed points for the next iteration. This process is repeated until the region can not be grown further. In practice, it demands high computational cost to the large amount of dependent data to be processed in SRG especially in the medical image analysis and still requires efficient solutions [5].

SRG is an iterative process. SRG is invoked continuously until region can not be grown further. Iterative process in SRG, when implemented on GPU requires terminating kernel and relaunching from CPU (Kernel Termination and Relaunch (KTRL)) and data transfers between CPU and GPU [1,4]. So our main objective is to reduce these data transfers using different inter block GPU synchronization (IBS) methods resulting in an efficient parallel implementation of SRG. IBS provides flexibility to move all the computations on GPU by providing visibility to updated intermediate data without any intervention from CPU.

* Corresponding author.
E-mail address: el2sasan@uco.es (N. Satpute).

https://doi.org/10.1016/j.cmpb.2020.105430
0169-2607/

Table 1
List of abbreviations with full forms.

List	Full Forms
SRG	Seeded Region Growing
GPU	Graphics Processing Unit
CPU	Central Processing Unit
RoI	Region of Interest
KTRL	Kernel Termination and Relaunch
IBS	Inter Block GPU Synchronization
CT	Computed Tomography
PT	Persistent Threads
SM	Streaming Multiprocessor
CUDA	Compute Unified Device Architecture
DS	Dice Score

In this paper, we propose persistent, grid-stride loop and IBS based GPU approach for SRG to avoid intermediate memory transfers between CPU and GPU. This also reduces processing over unnecessary image voxels providing significant speedup. Persistent thread block (PT) approach is basically dependent on number of active thread blocks and grid-stride loop becomes essential when the number of threads in the grid are not enough to process the image voxels independently [6,7].

We implement parallel image gradient using grid-stride loop and propose gradient and shared memory based fast parallel SRG implemented entirely on GPU without any intermediate transfers between CPU and GPU. This is inspired by parallel processing on static region of interest (RoI) of tiles on GPU. We compare the proposed persistent based parallel SRG with KTRL for accurate vessel segmentation. The gradient based fast parallel SRG for 2D vessel segmentation is 1.9 $\times$ faster compared to the state-of-the-art.

The rest of the paper is structured as follows. Section 2 briefs relevant works and state-of-the-art with respect to SRG. Section 3 explains GPU approaches (KTRL and Static) for SRG implementation using persistence and grid-stride loop. The application of parallel SRG to vessel segmentation is discussed in the Section 4. Performance results and comparison of persistent and grid-stride loop based parallel SRG for vessel segmentation are mentioned in the Section 5. Section 6 concludes summarizing the main conclusions of this paper and indicating future directions. List of abbreviations with explanations are mentioned in Table 1.

2. Background and motivation

There are many works done on image segmentation recently which are based on snake based model [8], gradient vector flow [9,10], and level set based Chan-Vese model [11]. Scientists have explored the snake model for segmentation. Snakes are defined as a set of points around a contour [8]. But the problem with the snake model is that the contour never sees the strong edges that are far away and the snake gets hung up due to many small noises in the image [8]. Hence researchers came up with the solution called gradient vector flow (GVF). In GVF, instead of using image gradient, a new vector field is created over the image plane [9,10]. Cost of GVF includes smoothness and edge map but it requires keeping track of the number of points and point distribution. Hence researchers came up with another solution called as level sets based Chan-Vese model for image segmentation [11]. In the absence of strong edges, a region based formulation for image segmentation is proposed by Chan-Vese model. Chan-Vese model for active contours is a powerful and flexible method which is able to segment many types of images. But amongst all, SRG is the simplest algorithm and plays a vital role in medical image segmentation [1,12].

Smistad et. al. [13,14] have discussed parallel SRG for image segmentation. The reference implementation is shown in Fig. 1. Medical image dataset is cropped before processing. Then the CPU allocates the memory equivalent to the cropped size to copy the data to the cropped image on the GPU. Further SRG is performed for image segmentation. This is the simplistic representation of the work by Smistad et. al. [14]. We have not considered pre-processing stage in this work assuming the images are pre-processed. Smistad et. al. [14] have proposed non persistent thread (non-PT) approach for SRG based vessel segmentation.

Smistad et al. [4] have proposed parallel region growing with double buffering algorithm based on the parallel breadth first search algorithm by Harish and Narayanan [15]. They have suggested a dynamic queue for SRG and mentioned that changing the number of threads (due to border expansion of the region) typically involves restarting the kernel, and this requires reading all the values from global memory again. But they have not recommended probable solution for this problem. Smistad et al. [14] have presented a data parallel version of the SRG based Inverse Gradient Flow Tracking Segmentation algorithm using KTRL. Zhang et al. [16] have implemented bidirectional region growing where they have used a dynamic queue (stack). Jiang et al. [17] have proposed improved branch based region growing vessel segmentation algorithm using stack.

GPU based implementation of SRG needs a dynamic queue (stack). CPUs provide hardware support for stacks but GPUs do not [7]. Any queuing system has a large number of pieces of work to do and a fixed number of workers corresponding to the fixed number of computing units. Pieces are then assigned dynamically to the workers. The problem is deciding the maximum number of pieces of work in the queuing system. If decided, persistent blocks iterate through these pieces of work in the queuing system.

GPU implementation of a stack requires continuous changes in memory allocations which in turn requires iterative GPU kernel invocation from CPU in other words kernel termination and relaunch as discussed in the algorithms IVM backtracking and work stealing phase by Pessoa et al. [18]. Task-parallel run-time system, called TREES, that is designed for high performance on CPU/GPU platforms by Hechtman et al. [19] have shown the invocation of GPU kernels from CPU iteratively for updating task mask stack (TMS) in TREES execution. The loop involved while implementing data flow through the stream kernels of the rendering system (involving stack) on GPU controlled by CPU (that is KTRL) is proposed by Ernst et al. [20].

Nevertheless, there is an alternate GPU implementation of queuing system (stack) using dynamic kernel launching. Chen et. al. [7] have proposed free launch based dynamic kernel launches through thread reuse technique [7]. This technique requires no hardware extensions, immediately deployable on existing GPUs. By turning subkernel launch into a programming feature independent of hardware support, free launch provides alternate approach for subkernel launch which can be used beneficially on GPUs.

KTRL includes terminating a GPU kernel and invoking it from the CPU if the region can be grown further [4,14]. GPU kernel SRG is called from CPU. Region grows from a seed based on the threshold criteria. SRG kernel is terminated and relaunched from CPU if region is not grown completely. This process continues until region can not be grown further. The process involves transfer of data to and fro from CPU and GPU. In KTRL, SRG kernel operates on each voxel of whole image data in all the iterations. It includes redundant memory transfers and unnecessary computations over complete image. Hence the main contributions of this paper are the implementation of persistence based approaches to improve the performance of SRG by reducing unwanted computations and avoiding intermediate memory transfers between CPU and GPU. Memory on the GPU is limited and may not be enough for processing large medical datasets. However, most medical datasets contain a lot of data that is not part of the RoI.

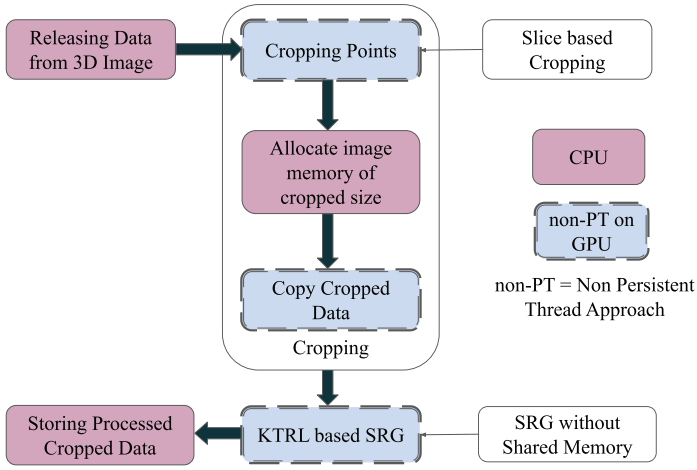


Fig. 1. Reference approach derived from Smistad et. al. [14]

The process of KTRL which involves iterative calling of the kernel is not efficient when implemented on GPU. Hence, as an optimized solution to KTRL, we propose persistent and grid-stride loop based GPU approaches. These approaches are based on processing over static RoI of tiles and dynamic RoI of tiles. We discuss the further details in the upcoming sections.

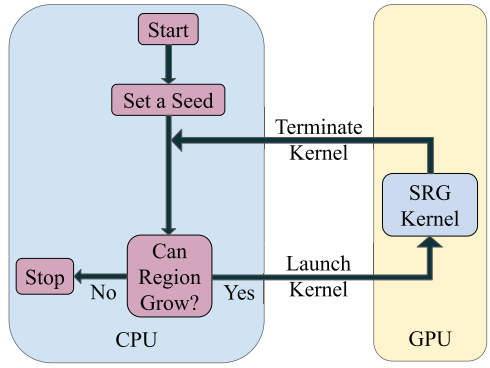
3. Parallel SRG

GPU is a grid of block of threads. Thread is the smallest computational unit mapped on the cores and block of threads are mapped on the streaming multiprocessors (SMs). Each SM can occupy more than one block. The threads from independent blocks can access data via shared memory in the SM [21]. In order to communicate valid data between the blocks, these persistent blocks need to be synchronized via IBS through device memory. Persistence implies maximum number thread blocks that can be active at the time of computation depending upon the GPU resources available [6,22].

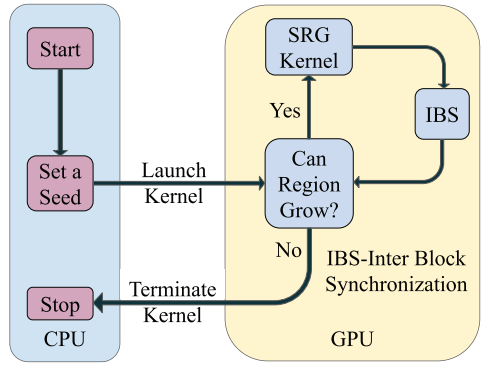
We use PT and shared memory based approaches for SRG implementations. Shared memory and grid-stride loop based SRG reduces total memory transfers and computations. Grid-stride is inspired when the grid is not large enough to occupy all the data elements [23,24]. Rather than assuming that the thread grid is large enough to cover the entire image elements, the kernel loops over the image one grid-size at a time. The stride of the loop is the total number of threads on the grid [23]. These threads (or block of threads) iterate over the image until the process of SRG terminates.

For each thread in parallel on GPU, SRG starts from the seed thread and finds similar neighbours surrounding it Region is grown by making similar neighbouring elements as new seeds. The process of SRG is repeated until similar neighbours can not be found. Normally SRG can be implemented on the GPU as a recursive or iterative kernel calling (KTRL) as shown in Fig. 2a. Kernel calling involves invocation of a grid of block of threads. The blocks are executed on streaming multiprocessors and threads are executed on cores. Park et el. [25] and Smistad et al. [4] have given brief introduction about CUDA (Compute Unified Device Architecture) architecture and GPU computing. They have detailed the information on grid, blocks, threads and memory hierarchy of CUDA architecture.

SRG can be recursive or iterative process. Recursive kernel calling can not utilize GPU cores efficiently due to hardware limita-

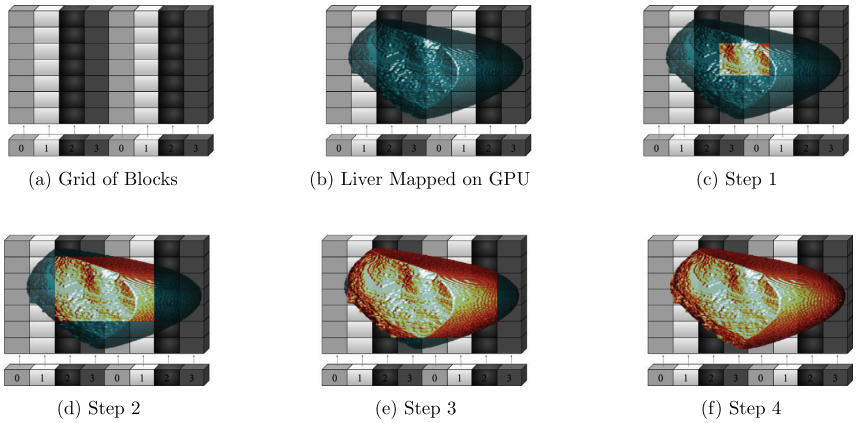


(a) using Kernel Termination and Relaunch [13, 14]



(b) using Proposed Grid-Stride Loop

Fig. 2. GPU implementations of seeded region growing (SRG).



(a) Grid of Blocks (b) Liver Mapped on GPU (c) Step 1

(d) Step 2 (e) Step 3 (f) Step 4

Fig. 3. SRG using persistence and grid-stride loop through complete image.

tions [26]. Iterative GPU kernel call from CPU is costlier due to memory transfers between CPU and GPU and it involves all the image elements to be considered in each step of SRG. GPU implementation of SRG using KTRL is shown in the Fig. 2a. It shows that, the kernel SRG is called on GPU continuously from the host CPU until the region can not be grown further. It starts from the seed, finds similar neighbours and grows the region. This process continues until the region can not be grown further. The process of the KTRL causes unnecessary image elements to be part of computations and intermediate memory transfers between CPU and GPU.

Hence in order to avoid these problems, we propose grid-stride loop through complete image based GPU approach as shown in the Fig. 2b. SRG starts from the seed and the control goes to GPU. The SRG kernel is launched if the region is not grown completely. IBS is needed in order to transfer valid data in between the active thread blocks. The number of active thread blocks on SMs are limited due to resource constraints. These maximum number of active blocks are persistent blocks [6,21,22]. The looping i.e. grid-stride loop terminate when the region can not be grown further and control returns to the host CPU as shown in the Fig. 2b. We have discussed KTRL based GPU approach for SRG implementation and its disadvantages. Now, we are going to analyze PT based GPU approaches for high performance SRG implementation. Proposed approaches exploit parallelism using persistence and IBS as detailed in the static and dynamic approaches.

3.1. Static approach

In the proposed approach, we apply grid-stride loop through static RoI (complete image) using persistence and IBS [6,22]. The complete liver image is mapped on the GPU as grid of block of threads as shown in the Fig. 3b. CPU invokes SRG kernel on GPU. Persistent blocks iterate through complete image and grow region from the seed in each and every iteration on GPU. This iteration of persistent blocks over the tiles of the image and the grid-stride loop based SRG is shown in Fig. 2b. Steps of SRG in Figs. 3c–f show the grown region of the liver. SRG kernel terminates when the region is grown completely. We copy the data from the device memory to the shared memory. This data is shared by all the threads inside the blocks. This is necessary to share the neighbouring elements between different voxels of the image. For each parallel thread in the block, if seed is found and is not the boundary element of the block, we calculate similar neighbouring elements. Region is grown by making similar neighbouring elements as new seeds.

There are four persistent blocks shown in Fig. 3. These four persistent blocks are iterated through liver elements. Tiles with the same color are iterated by same persistent block. In KTRL, these tiles are processed by the thread blocks randomly. Grid-stride loop by persistent blocks is applied on the tiles over complete liver image (Step 1 in Fig. 3c). Region grows around the seed containing similar elements. IBS is applied to communicate valid data in between the blocks for the next step of SRG as shown in Fig. 2b.

Persistent blocks iterate over the liver image and the region is grown again in step 2 as shown in Fig. 3d. IBS is applied and valid data is communicated in between the blocks so that the region can be grown further as shown in Fig. 3e and f. After step 4 in Fig. 3f, SRG stops as region can not be grown further. Each step contain many iterations where region starts growing when persistent blocks iterate through tiles of the image. This iterative process continues until region can not be grown further. Code snapshot of the complete process is provided in the Algorithm 1.

Algorithm 1: Grid-stride loop through complete image.

```
1:  unfinished=1;
2:  while unfinished==1 do
3:      unfinished=0;
4:      for int i=blockIdx.x;i <=
        width/(blockDim.x − 2);i=i+gridDim.x do
5:          for int j=blockIdx.y;j <=
            height/(blockDim.y − 2);j=j+gridDim.y do
6:              for int k=blockIdx.z;k <=
                depth/(blockDim.z − 2);k=k+gridDim.z do
7:                  Region_Growing(arguments, unfinished);
8:              end for
9:          end for
10:     end for
11:     Inter_Block_GPU_Sync();
12: end while
```

Global variable "unfinished" is 1 if region has to be grown further else it is 0. Persistent blocks in *x, y* and *z* directions iterate through complete image. Two is subtracted from block dimensions to avoid computations around boundary voxels (from left and right in each dimensions) from shared memory as region can not be grown further in the block. After each step of SRG, when the processing on complete image is done then all the persistent blocks are globally synchronized via "Inter_Block_GPU_Sync()" barrier. This ensures that valid data is communicated for the next step of SRG. This barrier can be Atomic(), Quasi(), LockFree() or can be implemented using NVIDIA CUDA API Cooperative-groups [22,27,28]. We use quasi based IBS because of its efficicient implementation [27].

The difference between static and dynamic approach by Nitin et. al. [1] can be explained in terms of static and dynamic RoI of tiles. In static approach, RoI remains constant and SRG happens within the constant RoI until the region can not be grown further. Whereas in dynamic approach, SRG starts within the initial RoI. RoI increases and includes more elements uniformly in all the directions for the next step of SRG. SRG takes place, RoI increases and the region is grown further. Hence RoI changes in each step of SRG until the region can not be grown further in a dynamic approach. In the next section, we present 2D vessel segmentation as an application to static RoI based SRG.

4. Application to 2D vessel segmentation

The 2D segmentation algorithm is inspired by the gradient based SRG algorithm developed by Rai and Nair [29]. We proposed the fast parallel SRG based segmentation algorithm on GPU for vessel segmentation. We discuss the two important modules i.e. image gradient and SRG for the fast parallel 2D segmentation of vessels from CT liver images.

4.1. Parallel image gradient

Rai and Nair [29] have presented homogeneity criterion selection and its impact on the quality of segmentation using SRG. In general, the threshold criteria include object contrast, region boundary, homogeneity of the region, intensities values and texture features like shape and color. But we include cost functions mainly based on intensity values and their gradient direction and magnitude.

The cost function exploits certain features of the image around the seed. Gradient based cost function requires gradient of the image, largest gradient magnitude (max_g) and minimum gradient (min_g) present in the image. The cost functions are:

$$cost1 = g/(k * \max\_g) \qquad 0 < cost1 < 1 \tag{1}$$

$$cost2 = (\max\_g - g)/(\max\_g - \min\_g) \qquad 0 < cost2 < 1 \tag{2}$$

where g is gradient magnitude of the pixel under consideration and k is the constant parameter which controls the region growth. The pixel under consideration is added in the growing region if it matches with the seed elements i.e. cost functions specified by Eqs. (1) and (2) are satisfied otherwise it is excluded from consideration.

We propose grid-stride loop based parallel image gradient method in Algorithm 2. For each pixel in parallel, we calculate its gradient magnitude (g) with respect to neighbouring element. Horizontal and vertical gradient components are given by gx and gy. The magnitude of maximum and minimum gradients are updated simultaneously. The gradient of the image is desired input for SRG based segmentation along with the seed. This is discussed in the next section.

Algorithm 2: Parallel image gradient using grid-stride loop.

```
1:  voxel.x = blockIdx.x * blockDim.x + threadIdx.x;
2:  voxel.y = blockIdx.y * blockDim.y + threadIdx.y;
3:  stridex = blockDim.x * gridDim.x;
4:  stridey = blockDim.y * gridDim.y;
5:  for int k=voxel.x; k < rows; k=k+stridex do
6:     for int l=voxel.y; l < cols; l=l+stridey do
7:        candidate.x = k + 1; candidate.y = l + 1;
8:        check if neighbour candidate is within image dimensions;
9:        gx = 0.5*(data[candidate.x*cols + l] - data[k*cols + l]);
10:       gy = 0.5*(data[k*cols + candidate.y] - data[k*cols + l]);
11:       g = sqrt(gx*gx + gy*gy);
12:       data_g[k*cols + l]=g;
13:       if(max_g < g) atomicMax(&max_g, g);
14:       if(min_g > g) atomicMin(&min_g, g);
15:    end for
16: end for
```

4.2. Parallel vessel segmentation

We propose fast parallel vessel segmentation as shown in Fig. 4. The algorithm is inspired from gradient based segmentation algorithm by Rai and Nair [29]. Fig. 4 shows parallel implementation of vessel segmentation where the user selects seed(s). These seed(s) along with the image are transferred to the GPU. Device kernel calculates the image gradient in parallel as discussed in the earlier section. The IBS is necessary to reflect the updated image gradients in the device memory.

Further we apply SRG algorithm. The cost functions based on gradient are shown in Eqs. (1) and (2). For each pixel in parallel, the pixel under consideration invokes SRG kernel if it satisfies the cost functions. The seeds are updated after IBS and the gradient based cost functions are verified again for new pixels. This process continues until no new seeds are formed i.e. no new pixels are added to the growing region.

The kernel is terminated and the control returns to the CPU when the region is grown completely. The segmented image is transferred to the CPU. The process of segmentation stops. This GPU implementation avoids iterative call of SRG kernel from CPU. We use gradient and persistent based parallel SRG for vessel segmentation.

5. Performance evaluation

We propose persistent and grid-stride based GPU approaches for fast parallel 2D vessel segmentation. The performance results are obtained from KTRL and proposed persistent based GPU approach. We compare proposed approaches with KTRL. We use Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz RAM 24 GB, NVIDIA GPU 1050 (RAM 4GB), OpenCL 1.2 (ref. [30]) and CUDA Toolkit 10.1 for the implementation.

5.1. Liver dataset and ground truth

Liver data for the research work has been acquired from The Intervention Center, University of Oslo, Norway [31]. The ground truths for vessel segmentation are provided by the clinician. The modality used is Computed tomography (CT). For the ground truth, images are pre-processed through locally developed application with 3D Slicer to enhance vessels [32]. In some cases, the same application is used for vessel segmentation and separation of portal and hepatic vessels although another possibility is to use active contour tool using ITK-SNAP and manual correction [1,31]. Table 2 shows information about images of different sizes including total

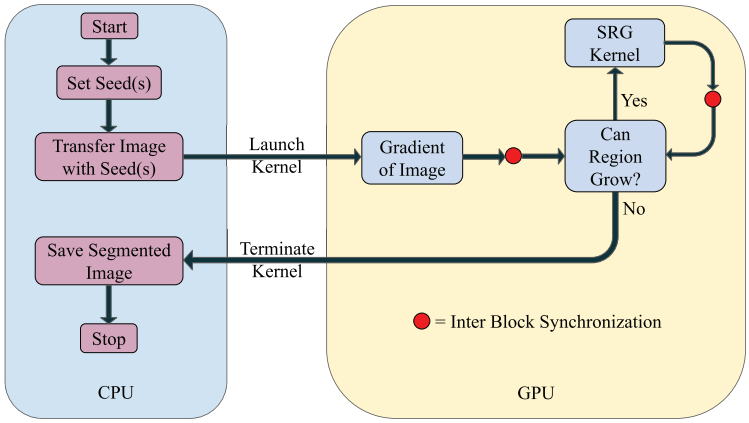


Fig. 4. Proposed parallel vessel segmentation.

Table 2
Liver dataset with vessels.

Volume #	Total # of Slices	Image Size ($w \times h$)	# of Slices with Vessels
10504	59	460 × 306	7
18152	139	512 × 512	5
23186	87	405 × 346	6
28059	59	462 × 321	6

number of vessel slices used for experimentation from a particular volume.

5.2. Parallel 2D vessel segmentation

We propose persistent and grid-stride based GPU approaches for fast parallel 2D vessel segmentation. Variations in vessel segmentation with parameter '*k*' using parallel SRG is shown in Fig. 5. Input to parallel SRG is CT liver slice as shown in Fig. 5a. Gradient of input CT image and the ground truth for the segmentation are shown in Fig. 5b and c respectively. Dice similarity coefficient (DS) and Precision [33,34] are used to assess the quality of vessel segmentation. Dice similarity coefficient measures the similarity between ground truth and the segmented output. If they are identical (i.e. they contain the same elements), the coefficient is equal to 1.0, while if they have no elements in common, it is equal to 0.0. Otherwise it is somewhere in between 0 to 1. Precision describes the number of positive detections with respect to the ground truth. Of all of the elements that are segmented in a given liver vessel image, the number of these elements actually had a matching ground truth annotation can be called as precision.

We show the two segmented vessels with change in parameter *k* i.e. 0.04, 0.05, and 0.06. The first segmented vessel as shown in Fig. 5d, e, f is accurate at 0.04 with high dice score i.e. $DS = 0.77$. Similarly we show the segmentation of second vessel from the same slice. The variations in the quality segmentation due to *k* are shown in Fig. 5g, h, i. The more accurate segmentation is obtained at 0.05 as the dice similarity coefficient value is higher i.e. $DS = 0.60$.

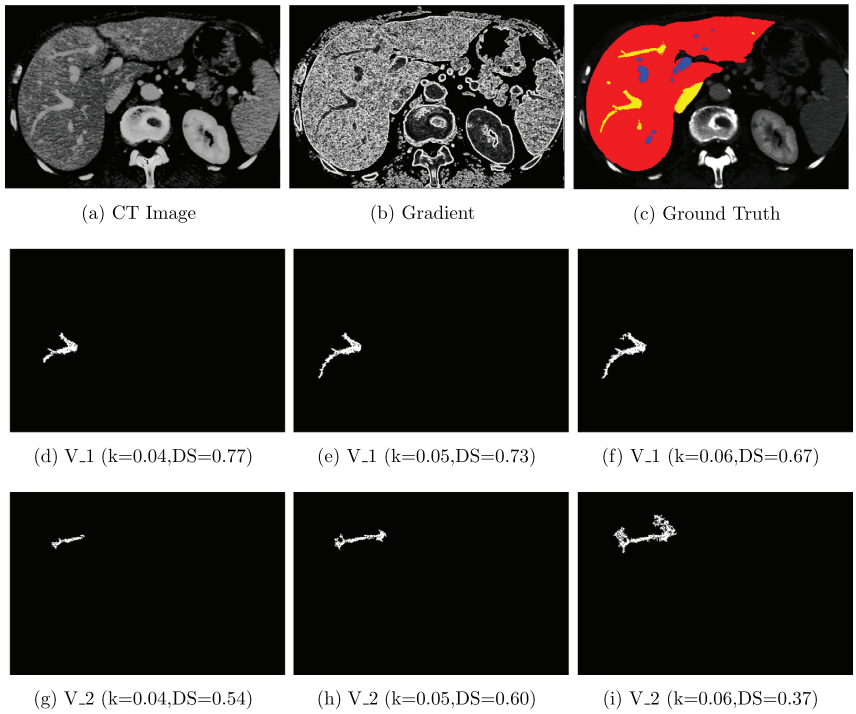
Further we show the quality of segmentation on another CT Slice as shown in Fig. 6a and calculate the gradient (Fig. 6b) of the input CT image. GPU computes parallel SRG using gradient based thresholding criteria giving more accurate results with high dice score at $k = 0.05$ for two vessels inside the CT slice as shown in Fig. 6c and d. The ground truth for the segmentation is shown in Fig. 6e. We analyze that the vessels are more accurately segmented due to better value of dice similarity coefficient when the parameter *k* takes the value 0.05 as shown in Fig. 6.

The speedup obtained by proposed parallel static approach over KTRL on first two CT liver slices are shown in the Table 3. The maximum speedup for vessel segmentation by proposed parallel static SRG is 1.67 × in comparison to KTRL on the first liver slice. But the average speedup obtained by proposed parallel static approach for all the vessels (in 6 slices tested) is 1.9 × compared to KTRL. We evaluate the speedup of the vessel segmentation when the vessel segmentation is more accurate with better Dice score value (Fig. 5).

Further we analyze the effect of parallel SRG on different slices for multiple vessel segmentation using multiple seeds as shown in Figs. 7–10. Segmentation of the long vessel as shown in Fig. 8d is slightly extended compared to the ground truth shown in Fig. 8e. It can be seen from input CT image and gradient image (Figs. 8a and b), long vessel has extension which is not shown in the ground truth. We show the thick vessel segmentation in Figs. 7d, 9 c, 8 c and thin vessel segmentation in Figs. 7c, 8 c and d. The results of segmentation in terms of Dice Score and Precision are provided in Table 4. The highest and lowest value of precision for these slices are 0.94 and 0.79 respectively. This implies the number of positive detections in the segmented images are higher. It is possible to use multiple seeds for the same vessel in the proposed vessel segmentation approach. We have the flexibility to provide two seeds on the same vessel and then the proposed approach can create a curve or a line of initial seeds (if needed) as an input for SRG. It is useful in order to increase the quality of vessel segmentation.

There are many works done on image segmentation recently which are based on snake based model [8], gradient vector flow [9,10], and level set based Chan-Vese model [11]. We validate the performance on 72 vessels from 24 vessel slices obtained from 4 different volumes in Table 5. Table shows the comparison of the vessel segmentation accuracy between three different models i.e. Snake model [8], Chan-Vese [11], and proposed SRG in terms of dice score and precision. Average dice score value and precision obtained by proposed SRG is outperforming the Chan-Vese and Snake based vessel segmentation.

Our proposed parallel implementations of SRG is not only fast but also accurate for vessel segmentation. The accuracy of the segmentation depends on the parameter '*k*'. Clinicians get the flexibility to decide which segmentation is more accurate. The process takes very less time (few ms). Hence this reduces the overall time for segmentation for various values of parameter '*k*' if the



(a) CT Image (b) Gradient (c) Ground Truth

(d) V_1 (k=0.04,DS=0.77) (e) V_1 (k=0.05,DS=0.73) (f) V_1 (k=0.06,DS=0.67)

(g) V_2 (k=0.04,DS=0.54) (h) V_2 (k=0.05,DS=0.60) (i) V_2 (k=0.06,DS=0.37)

Fig. 5. Variations in fast parallel vessel segmentation with constant parameter '*k*' using parallel SRG on first liver slice.

(a) CT Image (b) Gradient (c) V_1, DS=0.88 (d) V_2, DS=0.89 (e) Ground Truth

Fig. 6. Vessel segmentation (for $k = 0.05$) using parallel SRG on second liver slice.

Table 3
Time and speedup for vessel segmentation.

Data →	Vessel Segmentation	
GPU Approaches → Metrics ↓	KTRL	Static (Speedup)
Time in ms for kernel SRG - 1st Slice ($k = 0.05$) 1st vessel	5.7	3.4 (1.67 ×)
Time in ms for kernel SRG - 1st Slice ($k = 0.05$) 2nd vessel	2.1	1.5 (1.4 ×)
Time in ms for kernel SRG - 2nd Slice ($k = 0.05$) 1st vessel	1.5	1 (1.5 ×)
Time in ms for kernel SRG - 2nd Slice ($k = 0.05$) 2nd vessel	3.5	2.4 (1.45 ×)

(a) CT Image (b) Gradient (c) n=2, DS=0.70 (d) n=4, DS=0.88 (e) Ground Truth

Fig. 7. Fast parallel vessel segmentation using parallel SRG on 3rd liver slice using multiple seeds (n).

Fig. 8. Fast parallel vessel segmentation using parallel SRG on 4th liver slice using multiple seeds (n).

Fig. 9. Fast parallel vessel segmentation using parallel SRG on 5th liver slice using multiple seeds (n).

Fig. 10. Fast parallel vessel segmentation using parallel SRG on 6th liver slice using multiple seeds (n).

Table 4
Quality of vessel segmentation in terms of dice score and precision for Figs. 5–10.

Sr. No.	Ground Truth Figure	Segmented Image Figure	Value of 'k'	Dice Score	Precision
1	Fig. 5c	Fig. 5d	0.04	0.77	0.86
2		Fig. 5e	0.05	0.73	0.94
3		Fig. 5f	0.06	0.67	0.94
4	Fig. 5c	Fig. 5g	0.04	0.54	0.86
5		Fig. 5h	0.05	0.60	0.83
6		Fig. 5i	0.06	0.37	0.86
7	Fig. 6e	Fig. 6c	0.05	0.88	0.86
8		Fig. 6d	0.05	0.89	0.85
9	Fig. 7e	Fig. 7c	0.05	0.70	0.91
10		Fig. 7d	0.05	0.88	0.91
11	Fig. 8e	Fig. 8c	0.05	0.71	0.79
12		Fig. 8d	0.05	0.72	0.90
13	Fig. 9e	Fig. 9c	0.05	0.67	0.82
14		Fig. 9d	0.05	0.88	0.88
15	Fig. 10e	Fig. 10c	0.05	0.89	0.94
16		Fig. 10d	0.05	0.67	0.91

Table 5
Segmentation accuracy comparison for 4 volumes, 24 vessel slices and 72 vessels.

Volume #	Image Size ($w \times h$)	# of Vessel Slices	Total # of Vessels	Chan-Vese [11]		Snake model [8]		Proposed SRG	
				Average Dice	Average Precision	Average Dice	Average Precision	Average Dice	Average Precision
10504	460 × 306	7	21	0.78	0.82	0.77	0.81	0.85	0.83
18152	512 × 512	5	15	0.75	0.85	0.78	0.83	0.82	0.87
23186	405 × 346	6	18	0.77	0.83	0.81	0.82	0.84	0.86
28059	462 × 321	6	18	0.72	0.81	0.76	0.78	0.81	0.84

clinician wants to have more accuracy. The advantage of k is, by adjusting the value of k from the threshold criteria, clinicians have the flexibility to fine tune the accuracy of the vessel segmentation. But, vessels vary in shapes, sizes, texture features etc not only in different CT slices but also in the same CT slice. Even if the authors propose $k = 0.05$ provides better accuracy for the provided CT images, sometimes finding the same value of k for different vessels in the same CT slice becomes difficult. The range of k lies between 0.03 to 0.12 for the better accuracy of vessel segmentation.

5.3. Discussion

In this paper, we propose persistence and grid-stride loop based SRG implementation. In order to obtain significant speedup, we need to exploit parallelism by using persistence and IBS. It involves change in the large body of SRG algorithm. We want algorithms that require as less synchronization as possible. In general if algorithm requires IBS, it is probably not going to be particularly fast. The fastest algorithms on GPUs are ones that fit nicely into the GPU programming model, where blocks are independent from each other and do not require synchronization [35].

But the problem arises when iterative calling of the kernel can not be avoided. It incurs memory transfers from CPU to GPU when KTRL is used for global synchronizations. Hence it has to go through synchronizations as the next step of SRG which is dependent on the current step. Terminating a kernel and relaunching incurs data transfers from CPU to GPU and vice versa. It is time consuming.

If we use IBS method along with persistence, then we can map whole algorithm on GPU with synchronization. Control comes back to CPU only if the kernel task is over. CPU launches a kernel on GPU, GPU executes it and final results are copied to CPU. No intermediate data communication occurs in the proposed approach (unlikely in KTRL).

6. Conclusion

In this paper, we discuss SRG based vessel segmentation and its parallel implementation on GPU. We propose persistence and grid-stride loop based GPU approach for SRG providing significant speedup. Normally recursion/iterative calling of a kernel is generally a bad idea on GPUs. We use persistence and grid-stride approach as an alternate implementation for KTRL. We compare proposed GPU optimization strategy for SRG implementation. The proposed persistent and gradient based parallel SRG for 2D vessel segmentation is accurate with high dice scores and 1.9 $\times$ faster compared to the KTRL.

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgements

The work is supported by the project High Performance soft tissue Navigation (HiPerNav). This project has received funding from the European Union Horizon 2020 research and innovation program under grant agreement No. 722068. We thank The Intervention Centre, Oslo University Hospital, Oslo, Norway for providing the CT images with ground truths for the clinical validation of vessel segmentation.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2020.105430.

References

[1] N. Satpute, R. Naseem, E. Pelanis, J. Gomez-Luna, F. Alaya Cheikh, O.J. Elle, J. Olivares, GPU acceleration of liver enhancement for tumor segmentation, Comput. Methods Programs Biomed. 184 (2020) 105285, doi:10.1016/j.cmpb.2019.105285.

[2] R. Palomar, F.A. Cheikh, B. Edwin, Å. Fretland, A. Beghdadi, O.J. Elle, A novel method for planning liver resections using deformable Bézier surfaces and distance maps, Comput. Methods Programs Biomed. 144 (2017) 135–145.

[3] K.K. Delibasis, A. Kechriniotis, I. Maglogiannis, A novel tool for segmenting 3d medical images based on generalized cylinders and active surfaces, Comput. Methods Programs Biomed. 111 (1) (2013) 148–165, doi:10.1016/j.cmpb.2013.03.009.

[4] E. Smistad, T.L. Falch, M. Bozorgi, A.C. Elster, F. Lindseth, Medical image segmentation on GPUs–a comprehensive review, Med. Image Anal. 20 (1) (2015) 1–18.

[5] J. Wassenberg, W. Middelmann, P. Sanders, An efficient parallel algorithm for graph-based image segmentation, in: International Conference on Computer Analysis of Images and Patterns, Springer, 2009, pp. 1003–1010.

[6] K. Gupta, J.A. Stuart, J.D. Owens, A study of persistent threads style GPU programming for GPGPU workloads, in: Innovative Parallel Computing-Foundations & Applications of GPU, Manycore, and Heterogeneous Systems (INPAR 2012), IEEE, 2012, pp. 1–14.

[7] G. Chen, X. Shen, Free launch: optimizing GPU dynamic kernel launches through thread reuse, in: Proceedings of the 48th International Symposium on Microarchitecture, ACM, 2015, pp. 407–419.

[8] S. Roy, S. Mukhopadhyay, M.K. Mishra, Enhancement of morphological snake based segmentation by imparting image attachment through scale-space continuity, Pattern Recognit. 48 (7) (2015) 2254–2268.

[9] H. Zhou, X. Li, G. Schaefer, M.E. Celebi, P. Miller, Mean shift based gradient vector flow for image segmentation, Comput. Vis. Image Underst. 117 (9) (2013) 1004–1016.

[10] E. Smistad, A.C. Elster, F. Lindseth, Real-time gradient vector flow on GPUs using OpenCL, J. Real-Time Image Process. 10 (1) (2015) 67–74.

[11] S.K. Siri, M.V. Latte, Combined endeavor of neutrosophic set and Chan-Vese model to extract accurate liver image from ct scan, Comput. Methods Programs Biomed. 151 (2017) 101–109.

[12] R.P. Kumar, F. Albregtsen, M. Reimers, B. Edwin, T. Langø, O.J. Elle, Three--dimensional blood vessel segmentation and centerline extraction based on two-dimensional cross-section analysis, Ann. Biomed. Eng. 43 (5) (2015) 1223–1234.

[13] E. Smistad, Seeded region growing, 2015, (https://github.com/smistad/FAST/tree/master/source/FAST/Algorithms/).

[14] E. Smistad, A.C. Elster, F. Lindseth, GPU accelerated segmentation and centerline extraction of tubular structures from medical images, Int. J. Comput. Assisted Radiol. Surg. 9 (4) (2014) 561–575.

[15] P. Harish, P.J. Narayanan, Accelerating large graph algorithms on the GPU using CUDA, in: S. Aluru, M. Parashar, R. Badrinath, V.K. Prasanna (Eds.), High Performance Computing – HiPC 2007, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 197–208.

[16] X. Zhang, X. Li, Y. Feng, A medical image segmentation algorithm based on bi-directional region growing, Optik 126 (20) (2015) 2398–2404.

[17] H. Jiang, B. He, D. Fang, Z. Ma, B. Yang, L. Zhang, A region growing vessel segmentation algorithm based on spectrum information, Comput. Math. Methods Med. 2013 (2013).

[18] T.C. Pessoa, J. Gmys, N. Melab, F.H. de Carvalho Junior, D. Tuyttens, A GPU-based backtracking algorithm for permutation combinatorial problems, in: J. Carretero, J. Garcia-Blas, R.K. Ko, P. Mueller, K. Nakano (Eds.), Algorithms and Architectures for Parallel Processing, Springer International Publishing, Cham, 2016, pp. 310–324.

[19] B.A. Hechtman, A.D. Hilton, D.J. Sorin, TREES: a CPU/GPU task-parallel runtime with explicit epoch synchronization, arXiv:1608.00571. (2016).

[20] M. Greiner, Stack implementation on programmable graphics hardware, in: Vision, Modeling, and Visualization 2004: Proceedings, 2004, p. 255.

[21] V. Vineet, P.J. Narayanan, CUDA cuts: Fast graph cuts on the GPU, in: 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, pp. 1–8, doi:10.1109/CVPRW.2008.4563095.

[22] S. Xiao, W.C. Feng, Inter-block GPU communication via fast barrier synchronization, in: 2010 IEEE International Symposium on Parallel Distributed Processing (IPDPS), 2010, pp. 1–12, doi:10.1109/IPDPS.2010.5470477.

[23] M. Harris, Cuda pro tip:write flexible kernels with grid-stride loops, 2015.

[24] M. Sourouri, S.B. Baden, X. Cai, Panda: a compiler framework for concurrent CPU+GPU execution of 3d stencil computations on GPU-accelerated supercomputers, Int. J. Parallel Program. 45 (3) (2017) 711–729.

[25] S. Park, J. Lee, H. Lee, J. Shin, J. Seo, K.H. Lee, Y.-G. Shin, B. Kim, Parallelized seeded region growing using CUDA, Comput. Math. Methods Med. 2014 (2014).

[26] X. Tang, A. Pattnaik, H. Jiang, O. Kayiran, A. Jog, S. Pai, M. Ibrahim, M.T. Kandemir, C.R. Das, Controlled kernel launch for dynamic parallelism in GPUs, in: 2017 IEEE International Symposium on High Performance Computer Architecture (HPCA), 2017, pp. 649–660, doi:10.1109/HPCA.2017.14.

[27] Y. Komura, Y. Okabe, GPU-based single-cluster algorithm for the simulation of the ising model, J. Comput. Phys. 231 (4) (2012) 1209–1215.

[28] T. Sorensen, H. Evrard, A.F. Donaldson, Cooperative kernels: GPU multitasking for blocking algorithms, in: Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering, ACM, 2017, pp. 431–441.

[29] G. Rai, T. Nair, Gradient based seeded region grow method for ct angiographic image segmentation, arXiv:1001.3735. (2010).

[30] J.E. Stone, D. Gohara, G. Shi, OpenCL: a parallel programming standard for heterogeneous computing systems, IEEE Des. Test 12 (3) (2010) 66–73, doi:10.1109/MCSE.2010.69.

[31] Å.A. Fretland, V.J. Dagenborg, G.M.W. Bjørnelv, A.M. Kazaryan, R. Kristiansen, M.W. Fagerland, J. Hausken, T.I. Tønnessen, A. Abildgaard, L. Barkhatov, et al., Laparoscopic versus open resection for colorectal liver metastases, Ann. Surg. 267 (2) (2018) 199–207.

[32] R. Naseem, F.A. Cheikh, A. Beghdadi, O.J. Elle, F. Lindseth, Cross modality guided liver image enhancement of CT using MRI, in: 2019 8th European Workshop on Visual Information Processing (EUVIP), IEEE, 2019, pp. 46–51.

[33] Y. Zhao, H. Li, S. Wan, A. Sekuboyina, X. Hu, G. Tetteh, M. Piraud, B. Menze, Knowledge-aided convolutional neural network for small organ segmentation, IEEE J. Biomed. Health Inf. 23 (4) (2019) 1363–1373, doi:10.1109/JBHI.2019.2891526.

[34] D.N.H. Thanh, D. Sergey, V.B.S. Prasath, N.H. Hai, Blood vessels segmentation method for retinal fundus images based on adaptive principal curvature and image derivative operators, Int. Arch.Photogramm. Remote Sens. Spatial Inf. Sci. XLII-2/W12 (2019) 211–218, doi:10.5194/isprs-archives-xlii-2-w12-211-2019.

[35] O. Zachariadis, A. Teatini, N. Satpute, J. Gómez-Luna, O. Mutlu, O. Jakob Elle, J. Olivares, Accelerating B-spline Interpolation on GPUs: Applica-tion to Medical Image Registration, Comput. Methods Programs Biomed. (2020), doi:10.1016/j.cmpb.2020.105431.

Chapter 9

Paper D: GPU acceleration of liver enhancement for tumor segmentation

Computer Methods and Programs in Biomedicine 184 (2020) 105285

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

GPU acceleration of liver enhancement for tumor segmentation

Nitin Satpute[a,*], Rabia Naseem[b], Egidijus Pelanis[c,d], Juan Gómez-Luna[e], Faouzi Alaya Cheikh[b], Ole Jakob Elle[c,f], Joaquín Olivares[a]

[a] Department of Electronic and Computer Engineering, Universidad de Córdoba, Spain
[b] Norwegian Colour and Visual Computing Lab, Norwegian University of Science and Technology, Norway
[c] The Intervention Centre, Oslo University Hospital, Oslo, Norway
[d] The Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
[e] Department of Computer Science, ETH Zurich, Switzerland
[f] The Department of Informatics, The Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway

ARTICLE INFO

Article history:
Received 7 October 2019
Revised 27 November 2019
Accepted 16 December 2019

Keywords:
2D Histogram matching
Contrast enhancement
GPU
Image segmentation

ABSTRACT

Background and objective: Medical image segmentation plays a vital role in medical image analysis. There are many algorithms developed for medical image segmentation which are based on edge or region characteristics. These are dependent on the quality of the image. The contrast of a CT or MRI image plays an important role in identifying region of interest i.e. lesion(s). In order to enhance the contrast of image, clinicians generally use manual histogram adjustment technique which is based on 1D histogram specification. This is time consuming and results in poor distribution of pixels over the image. Cross modality based contrast enhancement is 2D histogram specification technique. This is robust and provides a more uniform distribution of pixels over CT image by exploiting the inner structure information from MRI image. This helps in increasing the sensitivity and accuracy of lesion segmentation from enhanced CT image. The sequential implementation of cross modality based contrast enhancement is slow. Hence we propose GPU acceleration of cross modality based contrast enhancement for tumor segmentation.

Methods: The aim of this study is fast parallel cross modality based contrast enhancement for CT liver images. This includes pairwise 2D histogram, histogram equalization and histogram matching. The sequential implementation of the cross modality based contrast enhancement is computationally expensive and hence time consuming. We propose persistence and grid-stride loop based fast parallel contrast enhancement for CT liver images. We use enhanced CT liver image for the lesion or tumor segmentation. We implement the fast parallel gradient based dynamic seeded region growing for lesion segmentation.

Results: The proposed parallel approach is 104.416 (± 5.166) times faster compared to the sequential implementation and increases the sensitivity and specificity of tumor segmentation.

Conclusion: The cross modality approach is inspired by 2D histogram specification which incorporates spatial information existing in both guidance and input images for remapping the input image intensity values. The cross modality based liver contrast enhancement improves the quality of tumor segmentation.

1. Introduction

Computed tomography (CT) images of abdomen often possess low contrast [1,2]. Radiologists often manually delineate lesions during segmentation of medical images, which can be difficult, time-consuming and prone to observer variability [3]. Some segmentation algorithms do not perform well when applied on the CT images and are time consuming [4,5]. However, their performance can be made better once the CT images are preprocessed [6,7]. Therefore, preprocessed CT images help in refining the lesions. One possible preprocessing step is image enhancement for the better visualization of tumors in undertaking surgical procedures [8–10].

Efficient preprocessing can certainly help to attain accurate segmentation of the critical structures in medical images [7,11]. High sensitivity and specificity indicates the improved quality of the segmentation [5,12]. The liver images obtained from the CT scans are sometimes noisy, low in contrast and contains high amounts of details. We consider contrast as the important feature. If the image is high contrast then it becomes easier to identify and segment the

* Corresponding author.
E-mail address: el2sasan@uco.es (N. Satpute).

https://doi.org/10.1016/j.cmpb.2019.105285
0169-2607/

object of interest [2,13]. In our case, lesion is necessary to be segmented.

There are many methods proposed to improve the contrast of the image. Histogram equalization, histogram specification and histogram matching are some of the ways to improve the contrast in the image as discussed by [1,14,15]. We apply 2D histogram matching where CT liver is the target image and the magnetic resonance imaging (MRI) liver slice is the guided image [16,17]. Cross modality based contrast enhancement exploits 2D histogram matching for liver enhancement. Once the image is enhanced then the task is to segment tumor from enhanced image. Seeded region growing for tumor segmentation is an easy and effective process. But the task of cross modality based liver enhancement is computationally expensive and time consuming [18]. Hence it becomes necessary to use GPU for real time performance of liver contrast enhancement and tumor segmentation. We propose accelerated cross modality guided liver enhancement scheme in this paper and demonstrate that our technique improves tumor segmentation on enhanced image.

The aim of this study is cross modality based liver enhancement to improve the contrast of CT liver image for tumor segmentation. We propose parallel implementation of liver contrast enhancement. This is accomplished by 2D histogram matching using CT and MRI liver images. We propose dynamic region of interest (RoI) based seeded region growing (SRG) for tumor segmentation from enhanced CT image. The overall average speedup obtained by parallel implementation is 104.416 ± 5.166 times compared to the sequential CPU implementation of the contrast enhancement and tumor segmentation. The enhanced liver image improves the sensitivity and specificity of the lesion segmentation. This is the first work targeted towards the high performance multi-modality guided liver enhancement for tumor segmentation to the best of our knowledge.

The rest of the paper is organized as follows. Section 2 briefs the related works, background and motivation with respect to the liver image enhancement. Section 3 explains the proposed methodology for liver contrast enhancement and its parallel implementation on the GPU. Further, we discuss dynamic RoI based fast parallel SRG for tumor segmentation in Section 4. Performance results and comparison of contrast enhancement and seeded region growing for tumor segmentation are mentioned in the Section 5. Section 6 concludes summarizing the main results related to the cross modality based contrast enhancement and tumor segmentation.

2. Background and motivation

Segmentation of lesions is a challenging problem in medical images because of the similar intensity values of structures of interest and the nearby regions in image. Research works are targeting various methods for the segmentation [19–21]. The results of the segmentation are subsequently used in patient specific model for diagnostics, surgery planning and navigation. One such approach using gradient based SRG has been presented to segment the aorta and rib bones in thorax images by Rai and Nair [21]. Inspired by this idea, we propose parallel SRG to segment tumors from CT liver images.

Image enhancement is regarded as a precursor to the accurate segmentation. CT scans are commonly used due to the availability and quicker imaging time compared to MRI. CT scans often suffer from low contrast which limit their utility [1,2]. In this work, we show through our experiments that corresponding MR image can be employed to improve the contrast of CT. The idea to enhance an image using another cross modal image has been witnessed in the literature for natural images [6–9]. The motivation to use cross modality guided image enhancement is to use the additional information contained in the other image having similar contents in different imaging times or position but better contrast or minimal noise. Ultimately, the details in the enhanced image can be improved from the perceptual perspective. In the context of liver images, tumors can be easily seen in the enhanced CT image.

In this regard, the contrast of photographs was improved using the corresponding near infra red images [6,14]. Histogram specification in combination with wavelet domain processing was used in this work. Yan et. al proposed a variational approach using anisotropic filter to eliminate noise in color images using infrared images [9]. The authors calculated cross correlation between input images and then used joint filtering for denoising in another approach [7,11].

Deep learning is applied to multimodal image denoising recently [8]. A deep learning method consisting of three convolutional neural networks has been applied to denoise natural images. Various deep learning based approaches for CT denoising have been presented in the last few years, however, they do not incorporate the multimodality guidance and use the CT image alone for supervised learning [16,17,22]. Histogram based methods are useful to enhance the global contrast of image [14], however, they introduce bad artifacts in the processed images. Since it does not consider the neighborhood of the pixels while remapping, it does not necessarily gives the desired contrast [2,14,23]. Two dimensional histogram specification is presented recently to improve the 1D histogram specification [18]. It uses 2D cumulative distribution function of the input and target images for remapping intensity values in the original image.

We apply same notion to CT liver images by applying 2D histogram matching based cross modality approach for liver contrast enhancement in the following section.

3. Methodology: liver contrast enhancement

We aim to improve the contrast of CT liver image considering MRI liver image as the guidance image to increase the quality of lesion segmentation. The methodology includes 2D contrast enhancement, gradient of enhanced image and segment the lesion using gradient based SRG. The parallel approach for liver enhancement and lesion segmentation makes the process faster in order to achieve real time implementation. In this section, we discuss parallel implementation of the cross modality based liver enhancement.

The flow of proposed GPU implementation of cross modality based contrast enhancement is shown in Fig. 1. We load CT and MRI images of liver on CPU and transfer it to the GPU. The first step of contrast enhancement of CT liver image is 2D (or pairwise) histogram calculation (Hist_2d). We calculate parallel 2D histogram of both CT (hist_CT) and MRI (hist_MRI) images. A 2D histogram is a plot of pixel and its neighbouring element which allows us to discover, and show, the underlying 2D frequency distribution (shape) of image. This shows how often each set of values (pixel and neighbour) in the image occurs. Instead of just considering the individual pixel values, it considers every possible pixel pair in the input and guidance image and calculate 2D CDF accordingly [18,24].

Further the calculation of cumulative distributive function (CDF_2d) of CT (CDF_CT) and MRI (CDF_MRI) images on GPU creates the input for the next step i.e. histogram equalization. 2D CDF is a function that describes the probability of a possible pixel pair in the input and guidance image. This helps in finding most frequent pairwise intensity values for histogram equalization [18].

Then we perform parallel histogram equalization (HE_2d). This step spreads out the most frequent pairwise intensity values increasing the global contrast of image. Hence it improves lower contrast areas to gain a higher contrast [18,24].

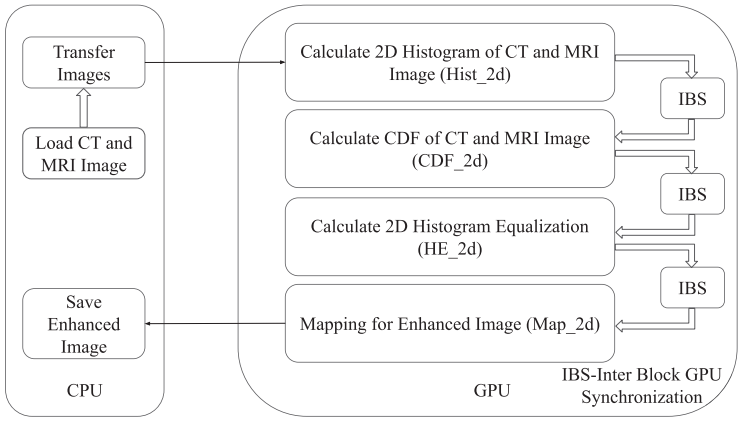


Fig. 1. GPU implementation of the cross modality based contrast enhancement.

The mapping (Map_2d) of histogram equalization onto the CT image gives the enhanced image. It maps the modified intensity values obtained from 2D histogram equalization to the corresponding pixels [18].

Inter block GPU synchronization (IBS) makes sure the updated values are sent to the next modules in GPU computing blocks. These parallel implementations of sub-modules of contrast enhancement are explained in following sections.

3.1. 2D Histogram

In this section, we discuss the 2D histogram implementation on GPU as the first step of the contrast enhancement of CT liver image. The histogram length (HL) is 256. We launch HLxHL parallel threads and find the histogram of neighboring elements in pairs. Hence it is called as pairwise histogram. Pairwise histogram is stored in an array of size HLxHL.

For each thread in parallel, it takes the pixel (x,y) and neighbouring pixel (x+1,y) value. This represents one of the indices in the range of (0-HLxHL-1) in histogram array given by variable temp as shown in Algorithm 1. We increment corresponding value

Algorithm 1: 2D Histogram of CT and MRI Image (Hist_2d).

```
HL=256 and launch HL x HL parallel threads
ti and tj can be any thread id between 0–255
while x<width_of_image do
    while y<height_of_image do
        if ti == I[x][y] and tj == I[x + 1][y] then
            temp=ti*HL+tj;
            atomicAdd(histogram[temp], 1);
        end if
    end while
end while
```

in the index position in histogram array as shown in Fig. 2. This function hist_2d for CT and MRI images gives hist_CT and hist_MRI histograms respectively. These 2D histograms are the input to the cumulative distributive function which is the next step of contrast enhancement.

3.2. Cumulative distributive function (CDF)

In this step of contrast enhancement, we calculate CDF of 2D histograms of CT and MRI liver images. The maximum number of

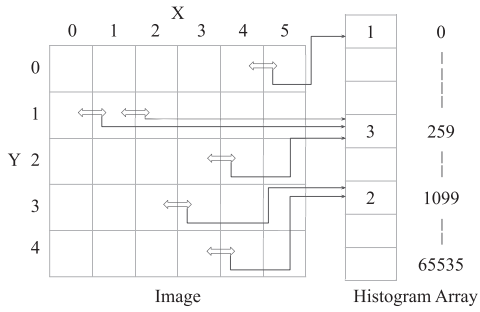


Fig. 2. 2D Histogram.

histogram pairs can be (w−1)×(h) where w and h are width and height of the image.

We launch HL × HL threads in parallel as shown in Algorithm 2. Each thread calculates its CDF from respective 2D histogram values. These CDF values for CT (CDF_CT) and MRI (CDF_MRI) images are the input to the next step of contrast enhancement which is 2D histogram equalization.

Algorithm 2: Calculate CDF of CT and MRI Image (CDF_2d).

```
count= (width-1)*height i.e. maximum number of pairs
HL=256 and launch HL x HL parallel threads
ti and tj can be any thread id between 0–255
temp=ti*HL+tj;
while temp<HL*HL do
    for int j=0; j<=temp; j++ do
        cdf[temp]+ = histogram[j]/count;
    end for
end while
```

3.3. 2D Histogram equalization (HE_2d)

2D Histogram Equalization technique improves the contrast of image. It spreads out the most frequent intensity values. This method increases the global contrast of image. This improves the lower contrast areas to gain higher contrast. The pseudocode for 2D histogram equalization is shown in the Algorithm 3. We launch

Algorithm 3: Calculate 2D Histogram Equalization (HE_2d).

```
HL=256 and launch HL x HL parallel threads
ti and tj can be any thread id between 0–255
index=ti*HL+tj;
for k=0; k<HL; k++ do
   for l=0; l<HL; l++ do
      temp8=k*HL+l;
      temp = cdf1[index]-cdf2[temp8]
      if temp is minimum then
         x = k
      end if
      if multiple minimum values found then
         temp2 = absolute((ti-k) + (tj-l))
         if temp2 is minimum then
            x = k
         end if
         if multiple minimum temp2 are found then
            temp3 = absolute((ti-tj) - (k-l))
            if temp3 is maximum then
               x = k
            end if
         end if
      end if
   end for
end for
HE[index]=x;
```

HLxHL threads in parallel. Each thread calculates the corresponding histogram equalization by taking the minimum difference between two CDFs (cdf1 for CDF_CT and cdf2 for CDF_MRI). It takes into the account the first minimum euclidean distance value between the indices when multiple minimum difference in CDFs are found. Again when multiple solutions are available, it further computes and find out the equalized value saved in array HE. This array is ready to get mapped for enhanced image which is the final step of contrast enhancement.

3.4. Mapping

The mapping of 2D histogram equalization is essential for obtaining enhanced CT image as an output. We launch wxh threads where w and h are width and height of the image respectively. This is reverse process of 2D histogram calculation as explained in the psuedocode given by the Algorithm 4. The index value is

Algorithm 4: Mapping for Enhanced Image (Map_2d).

```
launch (width)*(height) parallel threads
HL=256
tw can be any thread id between 0 to width-1
th can be any thread id between 0 to height-1
temp1 = I[tw][th];
temp2 = I[tw + 1][th];
index = temp1 * HL + temp2;
I[tw][th] = HE[index]; //EnhancedImage
```

generated from the neighbouring pixel values of the CT image. The pixel value in the CT image is changed by the corresponding value in the location (index) given by the 2D histogram equalization array. When all the threads are finished processing corresponding pixels, the enhanced image is sent back to the CPU.

4. Application to the tumor segmentation

Seeded Region Growing is an easy approach to segment the various objects in an image. The result of the region growing relies mainly on the initial seed(s) and the criteria defined to end recursive or iterative region growing process [4,19,25,26]. The parallel implementation of SRG based tumor segmentation is shown in Fig. 3.

We load CT and MRI images and transfer it to the GPU. GPU performs cross modality based contrast enhancement and stores the enhanced CT image in GPU memory. The control comes back to the CPU. This is essential for the selection of seed(s) and to change the number of persistent blocks. These persistent blocks (i.e. number of available computing resources on the GPU) differ depending on the application. The next task is tumor segmentation. GPU computes the gradient of enhanced CT liver image. The gradient of enhanced liver image is communicated through IBS to the next module for tumor segmentation. We apply SRG on the gradient of enhanced liver image. Region grows and new seeds are formed from initial seed(s) based on the threshold criteria. This process is iterative until the threshold criteria is satisfied. The process stops when new seed(s) can not be formed and region can not be grown further.

In this work, we use threshold criteria defined by the homogeneity of region and region aggregation considering the pixel values and their gradient direction and magnitude. The criteria is defined via a cost function that uses few features of the image around seed. Value of the cost function is compared with homogeneity criteria specified to check if the value is smaller than 1. The pixel becomes part of the region if there is a match; otherwise it is excluded from the region. The cost functions for threshold criteria are given by Rai and Nair [21]. They select homogeneity criterion using gradient based cost function which are dependent upon object contrast, texture features like shape and color, intensities values, gradient direction and magnitude. The cost function exploits features of image around the seed.

We apply parallel gradient based SRG algorithm on both enhanced images and original CT liver images. We propose dynamic RoI based parallel SRG.

4.1. Dynamic SRG

Dynamic SRG as the name suggests, it increases the region of interest (RoI) in each iteration of SRG. The initial RoI is decided by number of active computing blocks or persistent blocks that can be launched on GPU. This represents the phenomenon of persistence. In order to communicate valid data in between the blocks, inter block GPU synchronization (IBS) is necessary. Persistence and IBS provide flexibility to exploit parallelism using grid-stride loop through constant increase in RoI. One grid-stride is number of active computing threads that can be launched on GPU device.

Gupta et al. [27] have explored persistent thread based GPU programming. The idea behind this is once the SRG kernel launched from CPU, the control returns from GPU when the region is grown completely. Intermediate data transfers between CPU and GPU are avoided in this approach. SRG kernel on GPU is launched from the host CPU. Region is grown on GPU. Image elements are updated and communicated to the blocks via IBS. The region is grown again on GPU, if new similar neighbouring elements are found. This process continues until no similar neighbouring elements are available. The kernel terminates when the region can not be grown further and control returns to the CPU. Redundant data computations and communications are optimized on GPU using proposed approach. This process is explained in the Fig. 4.

There are four persistent blocks processing grid of blocks using grid-stride loop as shown in Fig. 4a. We map 3D liver on grid of

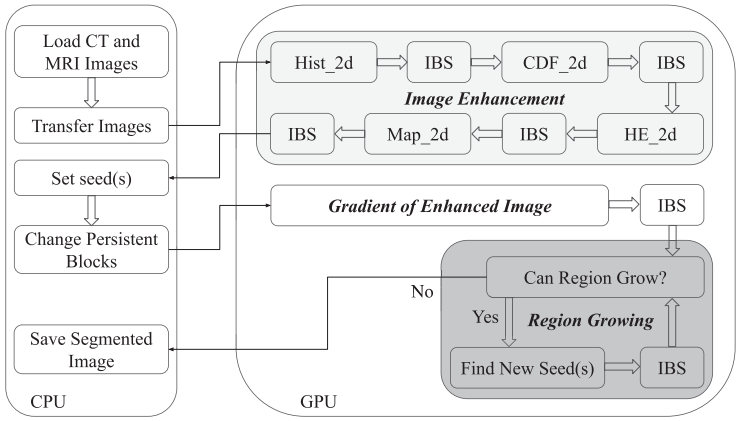
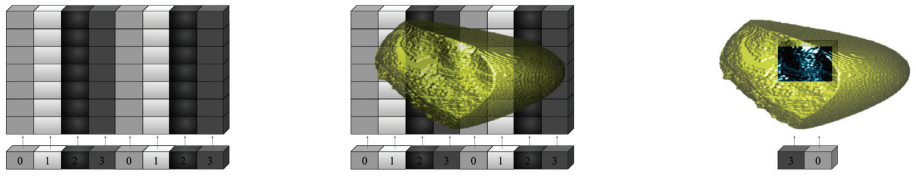


Fig. 3. GPU implementation of SRG based tumor segmentation.



(a) Grid of Blocks (b) Liver Mapped on GPU (c) Step 1

(d) Step 2 (e) Step 3 (f) Step 4

Fig. 4. SRG using dynamic RoI of tiles.

blocks as shown in Fig. 4b and initialize RoI of tiles around the seed as shown in Fig. 4c. Persistent blocks operate within RoI. First step of SRG takes place. Region is grown and RoI is incremented in all directions. This process makes necessary neighbouring voxels available for the second step of SRG as shown in Fig. 4d. New neighbouring voxels perform same function and RoI is incremented again. This flow is repeated until region can not be grown further as shown in Fig. 4e and f. This approach reduces compute and memory operations resulting in the increased performance. It is needed to ensure that the increase in RoI lies within the image dimensions.

Complete process is defined in the Algorithm 5 . RoI should be initialized in such a way that all threads are busy performing SRG. Variable “blockgrow” is essential to check the increase the RoI. Increase RoI of tiles if value of “blockgrow” is ”1”, otherwise stop SRG as region is grown completely. This variable “blockgrow” along with the variable “unfinished” are updated in the SRG segmentation step. Lower and upper values of RoI (in x, y, and z directions) are calculated when “blockgrow” is “1”. It has to be made sure that the RoI should not increase beyond image dimensions in the successive steps of SRG.

Persistent blocks operate inside the RoI. Kernel SRG is called for the voxels within the RoI. IBS makes sure only updated values are communicated to the persistent blocks in each step of SRG. IBS can be atomic, quasi, lock free or based on cooperative groups from NVIDIA toolkit CUDA 10.1 [28–30]. We use quasi IBS for our approach due to its efficient implementation [28].

5. Results and discussion

We discuss performance analysis of proposed parallel cross modality based liver enhancement for tumor segmentation. The enhanced liver images and segmented tumors are shown and the performance analysis of tumor segmentation is discussed based on quality assessment. We use Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz RAM 24 GB, NVIDIA GPU GeForce GTX 1050 (RAM 4GB), and CUDA Toolkit 10.1 to compare the proposed parallel GPU approach with CPU implementation.

Algorithm 5: Grid-stride Loop through Dynamic RoI.

```
blockgrow=1;
while blockgrow==1 do
    blockgrow=0;
    unfinished=1;
    Increase RoI of Tiles;
    To Increase RoI of Tiles
    w=w+1; h=h+1; d=d+1;
    Ensure RoI within image dimensions;
    while unfinished==1 do
        unfinished=0;
        for int i=blockIdx.x;i<=w/blockDim.x;i+=gridDim.x do
            for int j=blockIdx.y;j<=h/blockDim.y;j+=gridDim.y do
                for int k=blockIdx.z;k<=d/blockDim.z;k+=gridDim.z
                do
                    Region_Growing(arguments, unfinished,
                    blockgrow);
                end for
            end for
        end for
        Inter_Block_GPU_Sync();
    end while
end while
```

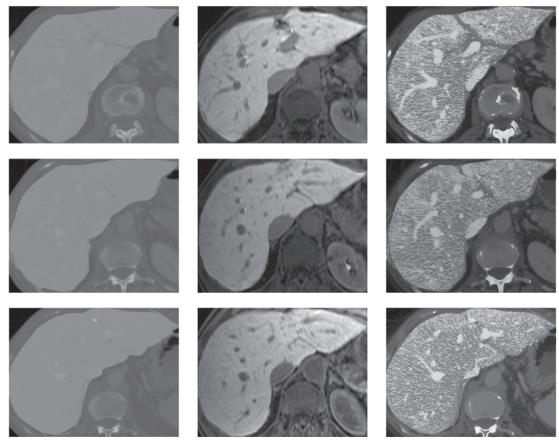


Fig. 5. CT, MR and enhanced CT images.

5.1. Liver enhancement

We propose fast parallel cross modality based contrast enhancement. 2D histogram of CT image is mapped to 2D histogram of guidance or MR image to get a better contrast image.

Fig. 5 shows input CT, MRI and enhanced CT liver images without any tumors. Fig. 6 shows enhanced CT liver images with tumors. Figures show the contrast is enhanced significantly to observe tumors clearly. Enhanced image is further processed for tumor segmentation using SRG. Average time taken by NVIDIA GPU GeForce GTX 1050 is 1.976 s ± 0.43 s providing the average speedup of 104.416 ± 5.166 times over CPU implementation (208.082s ± 55.799s) for tumor segmentation using 2D cross modality based contrast enhancement.

In order to enhance the contrast in CT images, we investigate quality improvements by fusing the information that is available in one modality (e.g. liver inner structures in MRI) to guide the adaptive enhancement in other image modality (e.g. CT in our case). This provides better control over the enhancement and is more effective and efficient than the state of the art technique used by clinicians. Clinicians generally use manual histogram adjustment technique based on 1D histogram specification on CT or MRI scans. This process does not provide efficient distribution of pixels for contrast enhancement of CT or MRI image. There are more chances of artifacts in 1D enhancement as it results in random histogram and is also a time consuming process.

However, 2D histogram specification incorporates spatial information while calculating 2D CDFs of both the guidance and input

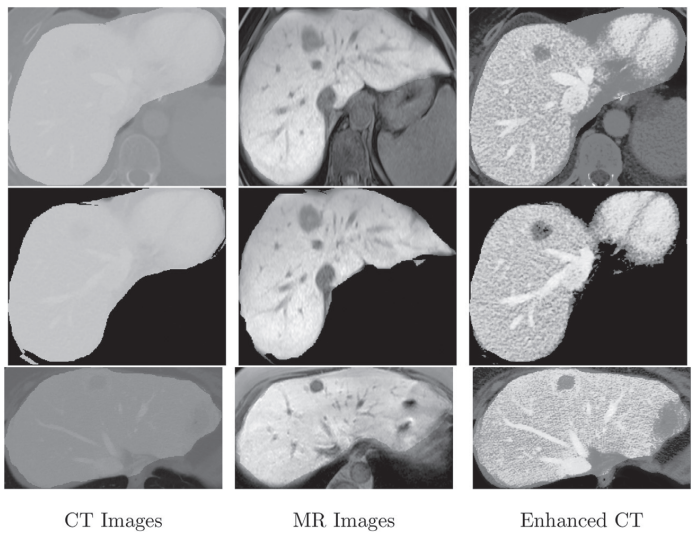
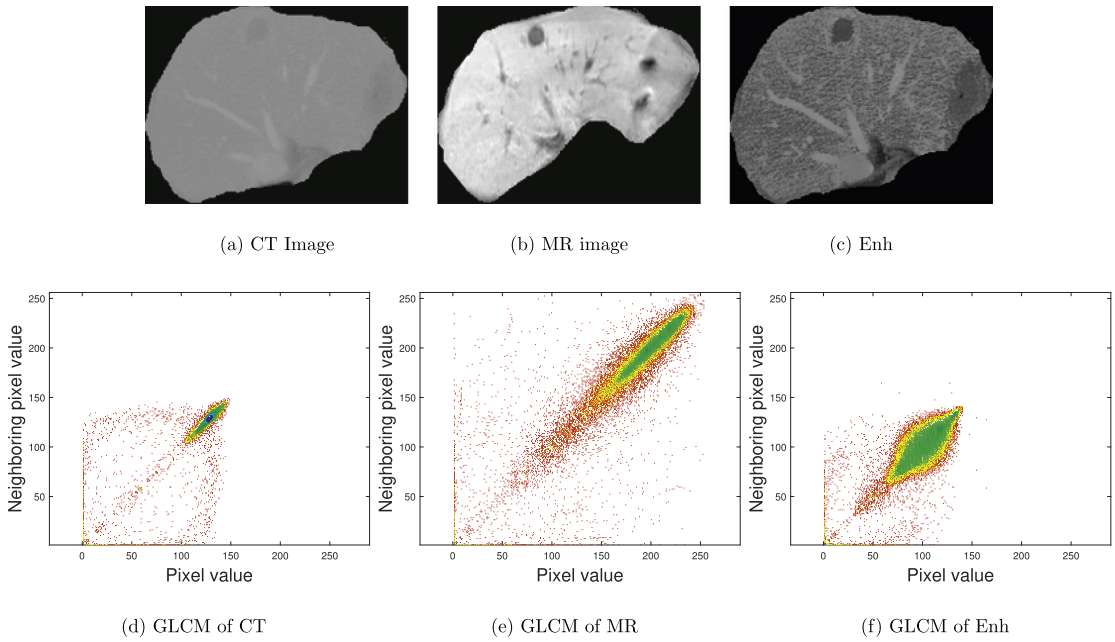


Fig. 6. CT, MR and enhanced CT images showing tumors.



(a) CT Image (b) MR image (c) Enh

(d) GLCM of CT (e) GLCM of MR (f) GLCM of Enh

Fig. 7. CT, MR and enhanced CT (Enh) with GLCM plots.

images and for remapping the input image intensity values. Instead of just considering the individual pixel values, it considers every possible pixel pair in the input and guidance image and calculate 2D CDF accordingly. Looking at the Gray Level Co-Occurrence Matrix (GLCM) plots in Fig. 7, it can be observed that the distribution of pixel pairs in GLCM plot of the resulting enhanced image (Fig. 7f) is expanded but concentrated along the diagonal in comparison to GLCM plots of CT and MR image (Fig. 7d and e), which means it does not introduce artificial artifacts unlike 1D histogram specification or histogram equalization.

We provide the histogram comparison of images using 1D and proposed 2D technique as shown in Fig. 8. The proposed 2D cross modality approach provides a proper distribution of pixel elements using guided MRI compared to 1D approach applied on CT or MRI image. 1D approach introduces unpleasant effects in the enhanced image. The histogram of enhanced CT using cross modality approach is similar to guided MRI image. There are more chances of artifacts in enhanced image using 1D approach as clinicians use manual adjustment which may result in any random histogram of the enhanced image. In the next section, we discuss the impact of cross modality based contrast enhancement for tumor segmentation.

5.2. Tumor segmentation

We propose fast parallel gradient based dynamic SRG for tumor segmentation. Our proposed parallel SRG is implemented on GPU. It does not involve transfer of data between CPU and GPU. The data for the research work have been acquired from The Intervention Center, University of Oslo, Norway [31]. The ground truths for tumor segmentation are provided by the clinician. We present the visual comparison of tumor segmentation on both enhanced and original CT liver images. The results in Figs. 9–11 show the tumor segmentation from original and enhanced liver images. Fig. 9a1 represents the original CT liver image. The gradient of input CT image is shown in Fig. 9a2. The tumor segmentation (Seg) and the ground truth (GT) for the original CT liver slice are shown in Fig. 9a3 and a4 respectively.

We enhance original CT liver image (Fig. 9a1) using cross modality based liver enhancement and the enhanced image (Enh_CT) is shown in Fig. 9b3. The tumor segmentation is performed on the enhanced CT liver image (Fig. 9b3) and segmented tumor from enhanced CT image is shown in Fig. 9b5. The quality of tumor segmentation is validated in our clinical validation section using Table 1. Tumor segmentation for other CT liver slices are shown in Figs. 10, and 11 and the segmentation quality is improved when the image is enhanced. Hence the cross modality based contrast enhancement on CT liver images improves the quality of tumor segmentation and it is faster. The proposed fast parallel liver enhancement based tumor segmentation is 104.416 ± 5.166 times faster compared to the sequential implementation. We include Table 2 showing experimental evaluation on 10 different datasets (including 107 tumor slices) obtained from The Intervention Centre, Oslo University Hospital, Oslo, Norway. It can be observed from the table that the cross modality based liver enhancement helps in improving the sensitivity, specificity (denoted by 'Sensi' and 'Speci' respectively in Table 2) and accuracy of tumor segmentation and GPU implementation of proposed approach is around 100 times faster compared to the CPU implementation. P value from ANOVA (analysis of variance) for the ten datasets is 3.31×10^{-14} which is less than 0.05. We reject the null hypothesis and conclude that not all means are equal which confirms the means are statistically significant for the concerned experiments.

5.3. Clinical validation

Tables 1 and 2 show the analysis of tumor segmentation before and after enhancement of CT liver images. Table 1 includes 5 liver slices with tumors from different datasets and Table 2 shows performance evaluation on 10 different datasets including 107 tumor

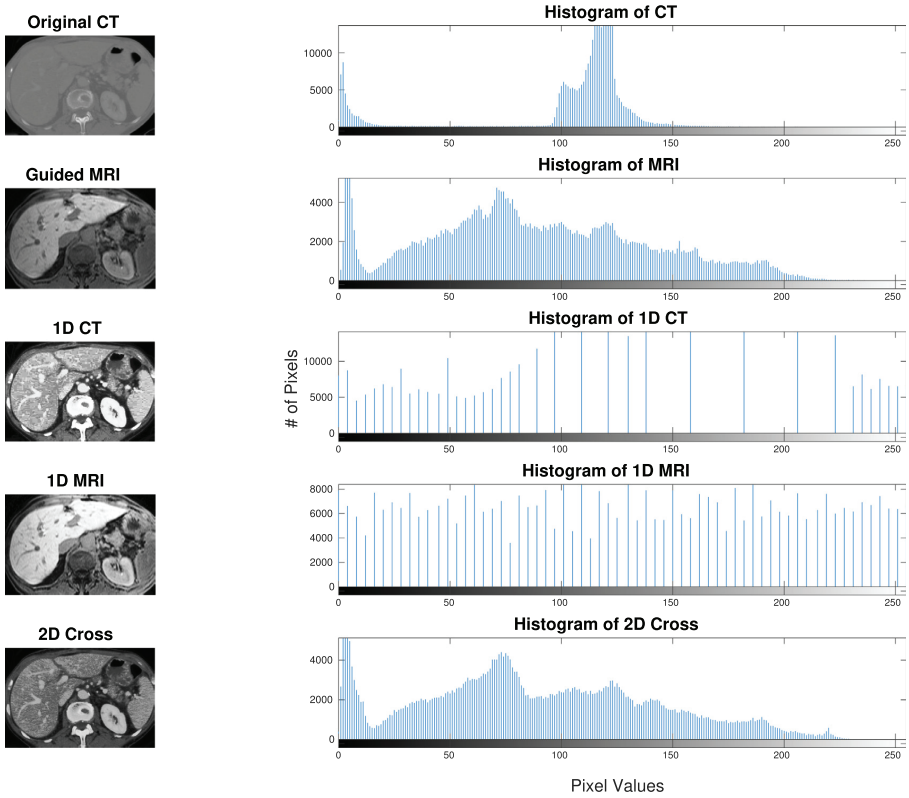
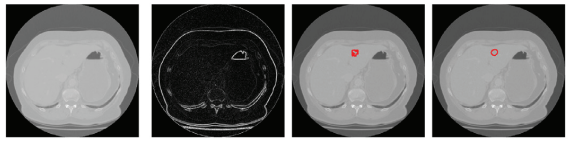
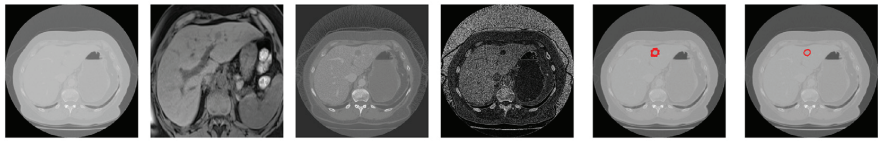


Fig. 8. Comparison between 2D cross modality and 1D histogram approach.



(a1) CT Image (a2) Gradient (a3) Seg (a4) GT

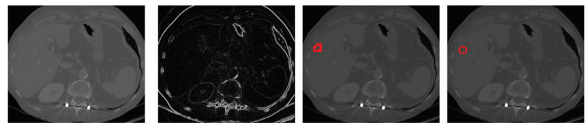
(a) Original Liver Image 1



(b1) CT Image (b2) MR (b3) Enh_CT (b4) Gradient (b5) Seg (b6) GT

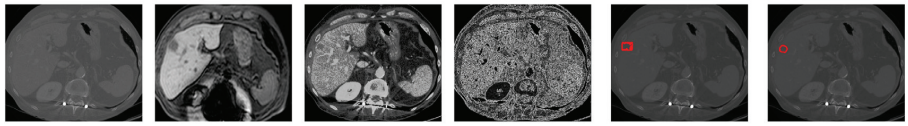
(b) Liver Enhancement and Tumor Segmentation from Enhanced CT Image 1

Fig. 9. Tumor segmentation from original and enhanced CT image 1.



(a1) CT Image (a2) Gradient (a3) Seg (a4) GT

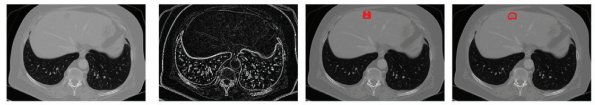
(a) Original Liver Image 2



(b1) CT Image (b2) MR (b3) Enh_CT (b4) Gradient (b5) Seg (b6) GT

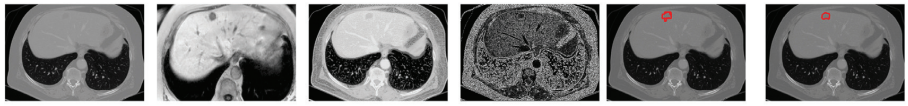
(b) Liver Enhancement and Tumor Segmentation from Enhanced CT Image 2

Fig. 10. Tumor segmentation original and enhanced CT image 2.



(a1) CT Image (a2) Gradient (a3) Seg (a4) GT

(a) Original Liver Image 3



(b1) CT Image (b2) MR (b3) Enh_CT (b4) Gradient (b5) Seg (b6) GT

(b) Liver Enhancement and Tumor Segmentation from Enhanced CT Image 3

Fig. 11. Tumor segmentation from original and enhanced CT image 3.

Table 1
Tumor segmentation analysis on five slices.

Tumor	Without any Enhancement		With Enhancement		Time-Enh+SRG(s)		Speedup
Slice #	Sensitivity, Specificity	Accuracy	Sensitivity, Specificity	Accuracy	CPU	GPU	
1	0.55	0.99899	0.82	0.99906	272.07	2.48	109.706
2	0.38	0.99918	0.81	0.99898	265.98	2.41	110.365
3	0.47	0.99769	0.58	0.9968	167.81	1.68	99.887
4	0.83	0.87091	0.50	0.99765	162.03	1.61	100.64
5	0.47	0.99786	0.74	0.99823	172.52	1.70	101.482
Average	0.54	0.973	0.69	0.998	208.082s	1.976s	104.416
Std. Dev.	0.173	0.057	0.143	0.001	55.799s	0.43s	5.166

slices. We chose sensitivity (true positive rate or recall) and specificity (true negative rate) as performance metrics for the evaluation of tumor segmentation [5,12]. It is observed that, the sensitivity and specificity are increased when the accuracy is nearly 1 on the enhanced image. This implies that when the tumor is actually present, then it is predicted more accurately when the image is enhanced.

5.4. Discussion

In this paper, we propose fast parallel cross modality based contrast enhancement for CT liver images. Further GPU performs dynamic RoI based tumor segmentation on enhanced CT liver image. These fast parallel implementations are based on persistence, grid-stride loop and IBS. The process of cross modality based contrast

Table 2
Tumor segmentation analysis on ten different datasets.

Dataset #	Size of each Slice (wxh)	Total # of Slices	# of Tumor Slices	Without any Enh (Average)		With Enh - Average (Avg.)		Enh+SRG Avg. Time (s)		Avg. Speedup
				Sensi, Speci	Model accuracy	Sensi, Speci	Model accuracy	CPU	GPU	
1	406 × 299	73	10	0.28	0.99132	0.36	0.99517	141.07	1.41	100.054
2	512 × 512	139	7	0.41	0.99213	0.52	0.99796	252.22	2.29	109.901
3	381 × 304	67	10	0.48	0.99412	0.65	0.99689	131.89	1.32	99.916
4	405 × 346	87	8	0.39	0.99325	0.47	0.99717	158.56	1.56	101.641
5	462 × 321	59	14	0.32	0.99173	0.50	0.99823	167.01	1.63	102.460
6	380 × 512	58	9	0.49	0.99112	0.64	0.99421	202.02	1.89	106.89
7	443 × 437	63	6	0.51	0.99201	0.71	0.99501	193.17	1.83	105.55
8	361 × 249	63	7	0.37	0.99312	0.57	0.99427	126.60	1.26	100.47
9	483 × 386	80	6	0.31	0.99415	0.59	0.99612	185.78	1.80	103.21
10	456 × 400	216	30	0.42	0.99178	0.62	0.99324	189.93	1.82	104.35

enhancement is computationally expensive and hence time consuming. This involves 2D histogram calculation, equalization and histogram matching [22]. They require several light weight tasks. The performance on GPU is improved compared to the CPU by dividing the tasks on several active threads.

The second part of the process is tumor segmentation. We propose gradient and dynamic RoI based SRG inspired from the works of Rai and Nair [21]. Initially, the process needs small part of the region to be accessed instead of whole image (as implemented previously on GPU). As soon as region grows, RoI should be increased to access more neighbouring elements. GPU implementation of SRG involves kernel termination and relaunch continuously from CPU. This is time consuming. We avoid this by using persistence and grid-stride loop and obtain the significant speedup i.e. 104.416 ± 5.166 times compared to the sequential implementation of liver enhancement and tumor segmentation.

6. Conclusion

In this paper, we discuss cross modality based contrast enhancement for CT liver images, application to tumor segmentation and their fast parallel implementation on GPU. Cross modality based liver enhancement includes CT liver image as an input and MRI liver image as a guided image. Pairwise 2D histogram implementation and histogram equalization spreads the intensity values across the image producing contrast enhanced CT image. We propose persistence and grid-stride loop based fast parallel implementation on GPU. The enhanced image then used for segmentation of tumors from enhanced CT liver images effectively. We propose gradient and dynamic RoI based seeded region growing for tumor segmentation. The parallel approach for liver enhancement and tumor segmentation is 104.416 ± 5.166 times faster compared to the CPU implementation.

Declaration of Competing Interest

We, the undersigned, confirm that the manuscript represents our own work, is original and has not been copyrighted, published, submitted, or accepted for publication elsewhere. We further confirm that we all have fully read the manuscript and give consent to be co-authors of the manuscript.

Acknowledgement

The work is supported by the project High Performance soft tissue Navigation (HiPerNav). This project has received funding from the European Union Horizon 2020 research and innovation program under grant agreement No. 722068. We thank The Intervention Centre, Oslo University Hospital, Oslo, Norway for providing images with ground truths for the clinical validation of tumor segmentation.

References

[1] S. Fu, M. Zhang, C. Mu, X. Shen, Advancements of medical image enhancement in healthcare applications, J. Healthc. Eng. 2018 (2018).
[2] H. Hassanpour, N. Samadiani, S.M. Salehi, Using morphological transforms to enhance the contrast of medical images, Egypt. J. Radiol. Nucl.Med. 46 (2) (2015) 481–489.
[3] Y. Zhao, H. Li, S. Wan, A. Sekuboyina, X. Hu, G. Tetteh, M. Piraud, B. Menze, Knowledge-aided convolutional neural network for small organ segmentation, IEEE J. Biomed. Health Inform. 23 (4) (2019) 1363–1373, doi:10.1109/JBHI.2019.2891526.
[4] R. Palomar, F.A. Cheikh, B. Edwin, Å. Fretland, A. Beghdadi, O.J. Elle, A novel method for planning liver resections using deformable Bézier surfaces and distance maps, Comput. Methods Program. Biomed. 144 (2017) 135–145.
[5] M.H. Yap, G. Pons, J. MartÃ, S. Ganau, M. SentÃs, R. Zwiggelaar, A.K. Davison, R. MartÃ, Automated breast ultrasound lesions detection using convolutional neural networks, IEEE J. Biomed. Health Inform. 22 (4) (2018) 1218–1226, doi:10.1109/JBHI.2017.2731873.
[6] S. Zhuo, X. Zhang, X. Miao, T. Sim, Enhancing low light images using near infrared flash images, in: International Conference on Image Processing 2010, IEEE, 2010, pp. 2537–2540.
[7] X. Shen, C. Zhou, L. Xu, J. Jia, Mutual-structure for joint filtering, in: Proceedings of the International Conference on Computer Vision, 2015, IEEE, 2015, pp. 3406–3414.
[8] Y. Li, J.-B. Huang, N. Ahuja, M.-H. Yang, Deep joint image filtering, in: European Conference on Computer Vision, 2016, Springer, 2016, pp. 154–169.
[9] Q. Yan, X. Shen, L. Xu, S. Zhuo, X. Zhang, L. Shen, J. Jia, Cross-field joint image restoration via scale map, in: ICCV, 2013, IEEE, 2013, pp. 1537–1544.
[10] W. Zhu, H. Jiang, E. Wang, Y. Hou, L. Xian, J. Debnath, X-ray image global enhancement algorithm in medical image classification, Discr. Cont. Dyn. Syst. 12 (2019).
[11] K. He, J. Sun, X. Tang, Guided image filtering, IEEE Trans. Pattern Anal. Mach.Intell. 35 (6) (2012) 1397–1409.
[12] Z. Yan, X. Yang, K. Cheng, A three-stage deep learning model for accurate retinal vessel segmentation, IEEE J. Biomed. Health Inform. 23 (4) (2019) 1427–1436, doi:10.1109/JBHI.2018.2872813.
[13] V. Jaya, R. Gopikakumari, Iem: a new image enhancement metric for contrast and sharpness measurements, Int. J. Comput. Appl. 79 (9) (2013).
[14] S.S. Bhairannawar, Efficient medical image enhancement technique using transform HSV space and adaptive histogram equalization, in: Soft Computing Based Medical Image Analysis, Elsevier, 2018, pp. 51–60.
[15] B. Subramani, M. Veluchamy, MRI brain image enhancement using brightness preserving adaptive fuzzy histogram equalization, Int. J. Imaging Syst.Technol. 28 (3) (2018) 217–222.
[16] J.M. Wolterink, T. Leiner, M.A. Viergever, I. Išgum, Generative adversarial networks for noise reduction in low-dose CT, IEEE Trans. Med. Imag. 36 (12) (2017) 2536–2545.
[17] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M.K. Kalra, Y. Zhang, L. Sun, G. Wang, Low-dose ct image denoising using a generative adversarial network with Wasserstein distance and perceptual loss, IEEE Trans. Med. Imag. 37 (6) (2018) 1348–1357.
[18] S.-W. Jung, Two-dimensional histogram specification using two-dimensional cumulative distribution function, Electron. Lett. 50 (12) (2014) 872–874.
[19] K.K. Delibasis, A. Kechriniotis, I. Maglogiannis, A novel tool for segmenting 3d medical images based on generalized cylinders and active surfaces, Comput. Methods Program. Biomed. 111 (1) (2013) 148–165, doi:10.1016/j.cmpb.2013.03.009.
[20] E. Smistad, Seeded Region growing, 2015. (https://github.com/smistad/FAST/tree/master/source)
[21] G. N. Rai, T. R. Nair. "Gradient based seeded region grow method for CT angiographic image segmentation." arXiv preprint arXiv:1001.3735 (2010).

[22] H. Chen, Y. Zhang, M.K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, G. Wang, Low–dose CT with a residual encoder-decoder convolutional neural network, IEEE Trans. Med. Imag. 36 (12) (2017) 2524–2535.
[23] G. Qi, S. Shu-Ting, R. Ping-Chuan, Medical image enhancement algorithm based on improved contourlet, J. Med. Imag. Health Inform. 7 (5) (2017) 962–967.
[24] T. Celik, Two-dimensional histogram equalization and contrast enhancement, Pattern Recognit. 45 (10) (2012) 3810–3824, doi:10.1016/j.patcog.2012.03.019.
[25] E. Smistad, M. Bozorgi, F. Lindseth, Fast: framework for heterogeneous medical image computing and visualization, Int. J. Comput. Assist.Radiol. Surg. 10 (11) (2015) 1811–1822.
[26] E. Smistad, A.C. Elster, F. Lindseth, Gpu accelerated segmentation and centerline extraction of tubular structures from medical images, Int. J. Comput. Assist. Radiol. Surg. 9 (4) (2014) 561–575.
[27] K. Gupta, J.A. Stuart, J.D. Owens, A study of persistent threads style GPU programming for GPGPU workloads, in: Innovative Parallel Computing-Foundations & Applications of GPU, Manycore, and Heterogeneous Systems (INPAR 2012), IEEE, 2012, pp. 1–14.
[28] Y. Komura, Y. Okabe, Gpu-based single-cluster algorithm for the simulation of the ising model, J. Comput. Phys. 231 (4) (2012) 1209–1215.
[29] S. Xiao, W.C. Feng, Inter-block GPU communication via fast barrier synchronization, in: IEEE International Symposium on Parallel Distributed Processing (IPDPS), 2010, 2010, pp. 1–12, doi:10.1109/IPDPS.2010.5470477.
[30] T. Sorensen, H. Evrard, A.F. Donaldson, Cooperative kernels: GPU multitasking for blocking algorithms, in: Proceedings of the 11th Joint Meeting on Foundations of Software Engineering, 2017, ACM, 2017, pp. 431–441.
[31] Å. Fretland, et al., Laparoscopic Versus Open Resection for Colorectal Liver Metastases: The OSLO-COMET Randomized Controlled Trial, Annals of Surgery (2018), doi:10.1097/SLA.0000000000002353.

Chapter 10

Paper E: Contrast Enhancement: Cross-modal Learning Approach for Medical Images

Contrast Enhancement: Cross-modal Learning Approach for Medical Images

Rabia Naseem [1], ***Akib Jayed Islam*** [1,2], ***Faouzi Alaya Cheikh*** [1], ***Azeddine Beghdadi*** [2],
[1] ***Norwegian Colour and Visual Computing Lab, Norwegian University of Science & Technology, Norway,*** [2] ***University Jean Monnet Saint-Etienne, France,*** [3] ***L2TI, Institut Galilée, Université Sorbonne Paris Nord, France***
rabia.naseem@ntnu.no, akibji@stud.ntnu.no, faouzi.cheikh@ntnu.no, azeddine.beghdadi@univ-paris13.fr

Abstract

Contrast is an imperative perceptible attribute embodying the image quality. In medical images, the poor quality specifically low contrast inhibits precise interpretation of the image. Contrast enhancement is, therefore, applied not merely to improve the visual quality of images but also enabling them to facilitate further processing tasks. We propose a contrast enhancement approach based on cross-modal learning in this paper. Cycle-GAN (Generative Adversarial Network) is used for this purpose, where U-Net augmented with global features acts as a generator. Besides, individual batch normalization has been used to make generators adapt specifically to their input distributions. The proposed method accepts low contrast T2-weighted (T2-w) Magnetic Resonance images (MRI) and uses the corresponding high contrast T1-w MRI to learn the global contrast characteristics. The experiments were conducted on a publicly available IXI dataset. Comparison with recent CE methods and quantitative assessment using two prevalent metrics FSIM and BRISQUE validate the superior performance of the proposed method.

Introduction

Different degradations are introduced during the image acquisition phase that reduces the lucidity of important details and ultimately affect the extraction of valuable information [26, 13]. Contrast Enhancement (CE) is a primary operation that allows the digital images to be visually perceptible. In the context of medical images, one of the objectives of CE is to improve the perceptual quality for superior visualization of specific structures [15]. Another objective is to facilitate feature extraction and other subsequent tasks such as detection and segmentation of critical structures [20, 21]. It has been reported that the performance of segmentation and detection in medical images can be augmented by employing effective pre-processing techniques on low-contrast images [16, 25].

It is important to mention that a single medical image does not carry complete structural information of the organ under inspection. Multi-modal image acquisition is therefore becoming a standard clinical practice [17]. It not only endorses the initial diagnosis, moreover, it also provides complementary information that can play an influential role in several stages of diagnosis and treatment. The multi-modal image information has been utilized to solve various problems in medical imaging such as segmentation, detection and denoising [24, 6]. The complementary information equips the image analysis tasks with additional capability enabling these methods to outperform those that rely on single images for these tasks [7, 18].

Cross-modal guidance-based enhancement has been applied to natural images [27, 23], where the cross-modality-guided CE methods generally perform well in preventing saturation and over-enhancement phenomena since they exploit the redundant complementary information in the corresponding better quality image [14]. A similar concept was applied to multi-modal medical image enhancement for better visualization of structures [15] and to facilitate tumor segmentation in liver CT images [16].

In this paper, a contrast enhancement method is proposed by learning from corresponding high-contrast multi-modal medical images. The multi-modal images employed in this work possess a better perceptual quality that can ameliorate the learning capability of the model. Since image enhancement is a subjective task and it is challenging to acquire the paired data for the supervised learning approaches targeted to contrast enhancement, we formulate the CE problem as an image to image translation problem. The low contrast T2-w brain MR image is transformed into an enhanced image inheriting the contrast of the corresponding T1-w image. A two-way GAN analogous to CycleGAN [32] is used for this purpose. The proposed method is inspired by the work of Chen et al. [4] where CycleGAN was used to embed the characteristics of high contrast natural images into low-contrast images under both paired and unpaired data configurations. In this work, the generator is basically a U-Net augmented with global features. The global features carry global contrast information to improve the low-contrast images specifically when acquiring high contrast paired ground truth is not feasible.

The paper is structured as follows. First, a review of related work regarding contrast enhancement and GANs is presented. Then, we elaborate on the proposed method followed by experiment results and discussion. The conclusion is drawn in the end of the paper.

Related Work

Contrast enhancement (CE) is one of the most instinctive and commonly applied solutions in medical image applications. There exist several contrast enhancement approaches, that improve the perceptual quality of the images, however, the need for controlled CE that does not over-enhance the images is a challenging problem. CE methods can be categorized as spatial or transform domain methods [3]. Among spatial domain approaches, histogram-based methods are widely researched for medical as well as natural image enhancement because of low complexity and reasonable performance [10], [2]. Transform domain methods are also widely investigated [22].

Followed by the idea of utilizing the information in a similar

image to enhance the original image [8], several cross-modality guided image enhancement approaches were proposed for natural images [27, 23]. For instance, Near-infrared (NIR) images were enhanced utilizing photographs [33]. Gradient-based histogram matching along with wavelet domain processing was performed to embed the contrast of NIR images in photos and to enhance texture information respectively. Recently, cross-modal guided enhancement has been extended and applied to medical images as well. A method using 2D histogram specification and morphological operations was employed to map the histogram of liver CT image to that of MR image [15]. In an optimization approach, 2D-HS was combined with structural similarity index metric to retain the structural information in the original image during enhancement [16].

Deep learning methods have been applied to contrast enhancement. These include Convolutional Neural Networks such as the primary work of Yan et al. [28] to adjust the contrast of photographs. [5] and another contrast enhancement approach suitable for real-time. Generative Adversarial Networks (GANs) were used by Ignatov et al. [9] to learn the mapping between phone and DSLR cameras. GANs have drawn incredible attention recently and are being applied to solve several difficult tasks including an image to image translation [32], super-resolution [29] enhancement [4] and many other problems.

All the approaches mentioned here require paired training data for network training. Since contrast enhancement is a subjective task, it is generally difficult to collect a huge amount of paired data. Moreover, different users have different preferences for contrast. To address this issue, cycle-GANs were introduced which eradicate the necessity of paired ground truth; instead, the network learns from the unpaired training data by incorporating several loss functions.

The availability of paired training data for medical image contrast enhancement is even challenging and it is difficult to acquire ground truth. However, the redundant complementary information acquired during clinical routine exams makes it possible to enhance the low contrast images using corresponding high perceptual quality multi-modal images. We exploit the capability of cycleGAN in this work to extract global contrast information from the corresponding multi-modal image to embed this information in enhancing the contrast of its corresponding low-contrast medical image.

Method

In this section, we discuss our proposed methodology. First, a general description of cycleGAN is provided, then we explain the generator architecture followed by loss functions.

As mentioned earlier, cycleGAN is particularly suited in scenarios where acquisition of paired input-ground truth data is challenging. Our proposed method discovers and learns global contrast from the ground truth images to embed this information in the generated images. The images enhanced as a result possess those characteristics while simultaneously possessing the content of the input image due to the loss functions used in cycleGAN. This kind of framework has shown drastic performance in the image to image translation domain due to its ability to learn the embedding of input data and generating output samples in the space spanned by training samples. This concept has been exploited for natural image enhancement as well, where the cycleGANs were used to learn the mapping between input and ground truth under paired supervision and unpaired supervision. Under unpaired supervision, the high contrast images, as well as HDR images with entirely different content, were used for training. Inspired by this work, we employ the corresponding high contrast multi-modal images for this purpose as ground truth that share similar objects contours as the input images.

In the proposed method, the source domain and target domain data are denoted by A and B respectively, the source domain consists of low contrast T2-weighted (T2-w) MR images and the target domain is a collection of high contrast T1-w images.

The general configuration of cycleGAN is depicted in the figure 1. Considering $a \in A$, the generator G_A converts a into b', where $b' = G_A(a) \in B$, The discriminator D_B discriminates between real samples (target domain) and generated data (fake samples). CycleGAN imposes cycle consistency loss, where G'_B accepts G_A-generated sample and applies backward mapping to transform it to source domain A. CycleGANs employ forward pass and backward pass represented as $a \xrightarrow{G_A} b' \xrightarrow{G'_B} a'' and b \xrightarrow{G_B} a' \xrightarrow{G'_A} b''$ to inspect the consistency between a and a' and b and b' respectively.

Now, we explain the design of our generator. U-Net has been used as generator in our work. Initially applied to medical image segmentation, U-Net has shown promising performance on several imaging problems. However, adapting the original U-Net as our generator cannot guarantee efficient enhancement considering the unpaired data. Therefore, global features have been added into the U-Net, the conjecture is that the global features manipulation works well in learning the global contrast of the corresponding multi-modal data samples. Figure 2 elaborates the detailed architecture of generator and discriminator.

Several losses used in the method are expressed below. The first that is Identity mapping loss I enforces the transformed image b content to be analogous to that of input a:

$$\mathrm{I} = \mathop{\mathbb{E}}_{a,b'}[MSE(a,b')] + \mathop{\mathbb{E}}_{b,a'}[\mathrm{MSE}(b,a')] \tag{1}$$

$$C = \mathop{\mathbb{E}}_{a,a''}\left[MSE\left(a,a''\right)\right] + \mathop{\mathbb{E}}_{b,b''}\left[MSE\left(b,b''\right)\right] \tag{2}$$

The adversarial losses A_D and A_G are expressed as:

$$A_D = \mathop{\mathbb{E}}_{a}[D_A(a)] - \mathop{\mathbb{E}}_{a'}\left[D_A\left(a'\right)\right] + \mathop{\mathbb{E}}_{b}[D_B(b)] - \mathop{\mathbb{E}}_{b'}\left[D_B\left(b'\right)\right] \tag{3}$$

$$A_G = \mathop{\mathbb{E}}_{a'}\left[D_A(a')\right] + \mathop{\mathbb{E}}_{y'}\left[D_B\left(b'\right)\right] \tag{4}$$

Conventionally, cycleGANs employ the same generator for G_A and G'_A since both transform the input samples in domain A to domain B (the same applies to G_B and G'_B). However, G_A accepts input from the original distribution (real samples) whereas G'_A accepts the generated (or fake) samples; both possess different distributions. Enabling the two to specifically adapt to their inputs results in higher PSNR in enhancement problems [4], therefore, individual batch normalization (iBN) layers were used for G_A and G'_A. Except for BN layers, the rest of the layers and parameters are shared between the two.

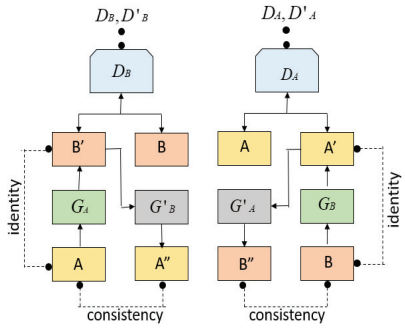


Figure 1: Network Architecture

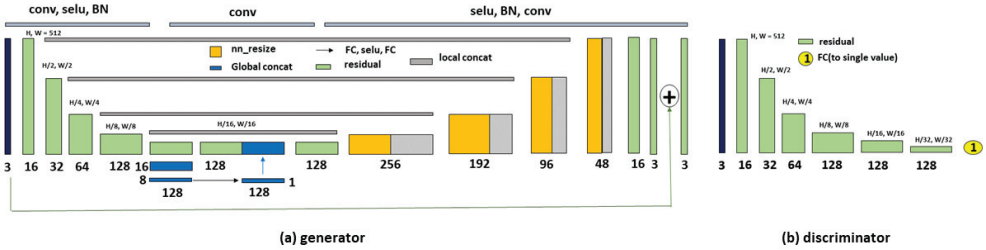


Figure 2: Network Architecture of a) generator and b) discriminator

Experiment

This section explains the dataset used for the experiment, pre-processing applied to data, and the methods selected for comparison.

Dataset

The public dataset of Hammersmith Hospital, United Kingdom accessible on the IXI database was used for analyzing the performance of the proposed method in comparison with other enhancement approaches. Total 3000 image pairs (T1-w, T2-w) were used for training, whereas 400 were used for testing. Input to our network (T2-w images) was darkened by applying morphological operations, whereas original T1-w images were used as reference or ground truth.

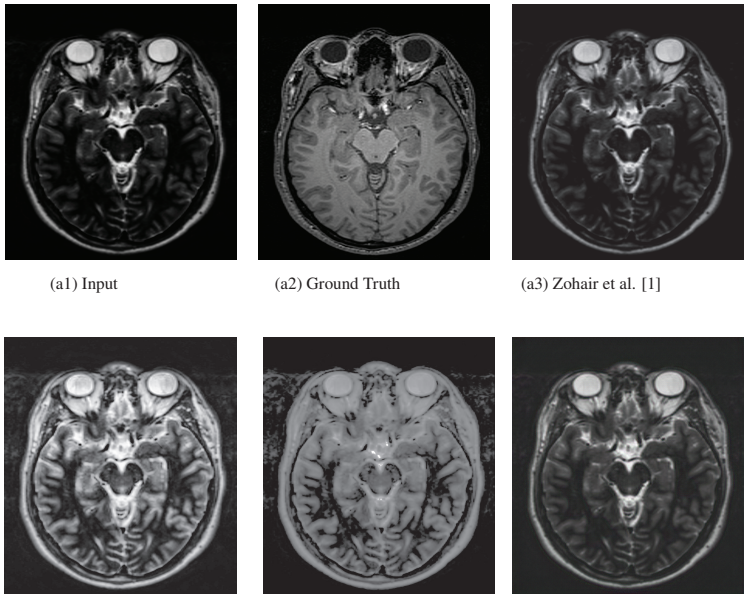
Implementation Details

The proposed method is implemented in PyTorch. The network was trained for 100 epochs with the learning rate 1e-5. Weight decay values was 0.5. All the images were 512×512. The network was trained on RTX Twin Titan with a batch size of 8.

Experiment Results

The proposed method was compared with three contrast enhancement methods. The first method, Contrast Limited Adaptive Histogram Equalization (CLAHE) is one of the well-known and widely accepted methods for CE. Cross-modality Guided Enhancement (CMGE) was proposed recently to improve the contrast of medical images using cross-modal guidance information in a 2D histogram-based approach. The third method is a modification of single-scale retinex with the inclusion of sigmoid function presented for low contrast medical images. The results of enhancement from all the methods are shown in figure 3. The input image is a low-contrast dark T2-w brain image. The method of zohair et al. [1] further darkens the image. CLAHE on the other hand improves the contrast; CMGE also improves image contrast in some regions, however qualitative analysis shows the over-enhancement phenomena in case of both the enhanced images. The proposed method improves the contrast without over-enhancing certain areas of the image. The quantitative assessment done to compare all the approaches is discussed below.

Image Quality Assessment (IQA) metrics Feature Similarity Index Metric (FSIM) [30] and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [31] are well accepted for the evaluation of medical imaging applications in addition to natural images [11, 12, 19]. FSIM is a full-reference IQA metric whereas BRISQUE is a reference-less metric. Both were used to evaluate the performance of the enhancement methods considered in this work. Table 1 presents the results of the quantitative assessment. It is important to mention that higher FSIM scores while lower BRISQUE scores imply superior contrast. Considering the quantitative results, we observe that the proposed method works best in preserving the important features in the enhanced image as shown by the highest FSIM values. Besides, it also prevents the artifacts in the enhanced images as pointed by the BRISQUE results.



(a1) Input (a2) Ground Truth (a3) Zohair et al. [1]

(a4) CLAHE [34] (a5) CMGE [15] (a6) proposed

Figure 3: Enhancement Results: Comparison with recent methods

Table 1: Quantitative Assessment

Metric	**Zohair et al. [1]**	**CLAHE [34]**	**CMGE [34]**	**proposed**
FSIM	0.812	0.714	0.71518	0.984
BRISQUE	47.132	45.172	52.582	32.838

Discussion and Conclusion

A cross-modal learning approach for contrast enhancement of medical images is proposed in this paper. The capability of cycleGAN coupled with global features in U-Net bypasses the need for paired ground truth. Instead, the complementary information and structural similarity of redundant multi-modal medical images has been exploited and effectively utilized in the learning framework. The proposed method improves reasonable contrast without introducing artifacts. The experimental results on the publicly available dataset prove that the method not only retains the features but also maintains the structure and naturalness of the original T2-w MR images as evaluated by the quality assessment metrics. This concept can be further extended to other multi-modal medical images including Computed Tomography and Positron Emission Tomography images.

Author Biography

Rabia Naseem and Faouzi Alaya Cheikh are affiliated with the Norwegian University of Science and Technology, Gjøvik, Norway. Rabia Naseem is a Ph.D. candidate, and Faouzi A. Cheikh is a Professor. Akib Jayed Islam is a master's student under the Erasmus+ Joint Master Program in Computational Colour and Spectral Imaging (COSI). Azeddine Beghdadi is a professor at University Sorbonne Paris Nord, France.

References

[1] Z. Al-Ameen and G. Sulong. A new algorithm for improving the low contrast of computed tomography images using tuned brightness controlled single-scale retinex. *Scanning*, 37 2:116–25, 2015.

[2] Zohair Al-Ameen, Ghazali Sulong, Amjad Rehman, Abdullah Al-Dhelaan, Tanzila Saba, and Mznah Al-Rodhaan. An innovative technique for contrast enhancement of computed tomography images using normalized gamma-corrected contrast-limited adaptive histogram equalization. *EURASIP Journal on Advances in Signal Processing*, 2015(1):1–12, 2015.

[3] Azeddine Beghdadi and Alain Le Negrate. Contrast enhancement technique based on local detection of edges. *Computer Vision, Graphics, and Image Processing*, 46(2):162–174, 1989.

[4] Yu-Sheng Chen, Yu-Ching Wang, Man-Hsin Kao, and Yung-Yu Chuang. Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6306–6314, 2018.

[5] Michaël Gharbi, Jiawen Chen, Jonathan T Barron, Samuel W Hasinoff, and Frédo Durand. Deep bilateral learning for real-time image enhancement. *ACM Transactions on Graphics (TOG)*, 36(4):1–12, 2017.

[6] Zhe Guo, Ning Guo, Kuang Gong, Quanzheng Li, et al. Gross tumor volume segmentation for head and neck cancer radiotherapy using deep dense multi-modality network. *Physics in Medicine & Biology*, 64(20):205015, 2019.

[7] Zhe Guo, Xiang Li, Heng Huang, Ning Guo, and Quanzheng Li. Deep learning-based image segmentation on multimodal medical imaging. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3(2):162–169, 2019.

[8] K. He, J. Sun, and X. Tang. Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35:1397–1409, 2013.

[9] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. Dslr-quality photos on mobile devices with deep convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3277–3285, 2017.

[10] Iza Sazanita Isa, Siti Noraini Sulaiman, Muzaimi Mustapha, and Noor Khairiah A Karim. Automatic contrast enhancement of brain mr images using average intensity replacement based on adaptive histogram equalization (air-ahe). *Biocybernetics and Biomedical Engineering*, 37(1):24–34, 2017.

[11] Zohaib Amjad Khan, Azeddine Beghdadi, Faouzi Alaya Cheikh, Mounir Kaaniche, and Muhammad Ali Qureshi. A multi-criteria contrast enhancement evaluation measure using wavelet decomposition. In *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6. IEEE, 2020.

[12] Zohaib Amjad Khan, Azeddine Beghdadi, Mounir Kaaniche, and Faouzi Alaya Cheikh. Residual networks based distortion classification and ranking for laparoscopic image quality assessment. In *2020 IEEE International Conference on Image Processing (ICIP)*, pages 176–180. IEEE, 2020.

[13] Amar Khellaf, Azeddine Beghdadi, and Henri Dupoisot. Entropic contrast enhancement. *IEEE transactions on medical Imaging*, 10(4):589–592, 1991.

[14] Yijun Li, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang. Joint image filtering with deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 41(8):1909–1923, 2019.

[15] R. Naseem, F. A. Cheikh, A. Beghdadi, O. J. Elle, and F. Lindseth. Cross modality guided liver image enhancement of ct using mri. In *2019 8th European Workshop on Visual Information Processing (EUVIP)*, pages 46–51, 2019.

[16] Rabia Naseem, Zohaib Amjad Khan, Nitin Satpute, Beghdadi Azeddine, Faouzi Alaya Cheikh, and Joaquin Olivares. Cross-modality guided contrast enhancement for improved liver tumor image segmentation. *IEEE Access*, In press, 2021.

[17] Maria Raquel Oliva and Sanjay Saini. Liver cancer imaging: role of ct, mri, us and pet. *Cancer imaging*, 4(Spec No A):S42, 2004.

[18] R Prieto, JM Pascual, and L Barrios. Topographic diagnosis of craniopharyngiomas: the accuracy of mri findings observed on conventional t1 and t2 images. *American Journal of Neuroradiology*, 38(11):2073–2080, 2017.

[19] Muhammad Ali Qureshi, Azeddine Beghdadi, and Mohamed Deriche. Towards the design of a consistent image contrast enhancement evaluation measure. *Signal Processing: Image Communication*, 58:212–227, 2017.

[20] N. Satpute, R. Naseem, E. Pelanis, Juan Gómez-Luna, F. A. Cheikh, Ole Jackov Elle, and Joaquín Olivares. Gpu acceleration of liver enhancement for tumor segmentation. *Computer methods and programs in biomedicine*, 184:105285, 2020.

[21] Nitin Satpute, Rabia Naseem, Rafael Palomar, Orestis Zachariadis, Juan Gómez-Luna, Faouzi Alaya Cheikh, and Joaquín Olivares. Fast parallel vessel segmentation. *Computer methods and programs in biomedicine*, 192:105430, 2020.

[22] B. Sdiri, M. Kaaniche, F. A. Cheikh, A. Beghdadi, and O. J. Elle. Efficient enhancement of stereo endoscopic images based on joint wavelet decomposition and binocular combination. *IEEE Transactions on Medical Imaging*, 38(1):33–45, 2019.

[23] Xiaoyong Shen, Chao Zhou, Li Xu, and Jiaya Jia. Mutual-structure for joint filtering. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3406–3414, 2015.

[24] Bernhard Stimpel, Christopher Syben, Franziska Schirrmacher, Philip Hoelter, Arnd Dörfler, and Andreas Maier. Multi-modal deep guided filtering for comprehensible medical image processing. *IEEE transactions on medical imaging*, 39(5):1703–1711, 2019.

[25] Shanmugapriya Survarachakan, Egidijius Pelanis, Zohaib Amjad Khan, Rahul Prasanna Kumar, Bjørn Edwin, and Frank Lindseth. Effects of enhancement on deep learning based hepatic vessel segmentation. *Electronics*, 10(10):1165, 2021.

[26] Congcong Wang, Ahmed Kedir Mohammed, Faouzi Alaya Cheikh, Azeddine Beghdadi, and Ole Jacob Elle. Multiscale deep desmoking for laparoscopic surgery. In *Medical Imaging 2019: Image Processing*, volume 10949, page 109491Y. International Society for Optics and Photonics, 2019.

[27] Qiong Yan, Xiaoyong Shen, Li Xu, Shaojie Zhuo, Xiaopeng Zhang, Liang Shen, and Jiaya Jia. Cross-field joint image restoration via scale map. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1537–1544, 2013.

[28] Zhicheng Yan, Hao Zhang, Baoyuan Wang, Sylvain Paris, and Yizhou Yu. Automatic photo adjustment using deep neural networks. *ACM Transactions on Graphics (TOG)*, 35(2):1–15, 2016.

[29] Chenyu You, Guang Li, Yi Zhang, Xiaoliu Zhang, Hongming Shan, Mengzhou Li, Shenghong Ju, Zhen Zhao, Zhuiyang Zhang, Wenxiang Cong, et al. Ct super-resolution gan constrained by the identical, residual, and cycle learning ensemble (gan-circle). *IEEE transactions on medical imaging*, 39(1):188–203, 2019.

[30] Lin Zhang, Lei Zhang, Xuanqin Mou, and David Zhang. Fsim: A feature similarity index for image quality assessment. *IEEE transactions on Image Processing*, 20(8):2378–2386, 2011.

[31] Zhicheng Zhang, Guangzhe Dai, Xiaokun Liang, Shaode Yu, Leida Li, and Yaoqin Xie. Can signal-to-noise ratio perform as a baseline indicator for medical image quality assessment. *IEEE Access*, 6:11534–11543, 2018.

[32] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[33] Shaojie Zhuo, Xiaopeng Zhang, Xiaoping Miao, and Terence Sim. Enhancing low light images using near infrared flash images. In *2010 IEEE International Conference on Image Processing*, pages 2537–2540. IEEE, 2010.

[34] Karel Zuiderveld. Contrast limited adaptive histogram equalization. *Graphics gems*, pages 474–485, 1994.

Chapter 11

Paper F: Cross-Modal Guidance assisted Hierarchical Learning based Siamese Network for MR Image Denoising

electronics

MDPI

Article

Cross-Modal Guidance Assisted Hierarchical Learning Based Siamese Network for MR Image Denoising

Rabia Naseem [1,*], Faouzi Alaya Cheikh [1], Azeddine Beghdadi [2], Khan Muhammad [3,*] and Muhammad Sajjad [1]

[1] Department of Computer Science, Norwegian University of Science and Technology, 2815 Gjøvik, Norway; faouzi.cheikh@ntnu.no (F.A.C.); muhammad.sajjad@ntnu.no (M.S.)
[2] L2TI, Institut Galilée, Université Sorbonne Paris Nord, 93430 Villetaneuse, France; beghdadi@sorbonne-paris-nord.fr
[3] Visual Analytics for Knowledge Laboratory (VIS2KNOW Lab), Department of Software, Sejong University, Seoul 143-747, Korea
* Correspondence: rabia.naseem@ntnu.no (R.N.); khan.muhammad@ieee.org (K.M.)

Abstract: Cross-modal medical imaging techniques are predominantly being used in the clinical suite. The ensemble learning methods using cross-modal medical imaging adds reliability to several medical image analysis tasks. Motivated by the performance of deep learning in several medical imaging tasks, a deep learning-based denoising method Cross-Modality Guided Denoising Network *CMGDNet* for removing Rician noise in T1-weighted (T1-w) Magnetic Resonance Images (MRI) is proposed in this paper. *CMGDNet* uses a guidance image, which is a cross-modal (T2-w) image of better perceptual quality to guide the model in denoising its noisy T1-w counterpart. This cross-modal combination allows the network to exploit complementary information existing in both images and therefore improve the learning capability of the model. The proposed framework consists of two components: Paired Hierarchical Learning (PHL) module and Cross-Modal Assisted Reconstruction (CMAR) module. PHL module uses Siamese network to extract hierarchical features from dual images, which are then combined in a densely connected manner in the CMAR module to finally reconstruct the image. The impact of using registered guidance data is investigated in removing noise as well as retaining structural similarity with the original image. Several experiments were conducted on two publicly available brain imaging datasets available on the IXI database. The quantitative assessment using Peak Signal to noise ratio (PSNR), Structural Similarity Index (SSIM), and Feature Similarity Index (FSIM) demonstrates that the proposed method exhibits 4.7% and 2.3% gain (average), respectively, in SSIM and FSIM values compared to other state-of-the-art denoising methods that do not integrate cross-modal image information in removing various levels of noise.

Keywords: cross-modal; guided; denoising; MRI; machine learning; siamese network; deep learning

Citation: Naseem, R.; Alaya Cheikh, F.; Beghdadi, A.; Muhammad K.; Sajjad, M. Cross-Modal Guidance Assisted Hierarchical Learning Based Siamese Network for MR Image Denoising. *Electronics* **2021**, *10*, 2855. https://doi.org/10.3390/electronics10222855

Academic Editor: Hyunjin Park

Received: 16 August 2021
Accepted: 11 November 2021
Published: 19 November 2021

1. Introduction

Magnetic Resonance Imaging (MRI) is preferred for the structural and functional analysis of several organs in the clinical setting thanks to its non-ionizing nature and ability to highlight structures with high contrast. In particular, MR neuroimaging is widely employed in the screening and diagnosis of brain cancers and neurodegenerative dysfunctions such as Alzheimer's disease and multiple sclerosis [1]. MRI can highlight tissue with various contrasts using different sequences of Radio-Frequency (RF) pulses. Specific pathologies are accurately analyzed and interpreted when captured using a particular RF pulse sequence. For instance, 'substantia nigra', a brain area affected due to Parkinson's disease can be visualized clearly on T2-w images compared to T1-w [2], whereas, T1-w images are preferred in the quantification of atrophy, an irreversible loss of neurons associated with multiple sclerosis [3]. However, certain pathologies possess uncertain features and assorted topography, whose existence needs to be validated by multiple modalities especially if their surgical resection is essential. A cohort study comprising 200 surgically

treated Craniopharyngiomas (CPs), an infiltrative brain tumor concluded that several key radiological variables recognized on both T1-w and T2-w MR images correctly predicted the CP topography in 86% of cases [4].

During MRI acquisition, noise is mainly introduced due to motion of charged particles in the radio frequency coils. This noise affects the reliability of diagnosis and image analysis tasks including feature extraction and segmentation [5,6]. Denoising of the images then becomes indispensable to make them suitable for further analysis. Let $X \in \mathbb{R}^{P\times Q}$ and $Y \in \mathbb{R}^{P\times Q}$ denote the ideal and observed MR images, respectively, and $N \in \mathbb{R}^{P\times Q}$ is the noise contained in the MRI signal. The noisy observation Y of X can be expressed in the case of an additive model as:

$$Y = X + N \tag{1}$$

The objective of denoising algorithms is to reduce the noise content N in Y to obtain an estimate of the original image X. The noise in MR images follows Rician distribution whose probability density function is expressed as:

$$p(N \mid X, \sigma_N) = \frac{N}{\sigma_N^2} \exp^{-\frac{\left(N^2+X^2\right)}{2\sigma_N^2}} J_0\left(\frac{XN}{\sigma_N^2}\right)\epsilon(N) \tag{2}$$

In the above equation, J_0 represents the 0^{th} order Bessel function, N is a Rician distributed random variable. $\epsilon(.)$ is the unit step Heaviside step function indicating that the pdf expression is valid for non-negative values of N. X is a non-noisy signal as stated above and σ_N is the noise variance. The Rician noise is a signal dependant noise and demonstrates gaussian distribution when Signal-to-Noise Ratio (SNR) is high and rayleigh distribution when SNR is low.

Despite the considerable amount of work devoted to image denoising during the two last decades, it is still a challenging problem particularly in the case of signal-dependant and correlated noise [7,8]. This is the case in medical imaging. Most often simplifying assumptions are made to make the denoising problem more or less tractable. This has led to a variety of denoising methods applied to various imaging modalities. Several denoising approaches have been proposed in the past that can be broadly grouped into two types: conventional methods and deep learning-based approaches. The conventional denoising methods include spatial domain methods such as bilateral filter [9], Non-Local Means filter (NLM) [10] and anisotropic filter [11] to name a few. Among these filters, the NLM filter specifically demonstrates superior performance when the image contains regions of various types of textures. Wavelet domain approaches were also widely researched for image quality enhancement [12–14]; one such approach applies thresholding on the detail coefficients. The wavelet-based denoising methods well preserve sharp edges in the images compared to spatial domain methods. Optimization-based denoising techniques including total-variation denoising [15] provide more control over preserving details in the image and the extent of noise reduction. Recently, data driven machine learning approaches, particularly deep learning methods are gaining incredible attention due to their promising performance in various areas such as biomedicine [16–19], video processing. These methods are able to mimic human cognition [20,21]. Similarly, these approaches clearly outperform the conventional approaches in the area of denoising [22–24].

Indeed, acquisition of multi-modal medical imaging data during therapeutics is becoming increasingly common [25,26]. Since these diagnostic imaging techniques are one of the largest sources of big data [27–31], their automated analysis is highly desirable to facilitate the computer aided diagnosis of several diseases [32–35]. For instance, Computed Tomography (CT) and positron emission tomography (PET) are concurrently acquired as a standard treatment protocol in oncology. Similarly, T1 and T2-w MRI provides anatomical and pathological information, respectively. The combination of this complementary information plays a significant role in therapy and surgical planning. The concept of 'weak learnability' in ensemble learning further motivates to exploit the strength of this

complimentarity. According to this concept, the learner (imaging modality here) can be incorporated into the learning system to elevate its performance, provided it can perform slightly better than random guessing [36].

With technical advancement and the availability of medical imaging techniques, using multi-modal data for the underlying computer-aided tasks is attracting several researchers. It has been exploited in segmentation, classification, super-resolution, and denoising [36–38]. For instance, in the context of lung nodule detection, CT and PET images were combined in a CNN-based approach [37]. Similarly, CT, PET, and MRI were also combined for tumor segmentation [36]. It is worth mentioning here that multi-modal information-based methods showed superior performance compared to those relying on a single modality (either CT or PET) [39].

The use of multi-modal medical imaging methods in improving segmentation and object detection motivates the researchers to employ the dual imaging in denoising as well. Few research works presented for medical image denoising [23,40] show improved performance over their single image denoising counterparts. Single image denoising approaches have an intrinsic limitation where the corrupted information in the original image is only hallucinated during the reconstruction process [41]. Consequently, these approaches over smooth certain critical structures in the image at the expense of removing noise [42]. It often leads to compromised performance of segmentation and object detection algorithms [43]. In this context, techniques that rely on cross-modal guidance offer the potential to overcome this limitation. Conventionally, cross-modal denoising methods use an image of better perceptual quality to facilitate the restoration process. Cross-modality guided medical image denoising is a relatively under-explored area; however, there exist a few approaches for natural images. One of the traditional denoising methods attempted to denoise depth maps using corresponding RGB images [44,45]. Deep learning-based cross-modal denoising approaches include [46,47]. One of these methods uses RGB-depth data pair to denoise depth images. Their proposed method consists of two CNNs; first to extract features individually from the RGB (guidance) and depth (target) images; the features are later concatenated to be fed to the third CNN, which selectively transfers the common structures in both images to generate the denoised image [46]. This work was further extended by adding a skip connection between the input image and the network prediction to enforce residual learning [47]. This modification brought significant improvement in the results by leveraging accurate details from the guidance to the target image.

A few cross-modal medical image denoising methods including [23,48,49] are found in the literature. One such work consolidated information from PET and MRI (T1 and T2 FLAIR) to denoise very low-dose PET images of the human brain [23]. The proposed method ResUNet was a residual encoder-decoder network, where residual learning was combined with U-Net. The PET, T1 and T2 FLAIR slices were stacked together and fed to the network. Using 2.5D information offers a way to discriminate structural information from noise. Compared to the ResUNet (PET without MRI), the combination of both modalities not only resulted in improved denoising performance but also improved lesion segmentation. In another similar CNN-based approach, amyloid PET images were concatenated with corresponding T1, T2, and T2 FLAIR images to learn denoising ultra-low-dose PET images using standard dose PET as ground truth. U-Net with residual learning was used in their approach. A similar idea was applied to T1 and T2 brain images [48]. The traditional guided filter was integrated with the deep learning framework, where guidance map generator takes guidance and cross-modal noisy images as input (T1 and T2 MR images). The guidance map generation component was realized using a modification of popular architecture, U-Net; where the encoding path was extended to dual branches for each modality followed by feature concatenation at the last encoding layer. The guidance filter was then incorporated as a differential layer and implemented as a linear combination of the guidance map and input image to yield the restored image. The method claimed to outperform approaches that do not include the guidance information from input im-

age directly in the restoration process and rather only rely on the network prediction as final output.

The above-mentioned approaches are not very effective since simply concatenating the images as network input as in [23] or combining features from all encoding layers [48] does not fully exploit the potential of cross-modal complementarity. It leaves a huge space to further explore the improvement of cross-modal denoising methods and advance in this direction. Therefore, there is a need for a more efficient way of manipulating and combining features. To address the denoising problem in MR images, we present a cross-modality-guided denoising approach *CMGDNet* in this paper. The proposed model is inspired by the work of Fu et al. [50], where a similar model was used to detect salient objects in RGB-Depth images. Cross-modal image denoising for brain MR images was earlier explored by Stimpel et al. [48]; however, simple feature concatenation at the last encoding layer of their proposed method does not effectively exploit the information in the non-noisy guidance image. Unlike the previous denoising approaches, *CMGDNet* extracts hierarchical features from the input and guidance image using a siamese network (mirror backbones) that are later combined in the complimentarity-aware mechanism. Although T1 and T2 images belong to different modalities; nonetheless, they capture similar structures and analogous object contours. The guidance image (T2 in our case) has better perceptual quality (noise-free), while T1 is of lower quality due to its sensitivity to acquisition noise. This scenario renders cross-modal feature learning viable in the presence of a guidance image. Our contributions in this work are listed as follows:

- A novel framework based on cross-modal guidance information is designed to denoise T1-w brain MR images. In particular, a siamese network is specifically modified to train the denoising network using both T1 and T2 MR images. By exploiting the diversity of information contained in the two modalities and in particular better perceptual quality of T2 images and the structural information contained in T1 images, the proposed approach seeks additional guidance from these images in the reconstruction process.
- Literature dictates that complementarity-aware cross-modal feature fusion is not well explored in the context of denoising, hence in this work, an effective cross-modal information fusion strategy is incorporated. The experimental results show that this fusion mechanism works well in comparison to single image denoising approaches.
- Comprehensive experiments have been conducted to analyze the performance of the proposed method on different noise levels both on registered as well as unregistered data. Moreover, the role of different loss functions is inspected to analyze their impact on denoising performance.
- In this work, two public datasets are customized keeping in view the requirement of denoising in medical image analysis. The dataset consists of both T1 and T2 MR images, meeting the requirement of learning models based on cross-modal guidance.

This paper consists of five sections. Section 1 gives an introduction to and motivations for the work followed by background and related work. The dataset and proposed methodology are elaborated on in Section 2. Experiments, comparisons with different techniques, and the results are discussed in Section 3. Section 4 summarizes the discussion of results. The conclusion and suggested future work are presented in Section 5.

2. Materials and Methods

In this section, the dataset, experimental setup and proposed methodology are explained.

2.1. Dataset

The experiments in this work are conducted on two datasets which are subset of a publicly available database IXI [51]. Both datasets are collections of T1, T2, and some other brain MR imaging modalities of healthy patients. The detailed configuration and scanning parameters can be found on the Brain IXI website: https://brain-development.org/ixi-dataset/. T1 and

T2 MRI have been used in this work. Further details of both datasets and experimental configuration are provided in the following subsections.

2.1.1. Dataset I: Hammersmith Hospital

The first experiment was done on the dataset acquired at Hammersmith Hospital, United Kingdom using Philips 3T system. 70 T1 and T2 volume pairs were randomly chosen, out of which 62 pairs were used for training and 8 for testing the denoising performance. It is pertinent to mention here that the proposed method was tested on registered as well as unregistered data and is explained in the following section. Furthermore, the role of different loss functions with both configurations was also investigated.

2.1.2. Dataset II: Guy's Hospital

The collection of T1 and T2 MRI acquired at Guy's Hospital, UK using Philips 1.5T system was also used in this work. Seventy T1 and 70 T2 volumes were randomly selected for this purpose, out of which 62 pairs were used for training and eight for testing.

2.2. Preprocessing

All the volumes (T1 and T2) were resampled to $256 \times 256 \times 150$. The proposed model is trained and tested on two types of input-guidance image combinations. In the first case, the model is trained on unregistered data, while in the second case, the corresponding T1 and T2 volumes are registered using 3D slicer, where T2 volume was moved/deformed with reference to the T1 volume that is fixed. Rigid registration with 12 degrees of freedom was used. Rician noise was added to the T1 slices. Min-max intensity normalization was applied to the data. We conducted experiments on unregistered as well as registered data at different levels of Rician noise. The detail of both configurations is elaborated in the next subsection.

2.3. Implementation Details

The proposed method was implemented using the PyTorch library and trained on NVIDIA TITAN RTX GPU with 24 GB RAM. The backbone was initialized using the pre-trained parameters of DSS [52], while other layers were randomly initialized. The network was fine-tuned through end-to-end paired learning. The learning rate and momentum values were 0.00005 and 0.99, respectively. Stochastic Gradient Descent learning was adopted and the network was trained using the loss functions described in Equations (6) and (8) for 50 epochs.

2.4. Proposed Methodology

The conventional deep learning models for image denoising are trained to learn the mapping from noisy image Y to non-noisy image X [41,53]. However, in cross-modality-guided denoising methods, the model incorporates an additional multi-modal image G to learn complementary information and facilitate the learning process.

$$\hat{X} = f(Y, G) \tag{3}$$

Therefore, in the proposed method, the model is trained to minimize the loss function $\mathscr{L}$ as:

$$f^* = \arg\min \mathscr{L}(f(Y, G), X) \tag{4}$$

The proposed framework PHL-CMAR (Paired Hierarchical Learning-Cross-Modal Assisted Reconstruction) employs CNNs for extracting features as well as combining these features efficiently in the restoration process. PHL-CMAR framework consists of two modules: PHL module and CMAR module. The PHL module is responsible for extracting the features from paired (T1 and T2) images using the Siamese network to conduct joint learning. It discovers commonalities between the dual inputs from a model-based

perspective, which is then incorporated in the model via back-propagation. The features extracted are then fed to the CMAR module, where they are combined to ultimately reconstruct the denoised image. Both components of the proposed model are explained in the subsequent sections.

2.4.1. Paired Hierarchical Learning (PHL)

The PHL module accepts two images, noisy T1 image Y and guidance image G, that is T2 as input. ResNet is used as a trunk architecture for feature extraction. For both images, the single-channel is copied three times to correspond with the RGB images which the general VGG or Resnet-like models accept. ResNet uses skip connections to address the issue of vanishing gradients and learns the residuals instead of the function [54]. Pre-trained ResNet-101 was used for feature extraction. Since ResNet's first convolution layer has a stride of 2, which gives the feature spatial size of 160×160 at the shallowest level, conv1_1 and conv1_2 layers from VGG-16 are used to obtain the full feature size of 320×320. Therefore, each hierarchical branch (1 to 6) is then connected to conv1_2 (borrowed from VGG-16), conv1, res2, res3(3), res4(22), and res5 layers of the ResNet-101, respectively. Axial slices were used where each image has dimension of $256 \times 256 \times 3$. The shared Siamese backbone extracts features from the dual images in the hierarchical side-output manner [52] and is briefly described below. Since feature extraction is accomplished using short connections, the feature set at every hierarchy contains varying dimensions. Another component Feature Pruning (FP) is therefore introduced in the PHL module [50], which ensures that the feature set from each hierarchy is of uniform size. Let us represent the feature set corresponding to guidance image and input image by F_g and F_y, respectively. Figure 1 shows the structure of PHL module. It is worth mentioning here that directly concatenating the two images has not been found as effective in detection tasks as combining features in a hierarchical way [50]. The combination of ResNet with hierarchical feature manipulation strategy allows the model to combine the complimentary information from both modalities that are later combined using densely connected FE module.

Short Connections

Generally, the feature maps at the shallow layers of CNNs are crude. As the DL network delves deeper, the successive convolution layers refine the feature maps obtained earlier. However, the deeper-level feature maps lack regularity. Using short-connection from following convolution layers to the earlier layers offers learning model a way to consolidate the information from multiple levels. Integrating features in this manner provides deep learning networks with information-rich multi-scale feature maps and therefore improves the results [52]. The idea of using short connections in the CNN to exploit multi-scale information was initially applied to edge detection [55] and later to salient object detection in RGB images [52]. This concept was further incorporated in the saliency detection framework using RGB-depth pair images [50], where the Siamese network was used for feature extraction.

Using short connections similar to [52] can be particularly beneficial in the denoising problem. A reasonable denoising algorithm should also be able to recover the corrupted information while simultaneously preserving the critical structures in the images. Considering medical images, the denoising problem becomes even more sensitive since the results would be later used in diagnosis and image analysis tasks such as segmentation. The capability of exploiting rich information at the feature extraction phase motivates us to include this strategy in our proposed model.

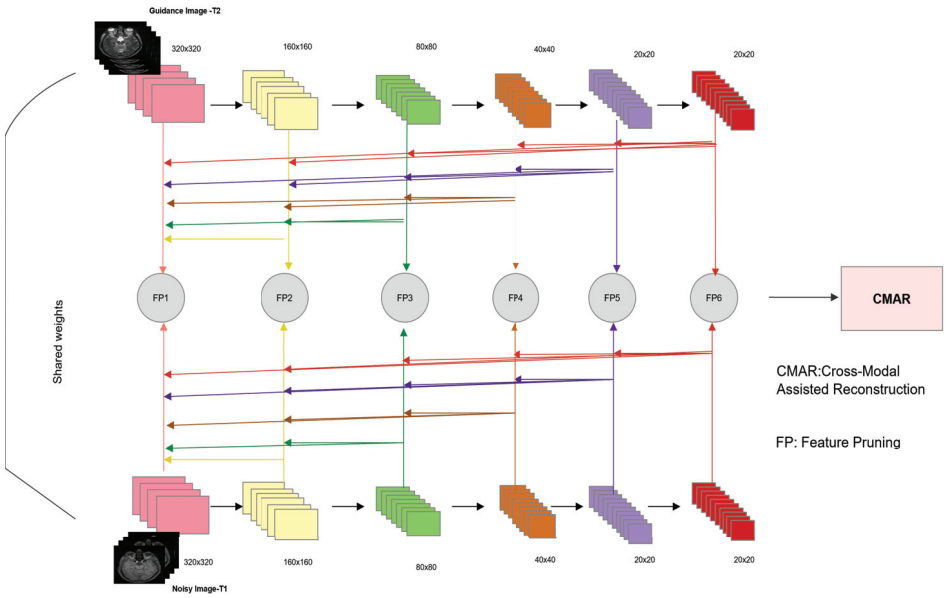


Figure 1. Paired Hierarchical Learning (PHL).

2.4.2. Cross-Modal Assisted Reconstruction (CMAR)

The CMAR module combines the hierarchical features extracted by the PHL module to perform upsampling. CMAR acts as a decoder in our proposed method and consists of two components, i.e., Cross-Modal feature Synthesis (CMS) and Feature Expansion (FE). The detail of the interactions among various elements of the CMS and FE components along with their relationship with the PHL module is depicted in Figure 2. Both components are explained in the subsections below:

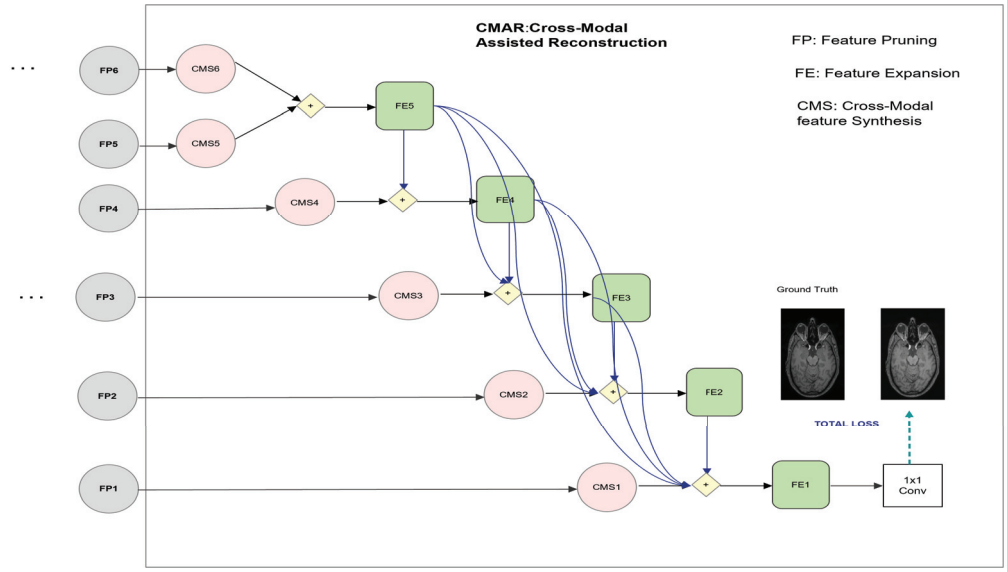


Figure 2. Cross-Modal Assisted Reconstruction (CMAR).

Cross-Modal Feature Synthesis (CMS)

Let us denote the feature set provided by the FP module as $\{F_y, F_g\}$, where F_y and F_g correspond to the features extracted from noisy image, and guidance image, respectively. CMS does feature multiplication followed by feature addition of the corresponding multi-scale features. This operation can be mathematically expressed as:

$$CMS(\{F_y, F_g\}) = F_y \oplus F_g \oplus (F_y \otimes F_g) \tag{5}$$

$\oplus$ and $\otimes$ symbols in Equation (5) represents addition and multiplication (element-wise). The addition operation exploits complementary information between both modalities in the feature space, while the multiplication operation combines common information in the cross-modal feature set. The complementarity-aware feature fusion in this manner genuinely exploits superior perceptual quality of guidance image and therefore embeds additional learning capability into the model.

Feature Expansion (FE)

The feature maps from the CMS component are passed to the FE component, which acts as a decoder in our framework and is embedded with dense connections. The dense connections enable effective information flow from each decoder block to the next. Propagating the multi-level features in a densely connected fashion has proven to improve the learning capability of the network [56]. Inception module [57] is incorporated in the FE module and is shown in Figure 3. Up-sampling in this module was done using simple bilinear interpolation. Leveraging varying filter sizes such as 1×1, 3×3, and 5×5 in Conv layers, the inception module allows the network to learn spatial patterns at several scales.

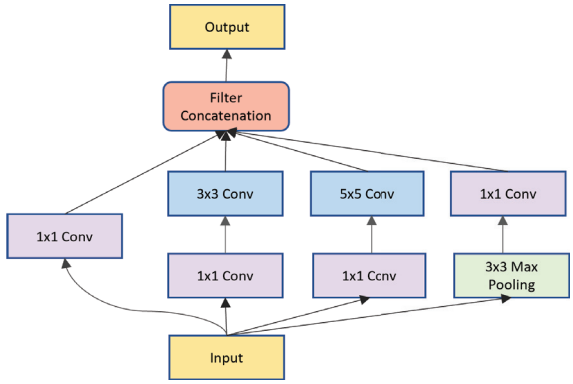


Figure 3. Structure of Inception module.

The output from the last FE module, i.e., FE1 is fed to a 1×1 convolution layer to acquire the reconstructed image in a supervised manner. The detail of the loss function is given as follows.

2.5. Loss Function

Mean square error (MSE) is a standard objective function used in several image processing problems including image super-resolution and denoising. Using MSE as a loss function allows minimizing the residual error between pixels in the ground truth and the network predicted image, which implies attaining a higher Peak Signal to Noise Ratio (PSNR). MSE loss is expressed as follows:

$$\mathscr{L}_{MSE}(\theta) = \|\hat{X}(\theta) - X\|^2 \tag{6}$$

However, it was observed that optimization using solely MSE sometimes generates blurred images. In this context, an objective function motivated by structural information can be integrated. SSIM is used to define the extent of local structural similarity between two images and can be incorporated with MSE as a loss function to address the over-smoothness phenomena associated with MSE. While the higher the SSIM value, the higher is the structural similarity in images, the objective function is therefore expressed as follows:

$$\mathscr{L}_{SSIM}(\theta) = 1 - SSIM(X, \hat{X}(\theta)) \tag{7}$$

The overall loss function of the proposed method is mathematically formulated as:

$$\mathscr{L}_{total}(\theta) = \alpha \mathscr{L}_{MSE} + \beta \mathscr{L}_{SSIM} \tag{8}$$

Equal values of α and β have been chosen for our experiment, that is 0.5.

3. Experiments and Results

The performance of the proposed method was validated by comparing it with five state-of-the-art methods including Non-local means filter (NLM) [10], Stein's unbiased risk estimate (SURE) [58], Block-matching and 3D filtering (BM3D) [59], Multi-channel Denoising convolutional neural network (MCDnCNN), referred as MCDN in the paper [24] and FFD-Net [53]. Among the methods chosen, NLM [10] is a popular denoising method that computes the weighted average of not only the local neighborhood but all pixels in the image. Wavelet-based denoising approach SURE does not rely on prior statistical modeling of wavelet coefficients [58]. Instead, it parametrizes denoising by computing parameters that minimize this MSE estimate. BM3D is a popular approach based on stacking similar 2D image patches followed by hard thresholding and Wiener filtering to denoise 3D stacks [59]. Although BM3D was originally developed for removing Gaussian noise in images; however, it has been applied to Rician noise removal as well [60]. MCDN is a 10 convolution layer network embedded with residual learning taking multi-channel input; however, we modified it to take identical slices. FFD-Net [53] is another CNN architecture that is capable of handling a variable range of noise levels in a single model.

Moreover, the denoising performance was also evaluated by using different combinations of loss functions on registered as well as unregistered data. Three metrics were used to quantitatively evaluate the performance. The first metric peak signal-to-noise ratio (PSNR) compares the root mean square error (RMSE) between the ground truth and denoised images. Another metric Structural Similarity Index (SSIM) was also included in the assessment that measures the structural affinity between denoised images and the ground truth. Feature Similarity Index (FSIM) [61] is a full reference image quality assessment (IQA) metric that is often used to evaluate the performance of denoising methods [62]. It computes feature similarity between the two images based on the low-level features including phase congruency and gradient magnitude.

In the following subsections, we describe in detail the experiments conducted on the brain MR images using the proposed method and state-of-the-art methods.

3.1. Configurations

The performance of the proposed method *CMGDNet* was evaluated on unregistered and registered data with different combinations of loss functions. Different configurations of data and loss functions tested in the proposed method are mentioned in Table 1 and briefly explained below:

3.1.1. $CMGDNet_{rs}$

Under this configuration, registration was not performed between T1-w and T2-w volumes. Using MSE as loss function, the model was trained and then tested on both datasets. The results of this configuration are referred as '$CMGDNet_{rs}$'.

3.1.2. $CMGDNet_{ss}$

The role of using an additional SSIM-based loss function was analyzed in case of unregistered data under this configuration. Therefore, both SSIM and MSE were combined here.

3.1.3. $CMGDNet_{rg}$

In this case, registration was performed between T1 and T2 volumes. Registration was done using 3D Slicer. Rigid Registration with 12 degrees of freedom was applied in all cases where T1 volume was fixed, while T2 was moved with reference to T1 in the registration process. The effect of registration can be better comprehended by visually inspecting the registered and unregistered T2 images with reference to T1 in Figure 4. It can be noticed that T1 and unregistered T2 slices are structurally similar; however, careful insight points to structural mismatches at various regions in the image. After applying registration, the structural similarity in the registered T2 image can be seen in the highlighted areas. In the following experimental sections, we further analyze the impact of registration on denoising and structural preservation in the presence of cross-modal image T2. The loss function used in this case is MSE.

(**a**) (**b**) (**c**)

Figure 4. Comparison of registered and unregistered images. (**a**) T1. (**b**) T2-Unregist. (**c**) T2-Regist.

3.1.4. $CMGDNet_{sg}$

SSIM was combined with MSE for analyzing the performance of the proposed method on the registered data in this configuration.

Next, we explain the experiments conducted to compare the performance of $CMGDNet$ with other denoising methods.

3.2. Experiment I

The first set of experiments was conducted by comparing the proposed method ('$CMGDNet_{sg}$' configuration was used in this set of experiments) with state-of-the art denoising methods. The experiments were conducted on the T1 images taken from two datasets, HH and Guy's, corrupted by Rician noise in the range 5% to 13%.

3.3. Experiment II

The second experiment was conducted to investigate the impact of registration on the denoising performance; besides, the role of using different loss functions was also evaluated. Therefore, the experiments were conducted using the four configurations $CMGDNet_{rs}$, $CMGDNet_{ss}$, $CMGDNet_{rg}$, and $CMGDNet_{sg}$, in Table 1.

Table 1. Variants of *CMGDNet*.

Configuration	Data	Loss Function
$CMGDNet_{rs}$	Unregistered	MSE
$CMGDNet_{ss}$	Unregistered	MSE + SSIM
$CMGDNet_{rg}$	Registered	MSE
$CMGDNet_{sg}$	Registered	MSE + SSIM

3.4. Experiment III

Another experiment was conducted to investigate the impact of integrating corresponding cross-modal images in the proposed framework and analyze its impact in denoising and preserving the structural information in the image. In order to do this, a noisy input image (T1) was fed to both the branches of the PHL module instead of the combination of noisy input and cross-modal (guidance) image. The model was then trained using MSE and SSIM losses on Guy's hospital dataset (contaminated with 13% noise).

4. Discussion

In this section, we summarize the discussion of our results. The results of Experiment I are shown in Figures 5–7, where the denoising performance of the proposed method is shown in comparison with state-of-the-art denoising methods. In Figure 5, the input images were contaminated using 13% noise. All the images denoised using different approaches suppress noise to some extent; however, NLM [10] removes important structural details in the image and oversmoothes the contents of the denoised image during the restoration. Wavelet-based technique SURE [58] and BM3D [59] preserve the structural details; however, they do not eradicate noise to a reasonable extent. The deep learning methods clearly show better performance compared to the traditional methods, both in removing noise and maintaining the morphology of the image. Both MCDN [24] and FFD-Net [53] effectively remove the noise. Similarly, *CMGDNet* also eradicates noise with reasonable preservation of the structural information. The enlarged ROIs are also shown in the figure for careful insight into the denoising performance of all the methods. Figure 6 shows the results of denoising applied on images contaminated with 8% noise. A similar trend can be observed in this case as well where the methods MCDN [24], FFD-Net [53], and *CMGDNet* preserve important structures in the denoised images. However, NLM [10] produces over-smoothing effects. The performance was quantitatively evaluated using PSNR, SSIM, and FSIM. BM3D [59] works better compared to NLM and SURE [58]; this claim is also supported by the higher PSNR value in Table 2. The performance of FFD-Net [53] and MCDN is very similar when quantitatively evaluated. However, *CMGDNet* performs best among all the techniques evaluated.

Table 2. PSNR(db)-SSIM values—HH Hospital Data.

Level	Noisy	NLM [10]	SURE [58]	BM3D [59]	MCDN [24]	FFD-Net [53]	$CMGNet_{sg}$
5%	25.71	28.84	28.89	29.45	36.32	36.15	**36.87**
8%	22.3	26.38	24.85	27.25	33.38	33.28	**33.74**
13%	19.8	23.224	22.821	24.671	31.116	30.81	**31.77**
Mean	22.6	26.148	25.52	27.12	33.61	33.41	**34.126**
5%	0.49	0.628	0.635	0.656	0.914	0.911	**0.951**
8%	0.413	0.565	0.532	0.628	0.834	0.822	**0.885**
13%	0.278	0.488	0.461	0.514	0.803	0.79	**0.836**
Mean	0.4	0.561	0.542	0.6	0.85	0.841	**0.89**

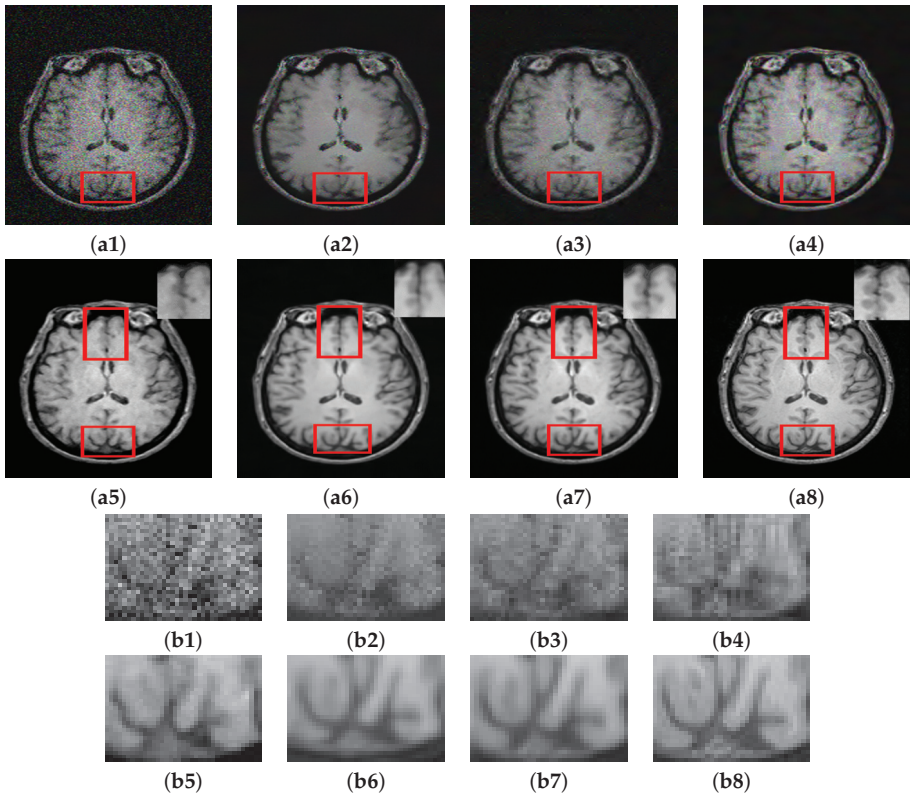


Figure 5. Comparison of proposed method with state-of-the-art denoising methods: (**a1**) Noisy. (**a2**) NLM [10]. (**a3**) SURE [58]. (**a4**) BM3D [59]. (**a5**) MCDN [24]. (**a6**) FFD-Net [53]. (**a7**) $CMGDNet_{sg}$. (**a8**) GT, Corresponding enlarged regions of the denoised images: (**b1**) Noisy. (**b2**) NLM [10]. (**b3**) SURE [58]. (**b4**) BM3D [59]. (**b5**) MCDN [24]. (**b6**) FFD-Net [53]. (**b7**) $CMGDNet_{sg}$. (**b8**) GT

The visual comparison of the performance of the proposed method with other denoising methods conducted on the Guys dataset (13% noise) is shown in Figure 7. NLM and SURE exhibit worse performance among all the methods tested. NLM eradicates significant details from the image while SURE removes minimal noise. BM3D performs slightly better than the two approaches. MCDN preserves structural information of the image; however, it leaves some noticeable noise in the image. The performance of FFD-Net visually in this case is comparable with $CMGDNet$. The quantitative assessment also validates the visual observations, which are shown in Table 3. For instance, NLM and SURE are ranked low at all the noise levels by PSNR and SSIM. BM3D performs better than both NLM and SURE. It is pertinent to mention that even the more robust conventional denoising methods such as BM3D leveraging the benefits of spatial and transform domains rely on pre-defined assumptions that do not work well under several types and levels of noise. On the other hand, deep learning approaches allow the underlying model to learn various levels of feature representations from raw to the higher level. In the context of denoising, the model thus learns the uncertain noise distributions from the input data. Consequently, these techniques can adapt to several types of noise efficiently. The deep learning methods in the proposed study perform better than the conventional methods on all the metrics. However, the cross-modal image information further enhances the network learning capability. Overall, the images denoised using all the methods still look

blurry compared to the ground truth. It is because it is not possible to recover the image contents completely that have been corrupted by noise without any loss of information. However, it can be sensed that the denoising at level 8% introduces less blur compared to the denoising applied to images containing 13% noise. Overall, the proposed method achieves the best performance among all the methods both in PSNR and SSIM. *CMGDNet* exhibits an average gain of 4.7% in SSIM value compared to the second-best MCDN (0.89 against 0.85).

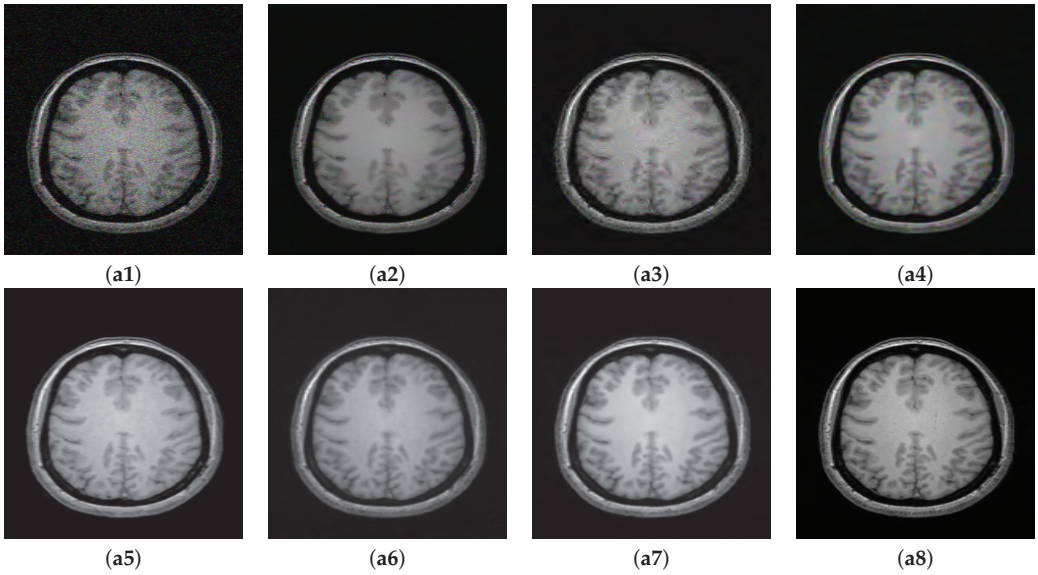


Figure 6. Denoising Results: Noise Level 8%. (**a1**) Noisy. (**a2**) NLM [10]. (**a3**) SURE [58]. (**a4**) BM3D [59]. (**a5**) MCDN [24]. (**a6**) FFD-Net [53]. (**a7**) $CMGDNet_{sg}$. (**a8**) GT.

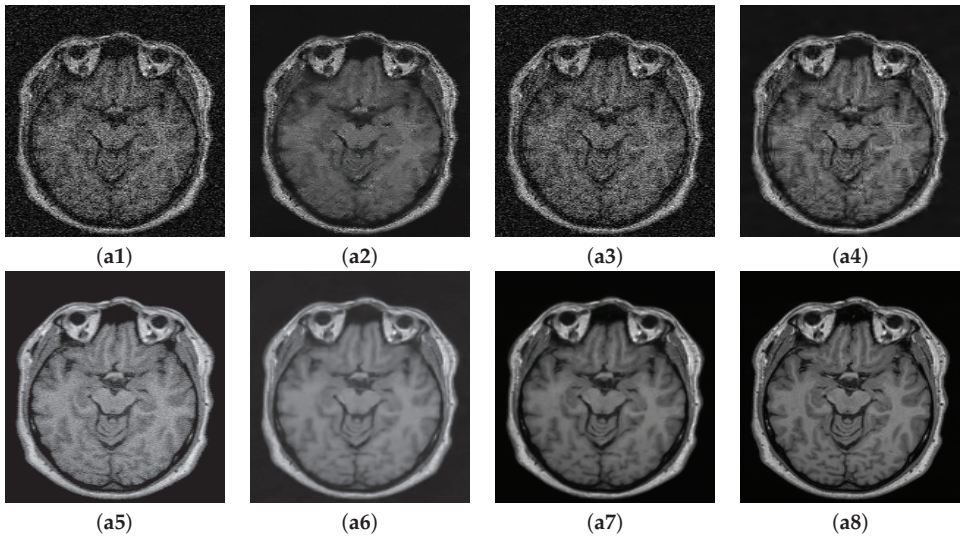


Figure 7. Comparison of proposed method with state of the art denoising methods. (**a1**) Noisy. (**a2**) NLM [10]. (**a3**) SURE [58]. (**a4**) BM3D [59]. (**a5**) MCDN [24]. (**a6**) FFD-Net. (**a7**) $CMGDNet_{sg}$. (**a8**) GT.

Table 3. PSNR(db)-SSIM values—Guys Hospital Data.

Level	Noisy	NLM [10]	SURE [58]	BM3D [59]	MCDN [24]	FFD-Net [53]	$CMGDNet_{sg}$
5%	23.45	26.96	28.07	30.41	31.67	**33.14**	33.12
8%	20.58	24.92	22.41	25.65	28.07	28.82	**30.37**
13%	17.37	22.74	21.33	23.28	25.68	27.56	**28.81**
Mean	20.46	24.87	23.93	26.44	28.47	29.84	**30.76**
5%	0.43	0.585	0.652	0.731	0.802	0.861	**0.864**
8%	0.367	0.519	0.481	0.543	0.782	0.794	**0.81**
13%	0.24	0.426	0.408	0.484	0.741	0.752	**0.786**
Mean	0.345	0.51	0.513	0.586	0.775	0.803	**0.82**

All the denoising methods included in this study bring improvement in preserving low-level features in the restored images when compared to the input noisy image as can be seen in terms of FSIM values (Tables 4 and 5). It is worth mentioning here that the FSIM scores for all the methods are very close particularly at low noise levels (5%). However, this difference is more pronounced at the higher noise levels (13%). For instance, at 13% noise, the proposed *CMGDNet* method shows the best performance on both datasets. The average gain in FSIM values in the case of *CMGDNet* (FSIM value 0.903) compared to the second-best performing method FFD-Net [53] (FSIM value 0.883) was 2.3%.

Table 4. FSIM values—HH Hospital Data.

Level	Noisy	NLM [10]	SURE [58]	BM3D [59]	MCDN [24]	FFD-Net [53]	*CMGDNet*
5%	0.868	0.9116	0.932	0.922	0.9414	0.946	**0.954**
8%	0.825	0.922	0.914	0.922	0.9385	0.936	**0.946**
13%	0.702	0.828	0.753	0.862	0.89	0.882	**0.92**
Mean	0.798	0.887	0.866	0.902	0.923	0.921	**0.94**

Table 5. FSIM values—Guys Hospital Data.

Level	Noisy	NLM [10]	SURE [58]	BM3D [59]	MCDN [24]	FFD-Net [53]	*CMGDNet*
5%	0.824	0.89	0.85	0.896	0.91	0.918	**0.925**
8%	0.77	0.884	0.877	0.888	0.881	0.89	**0.91**
13%	0.66	0.758	0.715	0.813	0.822	0.841	**0.874**
Mean	0.751	0.844	0.814	0.865	0.871	0.883	**0.903**

Another experiment (Experiment II) was conducted on the HH dataset using 13% noise. The denoising results are shown in Figure 8 along with enlarged regions for careful inspection. Table 6 shows the quantitative assessment results on different variants of data (i.e., registered and unregistered) using two different loss functions. Among the variants of the proposed method, it is observed that registration between the corresponding T1 and T2 images together with employing SSIM as loss function with MSE facilitates in improving the structural similarity between denoised image and ground truth as implied by the higher SSIM values in the case of $CMGDNet_{sg}$ compared to its corresponding variants $CMGDNet_{rs}$ and $CMGDNet_{ss}$; however, noticeable improvement in PSNR values was not observed under this configuration.

Table 6. PSNR-SSIM values—HH Hospital Data (Variants of *CMGDNet*).

Metric	Noisy	$CMGDNet_{rs}$	$CMGDNet_{ss}$	$CMGDNet_{rg}$	$CMGNet_{sg}$
PSNR	19.8	31.12	30.86	31.62	**31.77**
SSIM	0.278	0.787	0.764	0.811	**0.836**

Table 7. PSNR-SSIM values—Guys Hospital Data (Variants of *CMGDNet*).

Metric	Noisy	T1-T1	$CMGDNet_{rs}$	$CMGDNet_{ss}$	$CMGDNet_{rg}$	$CMGDNet_{sg}$
PSNR	17.37	26.52	28.44	28.21	28.63	**28.81**
SSIM	0.24	0.7	0.73	0.72	0.75	**0.786**

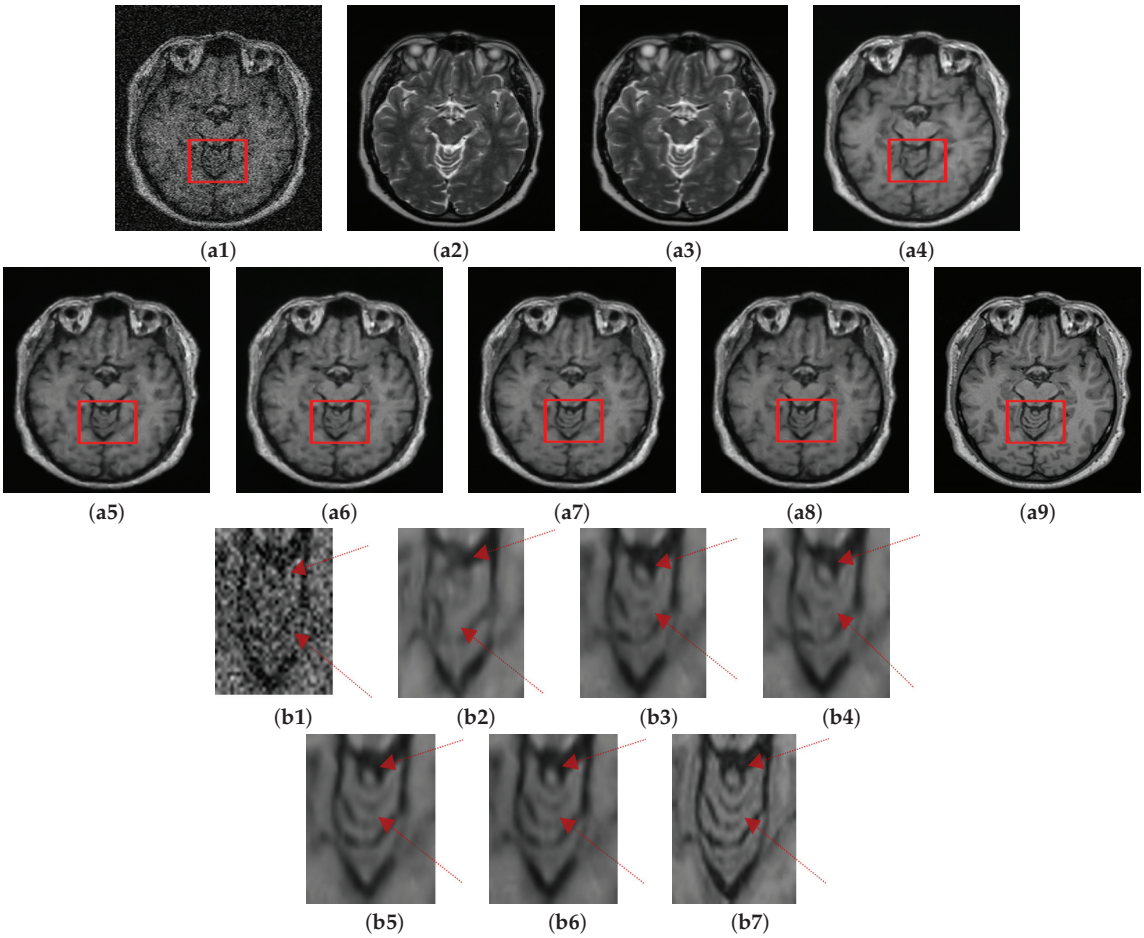


Figure 8. Role of using cross-modal guidance information in denoising: Results of *CMGDNet* method (**a1**) Noisy. (**a2**) Unreg. T2. (**a3**) Reg. T2. (**a4**) T1-T1. (**a5**) $CMGDNet_{rs}$. (**a6**) $CMGDNet_{ss}$. (**a7**) $CMGDNet_{rg}$. (**a8**) $CMGDNet_{sg}$. (**a9**) GT, Corresponding enlarged ROI (**b1**) Noisy. (**b2**) T1-T1. (**b3**) $CMGDNet_{rs}$. (**b4**) $CMGDNet_{ss}$. (**b5**) $CMGDNet_{rg}$. (**b6**) $CMGDNet_{sg}$. (**b7**) GT.

Role of Cross-Modal Guidance Information

To better understand the motivation of using cross-modal guidance information, the guidance image was bypassed and a noisy T1 image was fed to both branches of the PHL module as explained in Section 3.4. The results of this setup and its comparison with other variants of the proposed method are shown in Figure 9. Visually, the denoised images are similar on the whole; however, the enlarged ROI shows slight structural differences among the results. The model trained using identical noisy images fed to both branches (without guidance image) fails to recover various structures of the input image. Both $CMGDNet_{rs}$ and $CMGDNet_{ss}$ yield better results compared to the T1-T1 configuration; however, they also lack in recovering some structural information. $CMGDNet_{rg}$ shows better performance compared to the three variants in terms of retaining structural similarity with the ground truth. Incorporating SSIM in the registered configuration, that is $CMGDNet_{sg}$ configuration performs best. It not only retains structural similarity to a considerable extent, moreover, it also gives sharp edges compared to all the other variants. The PSNR-SSIM values for all the configurations tested are shown in Table 7.

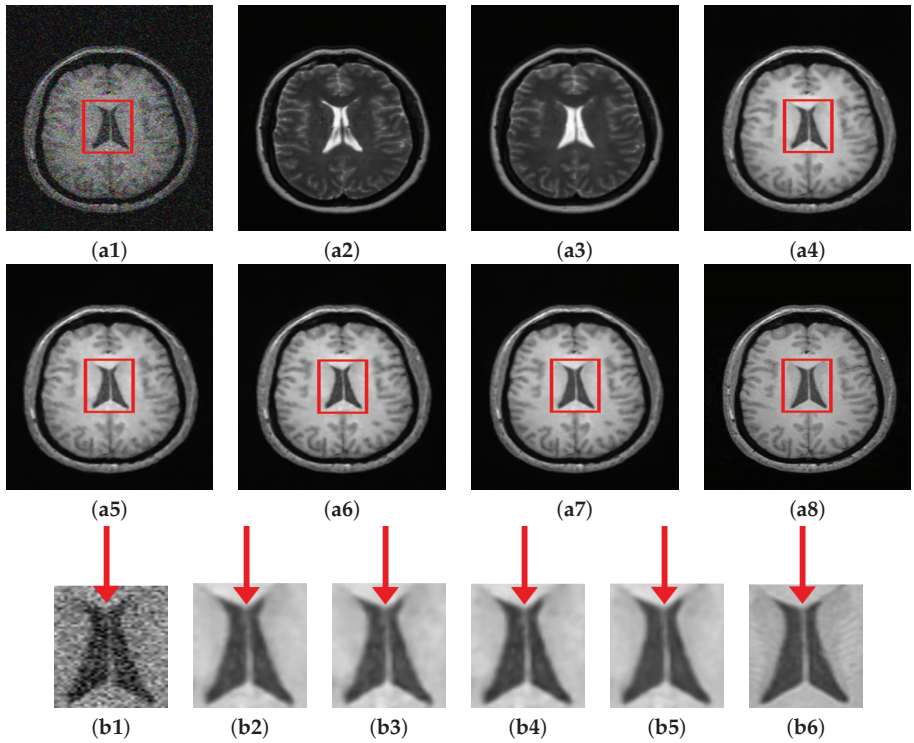


Figure 9. Comparison of proposed method (different configurations). (**a1**) Noisy. (**a2**) Unreg. T2. (**a3**) Reg. T2. (**a4**) $CMGDNet_{rs}$. (**a5**) $CMGDNet_{ss}$. (**a6**) $CMGDNet_{rg}$. (**a7**) $CMGDNet_{sg}$. (**a8**) GT, Corresponding enlarged RoI: (**b1**) Noisy. (**b2**) $CMGDNet_{rs}$. (**b3**) $CMGDNet_{ss}$. (**b4**) $CMGDNet_{rg}$. (**b5**) $CMGDNet_{sg}$. (**b6**) GT.

5. Conclusions

In this paper, a deep cross-modal guided denoising approach *CMGDNet* was presented for brain MR images, where the complementary information from the cross-modal image was exploited to embed the model with additional learning capability. Hierarchical feature manipulation combined with densely connected upsampling was particularly used to harness the additional guidance information effectively in image restoration process. Our quantitative and qualitative experimental analysis shows that the cross-modal denois-

ing shows superior results compared to single image denoising approaches. The capability of combining cross-modal image features in a systematic way, rather than simple concatenation proved to be influential in denoising. Furthermore, the experiments show that although the method works well on unregistered data; however, using registered data aids in recovering the structural information of the image. The proposed denoising approach can be used as an effective preprocessing step in various image analysis tasks.

In the future, it would be interesting to extend the research work to other organs such as the liver, lungs, and other multi-modal medical imaging modalities.

Author Contributions: Conceptualization, R.N., F.A.C. and A.B.; methodology, R.N., F.A.C. and M.S.; software, R.N.; validation, R.N. and A.B.; formal analysis, R.N. and M.S.; writing—original draft preparation, R.N., F.A.C. and A.B.; writing—review and editing, R.N., K.M. and F.A.C.; supervision, F.A.C. and A.B.; project administration and funding acquisition F.A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by H2020-MSCA-ITN Marie Skłodowska-Curie Actions, Innovative Training Networks (ITN)-H2020 MSCA ITN 2016 GA EU project number 722068 High Performance Soft Tissue Navigation (HiPerNav).

Data Availability Statement: Data available in a publicly accessible repository. The data used in this study is openly available in IXI database at: https://brain-development.org/ixi-dataset/.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Struyfs, H.; Sima, D.M.; Wittens, M.; Ribbens, A.; de Barros, N.P.; Vân Phan, T.; Meyer, M.I.F.; Claes, L.; Niemantsverdriet, E.; Engelborghs, S.; et al. Automated MRI volumetry as a diagnostic tool for Alzheimer's disease: Validation of icobrain dm. *Neuroimage Clin.* **2020**, *26*, 102243. [CrossRef]
2. Agosta, F.; Galantucci, S.; Filippi, M. Advanced magnetic resonance imaging of neurodegenerative diseases. *Neurol. Sci.* **2017**, *38*, 41–51. [CrossRef]
3. Rocca, M.A.; Battaglini, M.; Benedict, R.H.; De Stefano, N.; Geurts, J.J.; Henry, R.G.; Horsfield, M.A.; Jenkinson, M.; Pagani, E.; Filippi, M. Brain MRI atrophy quantification in MS: From methods to clinical application. *Neurology* **2017**, *88*, 403–413. [CrossRef]
4. Prieto, R.; Pascual, J.; Barrios, L. Topographic diagnosis of craniopharyngiomas: The accuracy of MRI findings observed on conventional T1 and T2 images. *Am. J. Neuroradiol.* **2017**, *38*, 2073–2080. [CrossRef]
5. Satpute, N.; Naseem, R.; Palomar, R.; Zachariadis, O.; Gómez-Luna, J.; Cheikh, F.A.; Olivares, J. Fast parallel vessel segmentation. *Comput. Methods Programs Biomed.* **2020**, *192*, 105430. [CrossRef]
6. Survarachakan, S.; Pelanis, E.; Khan, Z.A.; Kumar, R.P.; Edwin, B.; Lindseth, F. Effects of Enhancement on Deep Learning Based Hepatic Vessel Segmentation. *Electronics* **2021**, *10*, 1165. [CrossRef]
7. Gudbjartsson, H.; Patz, S. The Rician distribution of noisy MRI data. *Magn. Reson. Med.* **1995**, *34*, 910–914. [CrossRef] [PubMed]
8. Sagheer; Sameera, V.; George, S.N. A review on medical image denoising algorithms. *Biomed. Signal Process. Control* **2020**, *61*, 102036. [CrossRef]
9. Tomasi, C.; Manduchi, R. Bilateral filtering for gray and color images. In Proceedings of the Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), Bombay, India, 7 January 1998; pp. 839–846.
10. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; Volume 2, pp. 60–65.
11. Perona, P.; Malik, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 629–639. [CrossRef]
12. Weaver, J.B.; Xu, Y.; Healy, D., Jr.; Cromwell, L. Filtering noise from images with wavelet transforms. *Magn. Reson. Med.* **1991**, *21*, 288–295. [CrossRef]
13. Souidene, W.; Beghdadi, A.; Abed-Meraim, K. Image denoising in the transformed domain using non local neighborhoods. In Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing Processing, Toulouse, France, 14–19 May 2006; Volume 2, p. II.
14. Sdiri, B.; Kaaniche, M.; Cheikh, F.A.; Beghdadi, A.; Elle, O.J. Efficient enhancement of stereo endoscopic images based on joint wavelet decomposition and binocular combination. *IEEE Trans. Med. Imaging* **2018**, *38*, 33–45. [CrossRef] [PubMed]
15. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **1992**, *60*, 259–268. [CrossRef]
16. Yang, S.; Wang, J.; Hao, X.; Li, H.; Wei, X.; Deng, B.; Loparo, K.A. BiCoSS: Toward large-scale cognition brain with multigranular neuromorphic architecture. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**. [CrossRef]

17. Yang, S.; Deng, B.; Wang, J.; Li, H.; Lu, M.; Che, Y.; Wei, X.; Loparo, K.A. Scalable digital neuromorphic architecture for large-scale biophysically meaningful neural network with multi-compartment neurons. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 148–162. [CrossRef] [PubMed]
18. Yang, S.; Wang, J.; Deng, B.; Azghadi, M.R.; Linares-Barranco, B. Neuromorphic Context-Dependent Learning Framework With Fault-Tolerant Spike Routing. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**. [CrossRef]
19. Yang, S.; Wang, J.; Zhang, N.; Deng, B.; Pang, Y.; Azghadi, M.R. CerebelluMorphic: Large-scale neuromorphic model and architecture for supervised motor learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**. [CrossRef] [PubMed]
20. Bolkar, S.; Wang, C.; Cheikh, F.A.; Yildirim, S. Deep smoke removal from minimally invasive surgery videos. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 3403–3407.
21. Wang, C.; Mohammed, A.K.; Cheikh, F.A.; Beghdadi, A.; Elle, O.J. Multiscale deep desmoking for laparoscopic surgery. In Proceedings of the Medical Imaging 2019: Image Processing, International Society for Optics and Photonics, San Diego, CA, USA, 16–21 February 2019; Volume 10949, p. 109491Y.
22. Khan, S.; Sajjad, M.; Hussain, T.; Ullah, A.; Imran, A.S. A Review on Traditional Machine Learning and Deep Learning Models for WBCs Classification in Blood Smear Images. *IEEE Access* **2020**. [CrossRef]
23. Xu, J.; Gong, E.; Ouyang, J.; Pauly, J.; Zaharchuk, G. Ultra-low-dose 18F-FDG brain PET/MR denoising using deep learning and multi-contrast information. In Proceedings of the Medical Imaging 2020: Image Processing, International Society for Optics and Photonics, Houston, TX, USA, 17–20 February 2020; Volume 11313, p. 113131P.
24. Jiang, D.; Dou, W.; Vosters, L.; Xu, X.; Sun, Y.; Tan, T. Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network. *Jpn. J. Radiol.* **2018**, *36*, 566–574. [CrossRef]
25. Naseem, R.; Cheikh, F.A.; Beghdadi, A.; Elle, O.J.; Lindseth, F. Cross modality guided liver image enhancement of CT using MRI. In Proceedings of the 2019 8th European Workshop on Visual Information Processing (EUVIP), Rome, Italy, 28–31 October 2019; pp. 46–51.
26. Naseem, R.; Khan, Z.A.; Satpute, N.; Azeddine, B.; Cheikh, F.A.; Olivares, J. Cross-modality guided contrast enhancement for improved liver tumor image segmentation. *IEEE Access* **2021**, in press. [CrossRef]
27. Tahmassebi, A.; Ehtemami, A.; Mohebali, B.; Gandomi, A.H.; Pinker, K.; Meyer-Baese, A. Big data analytics in medical imaging using deep learning. In Proceedings of the Big Data: Learning, Analytics, and Applications, Baltimore, MD, USA, 13 May 2019; Volume 10989, p. 109890E.
28. Elhoseny, M.; Abdelaziz, A.; Salama, A.S.; Riad, A.M.; Muhammad, K.; Sangaiah, A.K. A hybrid model of internet of things and cloud computing to manage big data in health services applications. *Future Gener. Comput. Syst.* **2018**, *86*, 1383–1394. [CrossRef]
29. Tahmassebi, A.; Gandomi, A.H.; McCann, I.; Schulte, M.H.; Goudriaan, A.E.; Meyer-Baese, A. Deep learning in medical imaging: fMRI big data analysis via convolutional neural networks. In Proceedings of the Practice and Experience on Advanced Research Computing, Pittsburgh, PA, USA, 22–26 July 2018; pp. 1–4.
30. Yang, S.; Gao, T.; Wang, J.; Deng, B.; Lansdell, B.; Linares-Barranco, B. Efficient spike-driven learning with dendritic event-based processing. *Front. Neurosci.* **2021**, *15*, 97. [CrossRef]
31. Yang, S.; Wei, X.; Deng, B.; Liu, C.; Li, H.; Wang, J. Efficient digital implementation of a conductance-based globus pallidus neuron and the dynamics analysis. *Phys. A Stat. Mech. Its Appl.* **2018**, *494*, 484–502. [CrossRef]
32. Kumar, A.; Ramachandran, M.; Gandomi, A.H.; Patan, R.; Lukasik, S.; Soundarapandian, R.K. A deep neural network based classifier for brain tumor diagnosis. *Appl. Soft Comput.* **2019**, *82*, 105528. [CrossRef]
33. Sajjad, M.; Khan, S.; Muhammad, K.; Wu, W.; Ullah, A.; Baik, S.W. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *J. Comput. Sci.* **2019**, *30*, 174–182. [CrossRef]
34. Wang, C.; Cheikh, F.A.; Beghdadi, A.; Elle, O.J. Adaptive context encoding module for semantic segmentation. *Electron. Imaging* **2020**, *2020*, 27-1–27-7. [CrossRef]
35. Mohammed, A.; Wang, C.; Zhao, M.; Ullah, M.; Naseem, R.; Wang, H.; Pedersen, M.; Cheikh, F.A. Weakly-Supervised network for detection of COVID-19 in chest CT scans. *IEEE Access* **2020**, *8*, 155987–156000. [CrossRef]
36. Guo, Z.; Li, X.; Huang, H.; Guo, N.; Li, Q. Deep learning-based image segmentation on multimodal medical imaging. *IEEE Trans. Radiat. Plasma Med. Sci.* **2019**, *3*, 162–169. [CrossRef]
37. Teramoto, A.; Fujita, H.; Yamamuro, O.; Tamaki, T. Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique. *Med. Phys.* **2016**, *43*, 2821–2827. [CrossRef]
38. Guo, Z.; Guo, N.; Gong, K.; Li, Q. Gross tumor volume segmentation for head and neck cancer radiotherapy using deep dense multi-modality network. *Phys. Med. Biol.* **2019**, *64*, 205015. [CrossRef]
39. Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 424–432.
40. Kang, S.K.; Yie, S.Y.; Lee, J.S. Noise2Noise Improved by Trainable Wavelet Coefficients for PET Denoising. *Electronics* **2021**, *10*, 1529. [CrossRef]
41. Wang, Y.; Song, X.; Gong, G.; Li, N. A Multi-Scale Feature Extraction-Based Normalized Attention Neural Network for Image Denoising. *Electronics* **2021**, *10*, 319. [CrossRef]
42. Kang, E.; Min, J.; Ye, J.C. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med. Phys.* **2017**, *44*, e360–e375. [CrossRef]

43. Satpute, N.; Naseem, R.; Pelanis, E.; Gómez-Luna, J.; Cheikh, F.A.; Elle, O.J.; Olivares, J. GPU acceleration of liver enhancement for tumor segmentation. *Comput. Methods Programs Biomed.* **2020**, *184*, 105285. [CrossRef]
44. Yan, Q.; Shen, X.; Xu, L.; Zhuo, S.; Zhang, X.; Shen, L.; Jia, J. Cross-field joint image restoration via scale map. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 1537–1544.
45. Shen, X.; Zhou, C.; Xu, L.; Jia, J. Mutual-structure for joint filtering. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 3406–3414.
46. Li, Y.; Huang, J.B.; Ahuja, N.; Yang, M.H. Deep joint image filtering. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 154–169.
47. Li, Y.; Huang, J.B.; Ahuja, N.; Yang, M.H. Joint image filtering with deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 1909–1923. [CrossRef] [PubMed]
48. Stimpel, B.; Syben, C.; Schirrmacher, F.; Hoelter, P.; Dörfler, A.; Maier, A. Multi-Modal Deep Guided Filtering for Comprehensible Medical Image Processing. *IEEE Trans. Med. Imaging* **2019**, *39*, 1703–1711. [CrossRef] [PubMed]
49. Chen, K.T.; Toueg, T.N.; Koran, M.E.I.; Davidzon, G.; Zeineh, M.; Holley, D.; Gandhi, H.; Halbert, K.; Boumis, A.; Kennedy, G.; et al. True ultra-low-dose amyloid PET/MRI enhanced with deep learning for clinical interpretation. *Eur. J. Nucl. Med. Mol. Imaging* **2021**, *48*, 2416–2425. [CrossRef] [PubMed]
50. Fu, K.; Fan, D.P.; Ji, G.P.; Zhao, Q. Jl-dcf: Joint learning and densely-cooperative fusion framework for rgb-d salient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 3052–3062.
51. Brain IXI. Brain IXI Database. Available online: https://brain-development.org/team/ (accessed on 8 April 2021).
52. Hou, Q.; Cheng, M.M.; Hu, X.; Borji, A.; Tu, Z.; Torr, P.H. Deeply supervised salient object detection with short connections. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3203–3212.
53. Zhang, K.; Zuo, W.; Zhang, L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Trans. Image Process.* **2018**, *27*, 4608–4622. [CrossRef]
54. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
55. Xie, S.; Tu, Z. Holistically-nested edge detection. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1395–1403.
56. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
57. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
58. Luisier, F.; Blu, T.; Unser, M. A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding. *IEEE Trans. Image Process.* **2007**, *16*, 593–606. [CrossRef]
59. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [CrossRef]
60. Hanchate, V.; Joshi, K. MRI denoising using BM3D equipped with noise invalidation denoising technique and VST for improved contrast. *SN Appl. Sci.* **2020**, *2*, 1–8. [CrossRef]
61. Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A feature similarity index for image quality assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [CrossRef] [PubMed]
62. Kaur, R.; Juneja, M.; Mandal, A.K. A comprehensive review of denoising techniques for abdominal CT images. *Multimed. Tools Appl.* **2018**, *77*, 22735–22770. [CrossRef]

ISBN 978-82-326-6129-9 (printed ver.)
ISBN 978-82-326-5251-8 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology