Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

Pauline Hardeberg Zimmermann

Documenting skin bruises with smartphone imaging

Master's thesis in Electronic Systems Design and Innovation
Supervisor: Lise Lyngsnes Randeberg
June 2021

NTNU
Norwegian University of
Science and Technology

Pauline Hardeberg Zimmermann

Documenting skin bruises with smartphone imaging

Master's thesis in Electronic Systems Design and Innovation
Supervisor: Lise Lyngsnes Randeberg
June 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

NTNU

MASTER'S THESIS

Documenting skin bruises with smartphone imaging

Author:
Pauline Hardeberg ZIMMERMANN

Supervisor:
Lise Lyngsnes RANDEBERG

Department of Electronic Systems
Faculty of Information Technology and Electrical Engineering
Norwegian University of Science and Technology

June 30, 2021

Abstract

The aim of this thesis is to develop and test a standardized method of documenting skin bruises, for the purposes of providing better forensic evidence in court cases related to physical violence and aiding medical personnel in their decisions about suspected abuse. Skin bruises are the most common injury caused by physical violence, which is why they are important indicators and evidence of violence. Objective and accurate documentation and analysis of skin bruises is critical to recognize abuse in patients and to correctly determine injury mechanisms in court cases related to physical violence. However, there is no standardized method for documenting and analyzing skin bruises. Bruise images that are admitted into evidence in court cases are often acquired with smartphones. These images are often of low quality, which makes it difficult to draw conclusions based on them. To address this, this project aims to develop an objective method for documenting skin bruises using smartphone imaging, which includes estimation of bruise size, shape and color.

Participants in the study document their bruises using this standardized method. They will photograph their bruises following a set of instructions, using their own smartphone cameras. The color, shape and size of the bruises will be estimated using image processing techniques. The precision and accuracy of the feature extraction, as well as the quality of the images and the acquisition method, will be evaluated.

176 images of 26 bruises were acquired. The bruises in the data set are diverse in terms of size, shape and color. The performance of the feature extraction was evaluated with respect to precision, accuracy and robustness to changing illumination. Four image segmentation methods were evaluated based on execution time and the performance of feature extraction. Of the four methods, k-means clustering performed the best. The performance of the segmentation methods was found to depend on the type of bruise and image quality. The computational color constancy method white patch showed good results in correcting the color of the images. A ruler detection method was proposed, achieving low error.

Sammendrag

Målet med denne oppgaven er å utvikle og teste en standardisert metode for dokumentasjon av blåmerker. En slik metode kan bidra til å forbedre rettsmedisinsk bevis i rettsaker som omhandler vold og kan bistå helsepersonell i beslutninger om mistenkt mishandling. Blåmerker er den vanligste skaden forårsaket av fysisk vold, så dermed er de viktige indikatorer og bevis på vold. Objektiv og nøyaktig dokumentasjon og analyse av blåmerker er kritisk for å kunne gjenkjenne mishandling hos pasienter og for å kunne bestemme handlingsforløpet i rettsaker som omhandler fysisk vold. De bildene av blåmerker som blir brukt som rettsmedisinsk bevis i retten er ofte tatt med mobiltelefoner. Disse bildene er ofte av lav kvalitet, noe som gjør det vanskelig å bruke de til å trekke konklusjoner. Derfor har dette prosjektet som mål å utvikle en standardisert metode for dokumentasjon av blåmerker ved bruk av mobilfotografi som inkluderer estimering av blåmerkets størrelse, form og farge.

Deltakere i denne studien skal dokumentere blåmerkene sine med denne standardiserte metoden. De skal bruke egne mobilkameraer til å fotografere blåmerkene sine etter et sett med instruksjoner. Deretter, skal blåmerkenes farge, størrelse og form estimeres ved bruk av bildebehandling. Presisjonen og nøyaktigheten til estimeringen, i tillegg til kvaliteten til bildene og fotograferingsmetoden skal evalueres.

176 bilder av 26 blåmerker ble samlet inn. Blåmerkene i datasettet har et mangfold av størrelser, former og farger. Estimeringen av disse kvalitetene ble evaluert med hensyn til presisjon, nøyaktighet og robusthet til lysforhold. Fire bildesegmenteringsmetoder ble evaluert basert på kjøretid og ytelsen til estimeringingen av størrelse, form og farge. K-means clustering hadde best resultater av disse fire metodene. Ytelsen til segmenteringsmetodene avhenger av typen blåmerke og kvaliteten til bildene. Fargekonstansmetoden *white patch* ga gode resultater i å korrigere fargen i bildene. En linjalgjenkjenningsmetode ble utviklet og testet, og den oppnådde høy nøyaktighet.

Acknowledgements

I wish to thank Lise Lyngsnes Randeberg (my Master thesis supervisor and PhD co-supervisor) and Sony George (my PhD supervisor) for many fruitful discussions and helpful feedback.

I wish to thank all the study participants who provided data for this research. The Results chapter would be empty without you.

I also wish to thank my husband, Janik Hardeberg Zimmermann, for proof-reading this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introducion

1.1 Motivation

Skin bruises are the most common injury caused by physical child abuse [31, 93], intimate partner violence [91] and sexual assault [12, 89], which is why they serve as important indicators and evidence of abuse. Skin bruises can be the only visible indicator of more serious internal injuries. [118] They may also be a precursor to fatal or near-fatal injuries in children. [31] Medical personnel in Norway are required by law [81] to notify Child Protective Services if they suspect that bruises on their child patients are caused by abuse.

It is important that bruises are interpreted in an objective and accurate way, since the bruise interpretation may have significant medical and legal consequences. [133] Failing to recognize bruises as an indicator of physical abuse can be fatal. In 2005, the three-year-old child Karly Sheehan was taken to the doctor due to suspicious bruises. The doctor mistakenly concluded that the bruises were self-inflicted. Sheehan later died as a result of child abuse. [14] As a result of this case, the state of Oregon introduced *Karly's Law* [98] that introduced strict regulations to the investigation of suspected child abuse. Sometimes bruises on children are mistakenly classified as abusive, causing misplaced suspicion of child abuse. Parents in Michigan lost custody of their 6-week-old son after their doctor diagnosed the bruises on the boy's body as being a result of child abuse. The bruises were later found to be caused by the straps on a swing. [56] A British couple had their son removed from their care for more than a year due to bruises on the child's body. The son was not returned to their care until after he was examined by a geneticist, who diagnosed him with a syndrome that causes people to bruise easily. [126]

Accurate bruise documentation is important. [52, 105, 137] Melville et al. [90] found that the quality of bruise images has a significant effect on the ability of medical professionals to interpret the bruises. Bruise images may serve as forensic evidence in criminal court. Documentation of injuries, such as bruises, is associated with higher rates of conviction in sexual violence cases. [69] In State v. Hovig (2009) [6], a child abuse court case, photographs taken of bruises on the victim were considered persuasive evidence and led to a guilty verdict. In many cases, bruises are documented by the victims themselves, using smartphone cameras. The quality of these images is often poor. [34] In some cases, the bruise photographs admitted into evidence are of such low quality that it becomes difficult to draw any conclusions based on them. In another child abuse court case, Sweaney v. Ada County (1997) [5], the bruise photographs were unusable since they "did not accurately portray the bruise".

There is a need for an accurate and standardized documentation method of skin bruises. Smartphone camera quality and smartphone ownership have both increased

in the last years. [116] [103] If the skin bruise documentation method uses smartphone imaging, it will be an accessible method and potentially have a higher societal impact. Smartphone technology is becoming a valid alternative to traditional medical laboratories for gathering medical data, introducing significantly lower costs and higher availability. [80] The high availability, low cost and high connectivity of smartphone technology makes it a promising way of democratizing and decentralizing medical research and care, especially in developing nations and rural areas without access to medical laboratories. [54] Since bruises are already often documented by the victims themselves, creating a method of self-documentation of bruises with smartphone imaging can be beneficial.

1.2 Objectives

The main goal of this thesis is to answer the central research question:

Is self-documentation with smartphone imaging a suitable method for documenting skin bruises?

In order to answer this question, one must first determine was is required of skin bruise documentation in medical and legal settings and then test whether smartphone imaging is sufficient to achieve these requirements. The central research question will be answered by achieving the following objectives:

1. Create a data set of bruise images acquired by self-documentation with smartphone imaging.
2. Estimate the relevant bruise features (size, shape and color)
3. Evaluate the performance of the feature extraction
4. Evaluate the bruise self-documentation method

1.3 Outline

Chapter 2 covers the theoretical background about skin bruises and image processing that is necessary to follow this thesis. Chapter 3 covers the methodology of this project. The acquisition of the bruise image data set is covered in Section 3.1, the pre-processing is covered in Section 3.2 and Section 3.3, the bruise feature extraction is covered in Sections 3.4 - 3.6, and finally the evaluation is covered in Section 3.7.

The results are presented and discussed in Chapter 4. The data set is covered in Section 4.1, where the subjects, bruises and bruise images will be discussed. Then, the image pre-processing results are covered in 4.2. The results of the bruise feature extraction are covered in Sections 4.3 -4.5. Section 4.6 covers the bruise image segmentation results and the evaluation of the acquisition process is covered in Section 4.7. Finally, the findings are summarized and discussed in Section 4.8.

The final chapter, Chapter 5, of the thesis is the conclusion, in which the achievement of the aforementioned objectives will be discussed. The major challenges and possible future work will also be discussed.

Chapter 2

Theory

This chapter covers the necessary theoretic background to interpret the results of this thesis. The first section covers the basics of human skin and skin bruises. The second section covers how skin bruises are documented in forensic and medical settings, providing a basis for the choice of bruise documentation methods in Chapter 3. Section 2.3 covers existing work on bruise image segmentation. Finally, some image processing concepts are covered in Section 2.4.

2.1 Skin and skin bruises

2.1.1 The biology and optics of skin

Human skin consists of three main layers: the epidermis, dermis and subcutaneous tissues. [47, 74] The epidermis is the outermost layer and its most important function is to protect the body from its environment. [42] The middle layer, dermis, mainly consists of collagen, a type of fibrous protein. [62] The dermis shields the body from stress and strain. The innermost layer, the subcutaneous tissues, consists of fat tissue and provides temperature control and a connection to the underlying muscles and bones. [47] The majority of bruise bleeding occurs in the subcutaneous tissues. [74]

The optical properties of skin depend on the thickness and chemical composition of these three layers, which varies both between and within individuals. [114] Light interacts with tissue in two fundamental ways: absorption and scattering. [141] The optical properties of a material are quantified by its absorption coefficient μ_a and scattering coefficient μ_s. Depending on the chemical composition of the skin, the light will interact differently with it. Different chromophores (the compounds that contribute to the perceived color of tissue) interact differently with light. The absorption spectrum of the skin dictates the penetration depth of light into the skin. [82] The absorption spectra of several chromophores are known, see Figure 2.1.

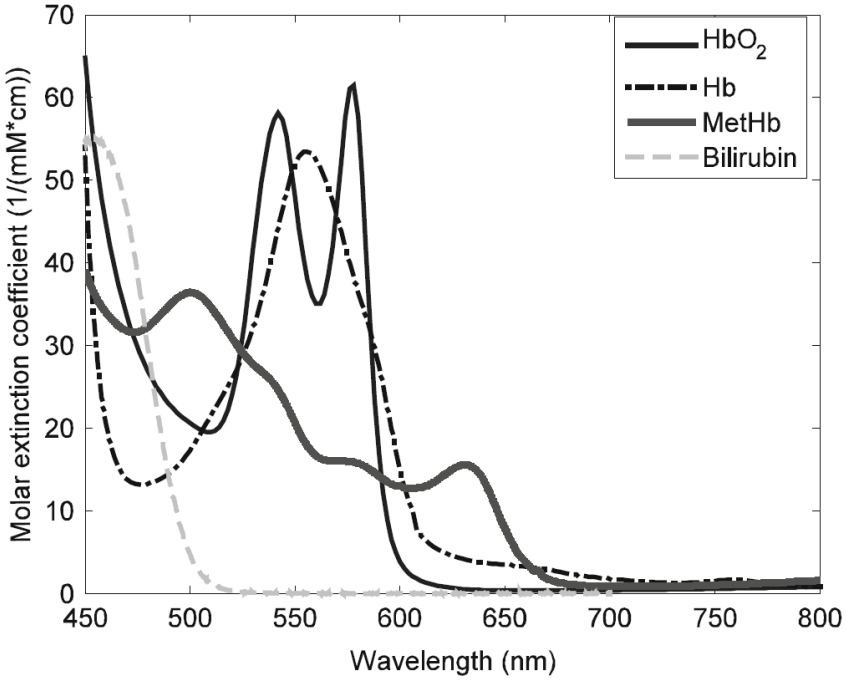


FIGURE 2.1: The molar extinction coefficient of chromophores: oxygenated hemoglobin, deoxygenated hemoglobin, methemoglobin and bilirubin. Taken with permission from [114]

Hemoglobin (red) and melanin (brown/black) are the most important chromophores in normal skin. [41, 112] Assuming the amount of absorption is sufficiently greater than the amount of scattering, it is possible to estimate the concentrations of chromophores by observing how light of different wavelengths interacts with tissue. Hemoglobin, which is found in blood, is one of the main chromophores contributing to the appearance of skin bruises.

2.1.2 Skin bruises: formation and appearance

Skin bruises are formed when a trauma to the skin causes blood vessels to break and blood escapes into the surrounding tissue. [124]. This leads to a visible discoloration of the skin - a bruise. [100] The epidermis must be intact for the injury to be considered a bruise. If the epidermis has been ruptured, the injury may instead be classified as a wound or laceration. [27] Some bruises are accompanied by an edema, or swelling, where water has gathered under the skin. [137] In these cases, the injury may be categorized as a hematoma. [27, 128] Small, pinpoint bruises are sometimes referred to as petecchiae. [127] Petecchiae bruises are caused by bleeding from the smallest blood vessels, and their size range up to 2 mm in diameter. [100]

Bruises change appearance over time, as the distribution of chromophores in the tissue changes. In some cases, it can take days for a bruise to become visible. [102] Some injuries might keep bleeding for some time, causing the amount of blood and thus the chromophore hemoglobin to increase over time. [128] This causes the size of the bruise to increase over time. Black et al. [18] generated 18 bruises and photographed them every other day for 3 weeks. They found bruises to be largest at an age of 3-7 days on average. Bruises may also change location over time, due to gravity. [102] Some time after the injury, the inflammatory response is initiated. [121] Immune cells called macrophages are recruited to the injury site to start breaking down the hemoglobin. [47] The breakdown of hemoglobin is one of the reasons why

bruises change color over time. [65] The hemoglobin breakdown process is shown in Figure 2.2.

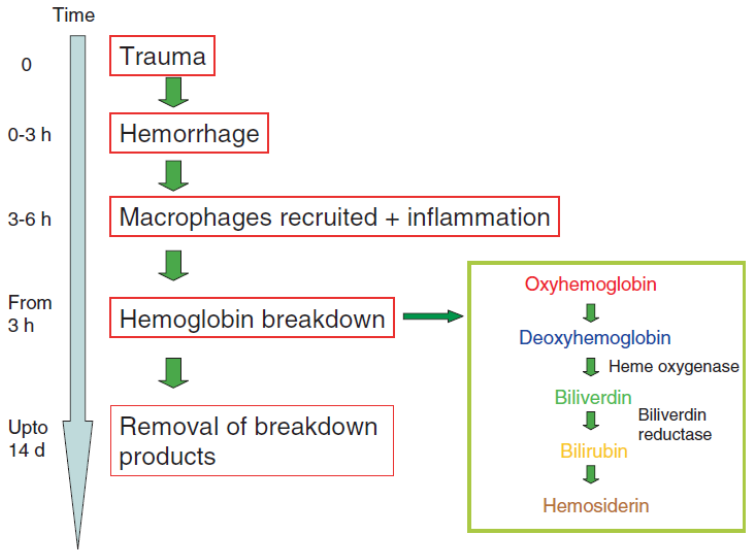


FIGURE 2.2: The hemoglobin breakdown process which causes the change in bruise color over time. Taken with permission from [111]

The temporal color change of bruises does not follow a predictable pattern, however. It was once believed that one could accurately determine the age of bruises from their color alone, but that theory has been disproved. [94, 106, 128] In their study of bruise color and age, Langlois and Gresham [74] found that yellow is the only bruise color that is significant with regards to its age. They found that if a bruise has yellow color, it is likely to be older than 18 hours. However, the same author later stated that it has not been rigorously established when yellow appears in a bruise. [73] In their study of 25 bruises over 2 months, Randeberg et al. [110] found that the freshest bruise with yellow coloration was 40 hours old. Stephenson and Bialas [128] studied 36 bruises in children, and found no presence of yellow in bruises younger than 1 day old. Schwartz and Ricci [123] state that yellow has been observed in bruises of ages between 7 days and 2 weeks.

There are other factors than the age of the bruise that determine its color. Shallow bruises tend to be more red and deeper bruises tend to be more blue, due to the wavelength dependence of light penetration in tissue. [19] The composition of the three skin layers varies throughout the body, which has an impact on bruise appearance. [36] The skin color of the individual has an effect on the color of skin bruises. [58] Bruises are less visible on darker skin than on lighter skin, due to the presence of melanin. [34, 144] [128] The appearance of a bruise might depend on the gender of the individual. According to Vanezis [137] and Black et al. [18], women bruise more easily than men. The age of the individual has an impact on bruise appearance. Young children bruise more easily than adults, [137] since their skin is thinner than adult skin. [114] Older adults also bruise more easily, due to an age-related thinning of the epidermis and increasing fragility of the blood vessels. [140] The appearance of the bruise will depend on the health of the individual. If the individual has a bleeding disorder [78], smokes [114] or uses blood-thinning medicine [65], it will impact the formation and appearance of bruises.

The perceived color of bruises, and all objects, also depend on the the observer.

Color is not a purely physical phenomenon, but also relies on psychological processes. [55] Color vision varies between individuals, due to the differing spectral sensitivities of the color receptors. [96] The ability to see the color yellow declines with age [59]. The human visual system is clearly not an objective color sensor, which further emphasizes that visual inspection of the color of bruises is not suitable to determine its age.

2.2 Bruises in forensics and medicine

2.2.1 Bruise documentation

Documenting bruises with imaging is recommended in the case of suspected child abuse [31, 52, 65, 139], suspected elder abuse [105] and postmortem examination [137]. The photographs both contribute to the medical record and may serve as forensic evidence. It is essential that the documentation is accurate [137] and objective [52].

Digital color imaging is the standard method for documenting bruises. Evans et al. [40], Harris et al. [51] recommend using the RAW file format when documenting skin bruises. JPEG is not recommended due to the compression associated with this file format. They also recommend the image plane to be parallel to the plane of the bruise. According to Evans et al. [40] Digital Single Lens Reflex (DSLR) cameras are preferable to smartphones, since they provide higher quality images.

Cross-polarized imaging has been found to to be the preferred imaging modality for pediatricians when assessing images of bruises in children. [75] Cross-polarization is obtained when light passes through two polarizing filters that are perpendicular to each other. [49] Cross-polarized photography of skin eliminates specular reflection caused by sweat and oiliness and allows for better visibility of internal tissue structures such as bruising. [13, 49] Thus the quality of the bruise image acquisition process can be improved by using cross-polarized light. Baker et al. [15] photographed 75 bruises with both non-polarized and cross-polarized light. They found that cross-polarized light enhanced the visibility of bruises. In contrast, Harris et al. [51] found no significant difference between the quality of bruise size estimation when cross-polarized and non-polarized light were used.

Documentation of bruises should include the recording of important bruise features. The four most commonly mentioned bruise features are color, shape, size and location on the body. [52, 65, 95, 102, 105, 108, 137, 139] Other factors, such as presence of swelling [95, 137] and pain [84, 95] are also recommended by some authors.

2.2.2 Bruise size

The size of bruises is an important factor to consider when determining whether a bruise has been caused by physical abuse. [142] According to Maguire et al. [86], abusive bruises tend to be larger and accidental bruises tend to be smaller. In their study of bruising in the elderly, Wiglesworth et al. [142] found that the bruises that had been caused by physical abuse were significantly larger than the accidental bruises. However, in a study comparing bruises in physically abused ($n = 350$) and non-abused ($n = 156$) children, [67] found that the mean bruise size in abused children was 1.53 cm and 1.57 cm in the non-abused children, indicating no significant size difference between the two classes.

A scale element such as a ruler should be included in the images, since this makes bruise size estimation possible. [24, 108, 143] A standard scale recommended by

the American Board of Forensic Odontology is used by many forensic examiners. [40, 65] It has also been used in several studies of bruise documentation. [51, 132]

The American Board of Forensic Odontology ruler is shown in Figure 2.3.

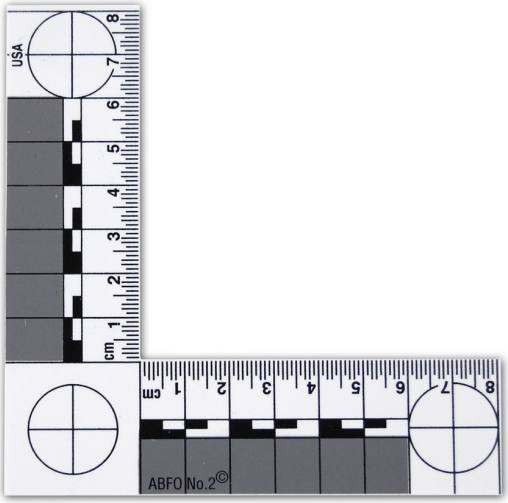


FIGURE 2.3: The American Board of Forensic Odontology ruler. Image from [3].

Payne-James et al. [102] recommend using a ForensiGraph ruler, which also includes a color calibration chart. (See Section 2.2.4 for a discussion of color calibration charts.) The scale or ruler should be placed next to the bruise and in the same plane as the bruise. [40, 65, 143] It should be rigid and have a matte finish. [40]

Scale detection is a prerequisite step to bruise size extraction. If a ruler is included in the image, it must be detected and the distance between the ruler ticks, the ruler spacing, must be estimated.

Ueda et al. [135] used Discrete Fourier Transform (DFT) to detect ruler spacing. Their method requires the ruler to be placed horizontally in the image. Bhalerao and Reynolds [17] also used Fourier analysis for ruler spacing detection. The requirement of the ruler being placed horizontally was removed by including ruler angle detection. However, their method requires the user to manually select an image block containing the ruler.

Harris et al. [51] and Kemp et al. [68] used the Feret diameter, the longest dimension, of bruises as a measure of their size. Vanezis [137] recommends describing the bruise size in terms of at least two dimensions, such as length and width. One method of making these measurements is measuring the maximum length of the bruise and defining width as the maximum bruise width perpendicular to the length axis. [33] Another measure of size is the area of the bruise. The area can be calculated by multiplying the length and width of the bruise. However, this assumes a rectangular bruise shape and will result in bruise areas that are too large. [11] The area can be computed by overlaying a metric grid over the bruise and counting the number of squares [72] or by counting the number of pixels in the bruise segment [64]. Lawson et al. [76] found that there is less variation in bruise size measurements when the size is measured using bruise images than when the size measurements are done in vivo.

There is a difference between how bruise size should be documented according to the literature and how they are described in practice. In Sweaney v. Ada County (1997) [5], the bruise in question is described as "about 1 1/4 inches long and 1/4

inch wide". In State v. Hovig (2009) [6] and State v Powell. (1995) [4], the size of the bruises are described using comparisons to coinage: "approximately the size of a quarter" and "somewhat larger than a silver dollar", respectively.

2.2.3 Bruise shape

The shape of bruises is an important factor to consider when determining their cause. [105] Certain bruise patterns are indicative of abuse. [31] The existence of bruises that have a central clearing may indicate physical abuse. [105] Tramline bruising is the phenomenon of a bruise with a central linear clearing, or alternatively, two parallel linear bruises. Such bruising can be caused by a forceful impact with a cylindrical object, like a rod or a baseball bat. [137] An open-handled slap can also cause tramline bruising, since each finger is a cylindrical object. [65] Bruises created by a high speed paintball impact tend to have a circular shape with a central clearing. [18, 112, 113] The central clearing observed in bruises induced by paintball is caused by the fact that the tissue compression by the paintball itself is not strong enough to cause bruising. Around the edges of the paintball, tissue shearing occurs, which causes the area around the paintball impact site to be more bruised than the central part of the bruise. Another mechanism that causes a central clearing to occur is a strong localized impact to the tissue. The immune reaction will be stronger where the damage was strongest, which causes this area to be healed faster and thus a central clearing develops. [111]

Cluster bruises are more commonly found in cases of physical abuse than in accidental bruise cases. [67, 86, 93] One type of cluster bruise is a bruise caused by a forceful grip or grab. Such bruises often consist of 2-4 small round bruises coinciding with the fingertips of a hand. A thumb bruise may often be found on the other side of the limb. [65, 124] Bruises that have the shape of a recognizable object are not likely to be caused by an accident. [38] Petechial bruising, see Section 2.1.2, are often seen as a part of a strangulation injury. [102]

In court cases, the shape of bruises is often described with geometric terms or by comparing the shape to an object that may have caused the bruise. In State v Powell. (1995) [4], the bruise is described as being the shape of a hand print and in [6], the bruise is described as being mouth-shaped.

2.2.4 Bruise color

As discussed in Section 2.1.2, the color of bruises changes as the bruise heals. The seemingly predictable temporal color change of bruises is the basis of the traditional method of determining the age of a bruise. However, various authors disagree on the specific connection between bruise color and age. Table 2.1 shows three different schemes for determining the age of bruises based on color.

TABLE 2.1: The relation between bruise color and bruise age, according to different authors. Adapted from Stephenson and Bialas [128] and Langlois and Gresham [74]

Author	Color	Age
Adelson [8]	Red/blue	Initial
	Blue/brown	1-3 days
	Yellow/green	1 week
Glaister [44]	Violet	Initial
	Dark blue	1-3 days
	Green	1 week
	Yellow	8-10 days
	Normal	2 weeks
Polson et al. [107]	Red/dark	$<$ 24 h
	Greenish	$\approx$ Day 7
	Yellowing	$\approx$ Day 14
	Normal	Up to 30 days

Even though directly determining the age of bruises from their color alone has been proven inaccurate, the color is still of interest when documenting skin bruises. The presence of many bruises with different colors may be indicative of child abuse. [38, 108, 134] Color is the second most mentioned bruise feature, after location, in a set of court documents analyzed by the author in an unpublished work.

Including a standard color calibration chart in the bruise images facilitates the estimation of bruise color. [40, 101] Payne-James et al. [102] recommends using the ForensiGraph [2] that also includes a ruler. Trefan et al. [132] used a GretagMacbeth Mini Color-Checker [1] in their documentation of bruises in children. Li et al. [77] used a ColorGauge Nano [37] color chart in their documentation of induced bruises.

Barata et al. [16] found that applying color constancy algorithms to skin lesion images improved the performance of melanoma classification. Sully et al. [129] used the eyedropper tool in Adobe Photoshop Elements [7] to apply color correction.

2.3 Bruise image segmentation

Image segmentation is the process of dividing an image into multiple non-overlapping regions. [131] It is an important prerequisite step to estimation of bruise size, shape and color. Johnson and Fazel-Rezai [64] used image segmentation as a prerequisite step to estimation of the size and color of bruises. The chosen image segmentation method was Otsu thresholding. Otsu thresholding, or Otsu binarization. is a segmentation method that uses grayscale histogram thresholding to divide the image into two classes: background and foreground. The threshold is chosen in a such a way that the variance between the two classes is maximized. [99]

Given the threshold t, the class probabilities of the background ω_0 and the foreground ω_1 are defined by equations (2.1) and (2.2).

$$\omega_0 = \sum_{i=1}^{t} p_i \tag{2.1}$$

$$\omega_1 = \sum_{i=k+1}^{L} p_i \tag{2.2}$$

In these equations, p_i is the probability of a given gray-level i, defined as the ration between the number of pixels having that gray-level n_i and the total number of pixels N. L is the number of gray-levels. Otsu thresholding is a global thresholding method, since the same threshold applies to the entire image. [83]

Otsu thresholding is one of the most popular image segmentation techniques, [79, 119, 145] since it is one of the image segmentation methods with the lowest computational complexity. [45] In addition to Otsu thresholding, Johnson and Fazel-Rezai [64] included a supervised processing step in which the user could tune a morphological opening filter to remove small specks from the bruise segment and thus improve the segmentation.

There is a lack of existing work on the topic of bruise image segmentation. Therefore, inspiration will be taken from two similar image segmentation problems: skin lesion segmentation and fruit bruise segmentation.

Skin lesions can be a sign of malignant melanoma, a type of skin cancer. Like bruises, lesions are a discoloration of the skin, but unlike bruises, the discoloration is situated in the outer layer of the skin - the epidermis. Lesion segmentation is often the first step in automated lesion image analysis. [28] Automatic lesion segmentation is challenging due to several factors, including low contrast between the lesion and the surrounding skin, confounding structures such as blood vessels and hairs and irregular lesion borders. [29]

Hameed et al. [48] used Otsu thresholding to segment skin lesions. Erkol et al. [39] used a combination of gray-scale Otsu thresholding and active contours to automatically segment skin lesion images. Active contours, or snakes, is a segmentation method introduced by Kass et al. [66]. The central idea of active contours is to define an initial curve, the *snake*, and then deform it while minimizing an energy function. [136] The energy function is defined by (2.3).

$$E_{\text{snake}} = E_{\text{int}} + E_{\text{image}} + E_{\text{con}} \tag{2.3}$$

Here, E_{int} is the internal energy of the snake that depends on the shape of the snake itself, E_{image} is the image energy that depends on edges in the image, and E_{con} depends on external constraints. The internal energy depends on two parameters: α and β. The first, α, dictates the tendency of the curve to contract. The second, β is a smoothness parameter. If β is set to 0, corners are allowed.

The initial curve is often defined around the object to be segmented, and then it contracts around the object while minimizing the energy function in (2.3). [30] The snake is attracted to areas with high image derivatives, which corresponds to edges in the image. A large number of iterations is required for the snake to get a stable configuration, making active contours a computationally complex algorithm. [92] It is recommended to smooth the image, for example with Gaussian blur, before applying active contours. [136]

Agarwal et al. [10] used k-means clustering to segment skin lesion images. k-means clustering, first introduced by MacQueen [85], is an unsupervised method of separating n observations into k clusters where each observation belongs to the cluster with the nearest cluster mean. The cluster centers are defined such that the within-cluster variance is minimized. The within-cluster variance $var(C_j)$ of cluster C_j is defined by (2.4), adapted from [104].

$$var(C_j) = \sum_{i=1}^{N_j} [d(x_{ij}, \bar{x}_j)]^2 \tag{2.4}$$

Here, N_j is the number of points in cluster C_j, x_{ij} is point number i in cluster C_j, and $\bar{x}_j$ is the cluster mean of cluster C_j, and $d(x_{ij}, \bar{x}_j)$ is the distance between the point and the cluster mean.

K-means clustering is popular because it is easy to implement and has low computational complexity. [61] Agarwal et al. [10] smoothed the lesion images with a Gaussian filter before clustering, to decrease the effect of small intensity changes such as skin hair.

Several authors evaluate the lesion segmentation results by comparing the automatic segmentation with a manual segmentation provided by dermatologists. [39, 122] Hasan et al. [53] used Intersection over Union (IoU) as a measure of similarity between the automatic lesion segmentation and the ground truth.

Upon searching for skin bruise segmentation methods, the majority of articles discuss the segmentation of fruit bruise images. Similarly to skin bruises, fruit bruises are a discoloration caused by an external impact. [23] Du and Sun [35] defined four categories of fruit bruise image segmentation: thresholding, region, edge and classification-based. Thresholding-based segmentation methods divide the image into segments by comparing pixel values to a given threshold value. One example is Otsu thresholding, described above. Region-based approaches define image regions based on some similarity criterion. There are two main sub-divisions of region-based segmentation methods: top-down or splitting and bottom-up or merging. Edge-based segmentation methods use edge detection to find the boundaries between segments. Classification-based methods divide the image pixels into segments, or classes, based on classification techniques such as Bayesian classification and neural networks.

Of the four categories of segmentation methods, thresholding is the most popular for segmentation of fruit bruises. [83] Satone et al. [120] used Otsu thresholding to detect bruises in apples. They pre-processed the images with histogram equalization, which is covered in Section 2.4.3 Roy et al. [117] used k-means clustering to segment fruit bruise images.

2.4 Image processing concepts

2.4.1 Computational color constancy

Human color constancy is our ability as humans to perceive the color of objects or surfaces to be constant even though the illumination changes. Computational color constancy is a class of methods attempting to emulate this ability of the human visual system in computers. [97] The vast majority of computational color constancy research focuses on estimating the unknown illumination and correcting the image accordingly. [9]

The most well-known computational color constancy method is the gray world method [22], which relies on the assumption that, under a neutral light source, the average color of an image should be achromatic - or gray. [43] Any deviation in the average color from gray is caused by the illumination. This deviation becomes the estimate of the illuminant.

Another well known computational method is the white patch method [71]. It relies on another assumption, namely that the whitest part of an image represents a perfect reflectance, and thus is equal to the illuminant. The main idea of both of these methods is to estimate the illuminant (either by the gray world assumption or the white patch assumption) and then adjust the image accordingly. Both methods assume that the illuminant is spatially homogeneous.

2.4.2 Color spaces

Color spaces are a way of representing colors as points in a (most often) three-dimensional space. [63] The most well known color space is RGB. It is based on the thrichromatic theory, that any color can be matched by mixing the three color primaries: red (R), green (G) and blue (B). The straight line that connects black and white is called the achromatic axis. [25] The International Commission on Illumination (CIE) defines the three primaries as monochromatic color stimuli with wavelengths of 700.0, 546.1 and 435.8 nm, respectively.

One way of measuring the difference between two colors is the Euclidean distance between two points in RGB space, shown in (2.5).

$$\Delta C_{12} = \sqrt{(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2} \tag{2.5}$$

However, this distance does not correspond with the perceived color difference. This is the motivation behind creating *perceptually uniform color spaces*, or color spaces in which the Euclidean distance between two colors corresponds by the color difference perceived by a human observer. One of the perceptually uniform color spaces recommended by the CIE is the Lab (or L*a*b*) color space. [32] In this color space, the three components L, a and b do not correspond to three color primaries. Instead, L represents the brightness, or luminance, and a and b represent the hue, or chromaticity. The color difference in Lab space is calculated by (2.6). [60]

$$\Delta E^*_{12} = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2} \tag{2.6}$$

According to Hardeberg [50], a color difference $\Delta E*_{12}$ larger than 3 is perceptible to the human eye.

2.4.3 Histogram equalization

Histogram equalization is a widely used method of contrast enhancement and a popular pre-processing step in image processing applications. [70] The main function of histogram equalization is to flatten the intensity histogram of the image, resulting in higher contrast and an increased dynamic range. The histogram of an image is a discrete function $h(r_k) = n_k$ where r_k is the k-th intensity level and n_k is the number of pixels in the image having that intensity level. [46] The probability of a pixel having a given intensity value is given by the fraction $\frac{n_k}{n}$, where n is the total number of pixels in the image.

Histogram equalization is a mapping $s = T(r)$ transforming the input image r into an equalized output image s. The transformation is obtained by mapping every pixel in the image with intensity level r_k to have intensity level s_k in the output image, according to equation (2.8).

$$s_k = T(r_k) \tag{2.7}$$

$$= \sum_{j=0}^{k} \frac{n_j}{n} \tag{2.8}$$

The sum $\sum_{j=0}^{k} \frac{n_j}{n}$ represents the discrete cumulative probability distribution of the image. In other words, histogram equalization is obtained by applying the discrete cumulative probability function of the image as an image transformation.

Chapter 3

Methods

The workflow of the project is shown in Figure 3.1.

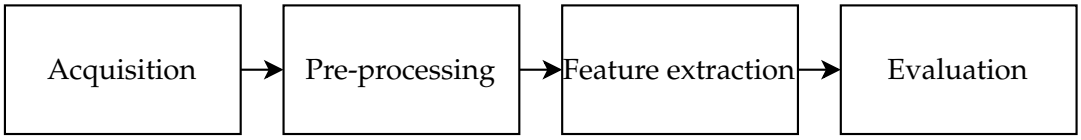


FIGURE 3.1: Project workflow.

The first section covers how the images of bruises are gathered from study participants, as well as the acquisition of other relevant data and management of data. Then, the image processing of these images is covered in the subsequent sections. Pre-processing includes extraction of region of interest (ROI) and image segmentation.

The three bruise features that have been selected to be extracted are size, shape and color. As discussed in Section 2.2, the size, shape, color and location on the body are the four most commonly mentioned features to be documented. Location has been excluded as a feature because recognizing bruise location from an image is a significantly different problem than estimating the size, shape and color of the bruise.

The final section of this chapter covers the evaluation of the image acquisition and processing. The precision and accuracy of each feature extraction will be evaluated.

3.1 Data acquisition and management

3.1.1 Subject selection

Study participants are selected based on the following criteria:

- Otherwise healthy adult
- Has a visible bruise
- The bruise is of an accidental nature (not caused by violence)
- The bruise is preferably of known age and origin

Participants are recruited by contacting sports groups and through social media.

3.1.2 Image acquisition

The images are taken by the study participants, using their own smartphones. Even though smartphones provide lower quality images than DSLR cameras, see Section **??**, it can be expected that more participants own a smartphone than own a DSLR camera. The bruise images that are used as forensic evidence in criminal court are often acquired with smartphones. [34]

The instructions given to the participants are included in Appendices A and B. An alternative version of the instructions was made as it became apparent that the original instructions were too time-consuming to follow for many participants.

In the original version of the instructions (Appendix A), the participants are asked to take 10 photos of their bruises (bruise images) and 10 photos of skin on the opposite side of the body (reference images). The reference images are included to make estimation of a skin color baseline possible. For both the bruise images and reference images, 5 images should be taken with flash and 5 without flash. The reason for having 5 repetitions for each imaging condition is to make the estimation of random error possible. In the simplified version of the instructions (Appendix B), the reference images are omitted. Further, the participants only take 2 repetitions of flash and no flash images.

A white paper should be included in each image. The purpose of the white paper is to provide a reference for computational color constancy, which can facilitate the color estimation of bruises under unknown illumination. As discussed in Section 2.2.4, a standardized color calibration chart is preferred. However, such color charts are not readily available, making them unsuitable for this self-documentation method. White paper, on the other hand, is commonly found in every home or workplace.

A ruler should be included in each image. The purpose of the ruler is to make size estimation of the bruise possible. Again, no standardized ruler is required, since such rulers are not readily available. Instead, the participants can use any ruler they have access to.

In the original version of the instructions, the participants are asked to measure the distance between the smartphone camera and the bruise. This step was removed in the simplified instructions, as it proved to be difficult.

Originally, participants were required to know when and how their bruise was caused. This requirement was removed in the simplified instructions, to increase the number of eligible participants.

There is a trade-off between quality and quantity of data. The original version of the image acquisition instructions provided better documentation of the bruises, but the complexity of the instructions potentially caused several participants to not finish the documentation. If the image acquisition process is too difficult, few people can be persuaded to do it. By decreasing the quantity of data provided by each participant, the number of participants could be increased.

The participants are asked to avoid strong shadows and fluorescent light, as these can cause difficulties when extracting the color of the bruise. They are also asked to keep the skin in focus, and to have the entire bruise be visible in the bruise images. Finally, they are asked to try to keep the distance and angle of the phone constant.

3.1.3 Questionnaire

The images are collected using Nettskjema, a secure platform for gathering research data. In addition to the images, relevant data about the participants and their bruises are gathered through a questionnaire. The following data is gathered:

- **Personal:** Sex and birth year
- **Health:** Whether they smoke and whether they use blood-thinning medicine
- **About the bruise**: Date and time of bruise formation, cause of bruise
- **Technical**: Type of phone, type of illumination

The questions about the age and cause of the bruise are optional to answer, in case the particpant does not know the origin of their bruise. They also provide a possibility for the participant to provide a range of possible times for the creation of the bruise, in the case of uncertainty. The health questions are motivated by the fact that health factors impact the appearance of bruises, see Section 2.1.2.

At the end of the questionnaire, some questions about the method itself are included. The participants are asked to rank the difficulty of the following tasks:

1. Understand the instructions
2. Find a ruler and white paper
3. Measure the distance between the bruise and camera
4. Photograph the bruise
5. Upload the photos

Each task is ranked on a scale from (1) very easy to (5) very difficult. They are also given the possibility to provide any other feedback in an open text box.

3.1.4 Data management and privacy

Participants are identified by a Participant ID consisting of one letter and four digits, for example *A0001*. The document that connects the participant's personal information with the Participant ID's is stored separately and will be destroyed within five years after project completion.

Participants sign an informed consent form before submitting their data. An ethics approval application was submitted to the Regional Committee for Medical and Health Research Ethics (REK) [115]. Due to the nature of the project, no ethical approval is needed.

Bruises are referred to by a separate Bruise ID. By not using the Participant ID, it is possible for one participant to submit more than one bruise.

3.2 ROI extraction

The region of interest (ROI) must be extracted for each bruise image. ROI extraction is done according to these criteria:

- The entire bruise should be included in the ROI, unless the bruise is not entirely visible in the image

- If possible, only the bruise and surrounding skin should be included in the ROI
- Images from the same series should have as similar ROI's as possible

Each bruise has two series of images: flash images and no flash images. The first images in each series is shown to the user using a graphical interface. The user manually defines a ROI by selecting two corners of the ROI rectangle. Then, the coordinates of these corners are used to extract ROI's of the remaining images of the series. The purpose of this is to increase the similarity of the ROI's within the same series, assuming that the camera has not moved significantly between images. However, this will not always be the cause. Therefore, the user is given the chance to reject the ROI selection and manually select a new ROI for the remaining images in the series.

3.3 Segmentation

The bruise images must be segmented before the size, shape and color can be estimated. Image segmentation means dividing an image into segments that belong to a certain class. More background on segmentation is found in Section 2.3.

The two segmentation classes are bruised skin and healthy skin. Four segmentation methods were selected for comparison: Otsu binarization, k-means clustering without histogram equalization, k-means clustering with histogram equalization and active contours. These methods have been selected based on the existing work on skin bruise, fruit bruise and skin lesion segmentation, see Section 2.3.

3.3.1 Otsu thresholding

The flowchart of Otsu thresholding, also called Otsu binarization, is shown in Figure 3.2. The theory of Otsu thresholding is covered in Section 2.3.

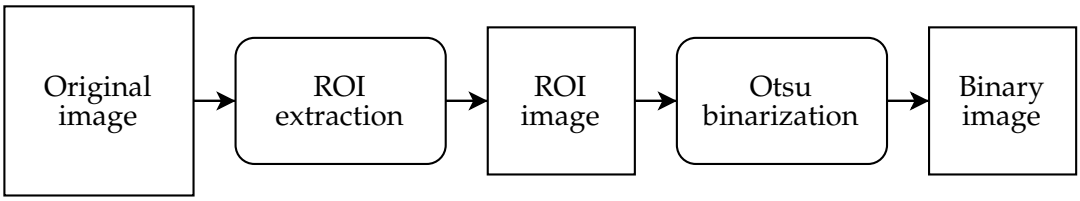


FIGURE 3.2: Otsu thresholding.

Once the region of interest has been extracted, the ROI image is divided into two classes by Otsu thresholding. Otsu thresholding is implemented in Python [109], using the `cv2.threshold()` method from the OpenCV library [20].

3.3.2 K-means clustering

The flowchart of k-means clustering is shown in Figure 3.3 The theory of k-means clustering is covered in Section 2.3.

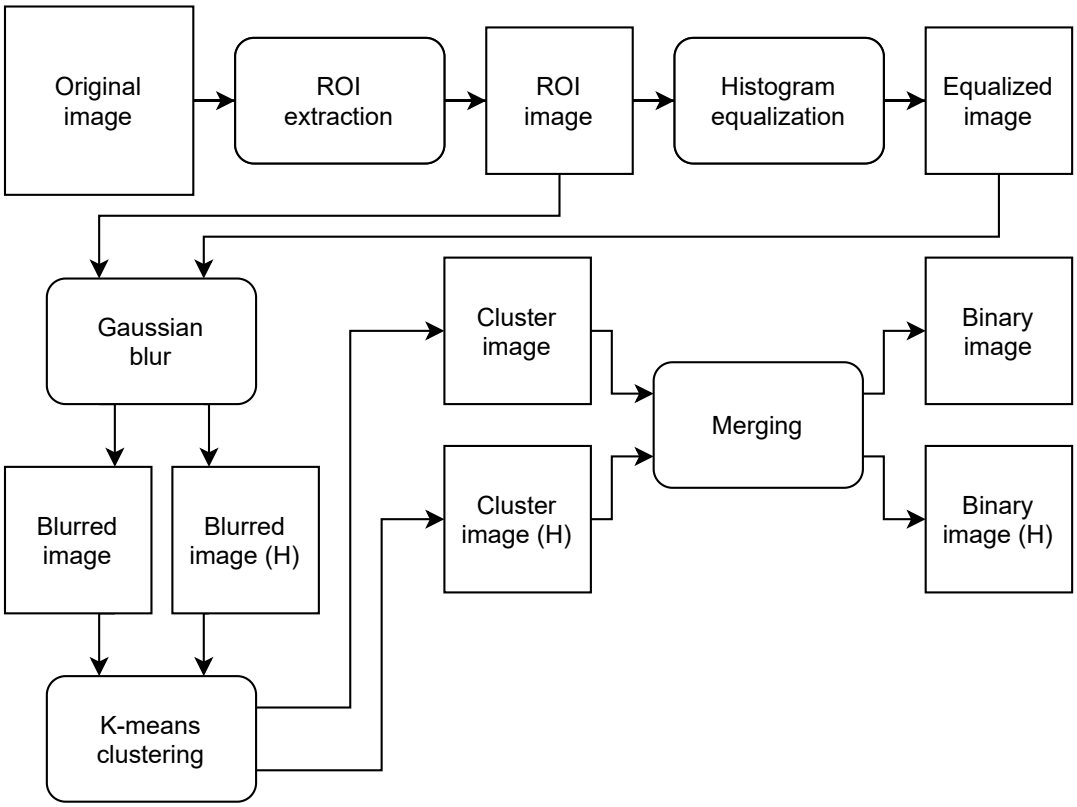


FIGURE 3.3: K-means clustering with and without histogram equalization.

First, the region of interest is extracted. Then, Gaussian blur is applied to two versions of the image: one version with histogram equalization and one without. In Figure 3.3, the images with histogram equalization are marked with an H. Histogram equalization is covered in Section 2.4.3. The reason for including histogram equalization is that it increases the contrast of images and thus may increase the visibility of the bruise, improving the segmentation results.

The reason for including a Gaussian blur step is to reduce the effect of small intensity variation in the image, as discussed in Section 2.3. The standard deviation of the Gaussian blur is chosen dynamically based on image size. The blurred images are fed into the k-means clustering method. The output of the k-means clustering method is a cluster image: an image that has been divided into k classes. The segment that includes the middle pixel is defined as the bruise class and the remaining segments are merged and defined as the not bruise class. After this merging step, what remains is a binary image.

The chosen implementation of k-means clustering is the `slic()` method from the **scikit-image** segmentation library [136] in Python [109] Histogram equalization is also implemented in Python, using the `equalize_hist()` method from the **scikit-image** exposure library. The Gaussian blur is implemented using the `gaussian()` method from the **scikit-image** filters library.

3.3.3 Active contours

The flowchart of active contours segmentation is shown in Figure 3.4. The theory of active contours is covered in Section 2.3.

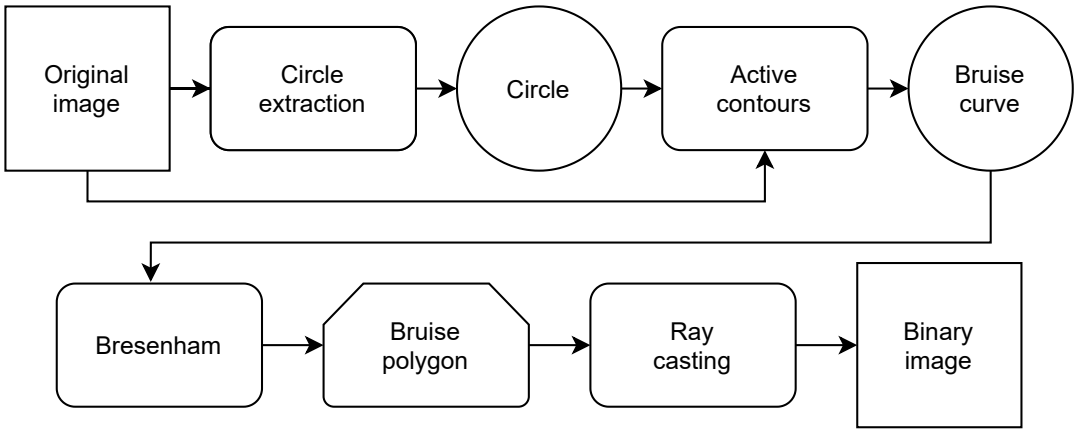


FIGURE 3.4: Active contours.

First, the initial curve of the algorithm is chosen. In this implementation, the curve is a circle around the bruise defined by the user. The circle and the original image are fed into the active contours algorithm. The curve contracts around the bruise, tending towards areas of the image with high derivatives. When the algorithm stops, it returns a set of points along the final curve. This is an estimate of the boundary of the bruise.

The points are converted to integers to become pixels in the image. A polygon representing the bruise is created by connecting each point with a straight line. The Bresenham algorithm [21] was used to find the connecting lines. Then, in order to transform the polygon into a bruise segment, the ray casting point in polygon algorithm [125] was used to fill the polygon.

There are several algorithm parameters that affect the outcome of the active contours method:

- α: A parameter that controls how fast the contour contracts. Higher values = faster contractions.
- β: Smoothness parameter. Higher values = smoother contour.
- **N**: Number of points in starting circle.

These parameters are selected before segmenting and not tuned during the segmentation process. The chosen implementation of active contours is the `active_contours()` method from the **scikit-image** segmentation library [136] in Python [109]. Petr Viktorin's Python implementation [138] of the Breseham algorithm was used. Ray casting was implemented in Python by the author, as no existing implementation was found.

3.4 Size estimation

Size is the first of three bruise features that are estimated in this project. Two different size measures have been selected: bruise area and the length and width of the bruise. As discussed in Section 2.2.2, extracting the size of bruises from images may result in higher precision than manual measurement of the bruise. [76]

3.4.1 Ruler detection

In order to calculate the size of the bruise, the size of each pixel must be found. This is done by measuring the number of pixels between each millimeter tick on the ruler. The automatic ruler detection method works as follows.

1. **Find ruler angle**. The image is converted to gray-scale. Then, the edges of the image are detected using Canny edge detection [26]. Then, the Hough transform [57] is used to detect lines in the edge image. The angles of each line is computed. The ruler angle θ_R is defined as the median angle of all detected lines.

2. **Find ruler location**. The image is rotated by θ_R, such that the ruler is horizontal in the image. Lines are detected in the same way as in step (1). All lines with an angle that deviates less than a defined threshold Δ_θ from the horizontal are counted as probable ruler lines. The median y coordinate of these lines is chosen as the probable y coordinate of the ruler y_R.

3. **Find ruler tick separation**. N_l horizontal lines are scanned around the probable y coordinate of the ruler y_R. The autocorrelation of each line is computed. The median distance between the autocorrelation peaks d and the median prominence of the autocorrelation peaks p is computed.

4. **Repeat** steps (1) to (3) N times. Chose the median distance d that corresponds to the highest median prominence p. This is the number of pixels per millimeter.

This method has been inspired by existing work by Ueda et al. [135] and Bhalerao and Reynolds [17], described in Section 2.2.2. The existing methods either require the ruler to be horizontal or the user to mark the location of the ruler, which does not allow for flexibility in acquisition process. The proposed method does not have such requirements.

Automatic ruler detection is performed on all bruise images. The prominence of the autocorrelation peaks is used as a confidence heuristic. The images that have a low prominence value are deemed as ruler detection failures, and in these cases manual ruler detection is done. Manual ruler detection is done by asking the user to mark two points on the ruler. This bypasses the two first steps: finding the ruler angle and finding the ruler location. The only part that remains is finding the ruler tick separation.

The conversion to gray-scale is done using the `rgb2gray()` method from the **scikit-image** color library [136]. The Canny edge detection is implemented by using the `canny()` method from the **scikit-image** feature extraction library. The **scikit-image** signal processing library is used to compute the ruler tick separation, using the `correlate()`, `find_peaks()` and `peak_prominences()` methods.

3.4.2 Area

To compute the area of the bruise, one must first compute the area of one pixel. The area of one pixel is defined by equation (3.1), where A_p is the area of the pixel in mm^2 and $ppmm$ is the number of pixels per millimeter.

$$A_p = \frac{1}{ppmm^2} \tag{3.1}$$

Then, the area of the bruise is given by the number of pixels within the bruise segment N_p multiplied by the area of one pixel A_p.

$$A_b = N_p \times A_p \tag{3.2}$$

Bruise area is a computationally simple description of the bruise size, but is not a typical descriptor of bruise size. A more common approach is describing the size of the bruise in terms of its dimensions, or length and width. This is covered in the next subsection.

3.4.3 Length and width

Bruises are often described by their length and width [95, 137]. This measure serves as a description of both the size and shape of the bruise. As described in Section 2.2.2, one way of measuring the length and width is first identifying the longest dimension of the bruise and then measuring the width of the bruise along a line perpendicular to the length axis. The following method is used to compute the length of the bruise.

First, the top, bottom, leftmost and rightmost points of the bruise are identified. These are shown as blue points in Figure 3.5.

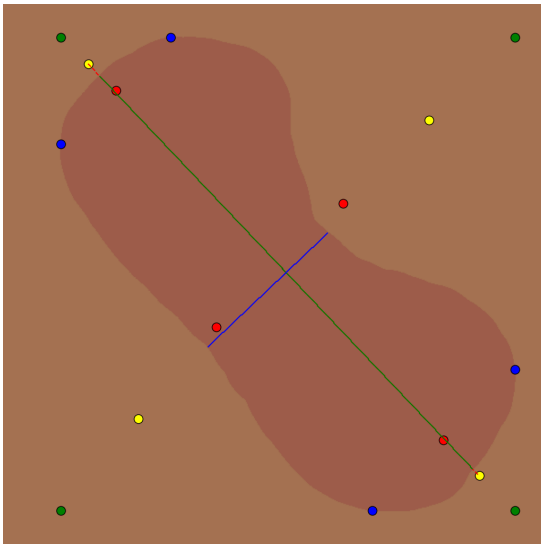


FIGURE 3.5: Bruise length extraction. The extracted length of the bruise is the green line and the width is the blue line.

Then the midpoints are found: the top-left, top-right, bottom-left and bottom-right points. These are shown as red points in Figure 3.5. Then, the "corner points" are identified. These are the points with the y-value of the top/bottom points and x-value of the left-right points. The rectangle defined by these four corner points becomes a bounding box around the bruise. The corner points are shown as green points in Figure 3.5. Then, the midpoints between the red midpoints and the green corner points are calculated. These are shown as yellow points in Figure 3.5.

Once all the points have been found, the following distances are calculated:

- The distance between the top point and the bottom point
- The distance between the leftmost point and the rightmost point
- The distance between the top-left yellow midpoint and the bottom-right yellow midpoint

- The distance between the top-right yellow midpoint and the bottom-left yellow midpoint

The longest distance is identified. A line is drawn between the two points with the longest distance. This is shown as a red dashed line in Figure 3.5. In this case, the longest line was the distance between the top-left yellow midpoint and the bottom-right yellow midpoint. Finally, the line is shortened to only include points within the bruise segment. The result is the green line in Figure 3.5.

In order to find the width, the length line is rotated 90 degrees around its midpoint and shortened to only include points within the bruise segment. The result is the blue line in Figure 3.5.

The width and length in millimeters are computed by multiplying the width and length in pixels by the length of one pixel, given by $\frac{1}{ppmm}$, where $ppmm$ is the number of pixels per millimeter.

3.5 Shape estimation

Shape is the second of three bruise features that are estimated in this project. Roundness has been selected as a shape descriptor.

3.5.1 Roundness

The shape of the bruise can be described by its roundness. One measure of bruise roundness R_b is the ratio between the width w_b and length l_b of the bruise, as defined by (3.3).

$$R_b = \frac{w_b}{l_b} \tag{3.3}$$

Length and width extraction is covered in Section 3.4.3. If the ratio is 1, the bruise has the same length and width. An example of a bruise with a width/length ratio close to one is shown in Figure 3.6, where the green and blue lines are the extracted length and width lines.

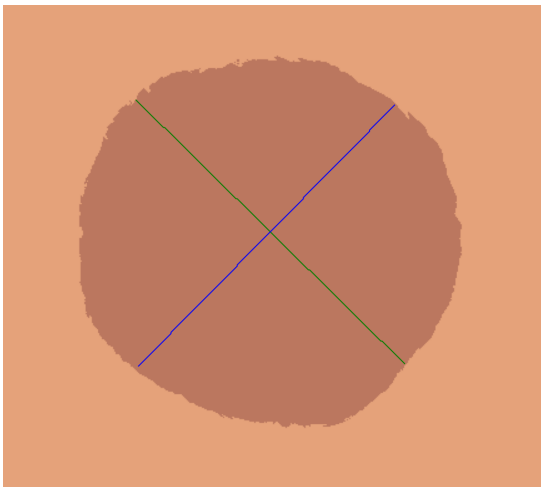


FIGURE 3.6: An example of a round bruise.

3.6 Color estimation

Color is the third and final bruise feature that is estimated in this project. As discussed in Section 2.1.2, the human visual system is not suited to objectively measure the color of bruises. Therefore it is interesting to develop an imaging-based approach to do so. Two color measures are used: bruise contrast and average bruise color.

3.6.1 Contrast

The first bruise color descriptor is the contrast between the bruise and the surrounding skin. This is done by finding the average RGB value of the bruise segment and of the surrounding skin. Then, these values are converted to *Lab*-space. See Section 2.4.2 for a discussion of color spaces. The bruise contrast is defined by equation (3.4), where L_s is the average luminance of the skin and L_b is the average luminance of the bruise.

$$\Delta L = L_s - L_b \tag{3.4}$$

Bruises are in general darker than the surrounding skin, so ΔL is expected to be a positive number. Imaging factors such as shadows can cause this expectation to be violated.

3.6.2 Average bruise color

The second chosen color descriptor is average bruise color, or the average RGB value of the bruise segment. Since the color of an object in an image depends on the illumination as well as the reflectance of the object itself, computational color constancy is applied to remove the contribution of the illumination. Section 2.4.1 covers the theory of computational color constancy.

Two computational color constancy methods were selected: white patch and gray world. The two methods are implemented in MatLab [88], using the Color Constancy Toolbox [87]. The performance of the computational color constancy methods will be evaluated by sampling the color of the white paper in the image. The method that provides the whitest paper is considered to have performed the best. Here, whiteness is defined as the closeness to the achromatic axis. (See Section 2.4.2.

3.7 Evaluation

3.7.1 Evaluation of data set

The quality of the bruise images will be evaluated with respect to the following criteria:

1. Include paper and ruler in image
2. Entire bruise visible
3. Good lighting without prominent shadows
4. Bruise and ruler in focus
5. Ruler parallel to camera and bruise
6. Keep the angle and distance of the camera constant

3.7.2 Evaluation of acquisition method

The acquisition method will be evaluated by gathering feedback from the participants, see Section 3.1.3. The average difficulty of each task will be computed, as well as the standard deviation. The acquisition method will also be evaluated by investigating how well the instructions were followed, see Section 3.7.1

3.7.3 Evaluation of size estimation

The accuracy of the pixel per millimeter extraction will be evaluated by comparing to the ground truth provided by measuring the ruler in GIMP [130]. The performance of the pixel per millimeter extraction method will be evaluated by counting how often manual ruler detection is needed.

The precision of the bruise area estimation is evaluated by computing the standard deviation of the area estimates for each bruise. The accuracy of the length and width estimation is evaluated by comparing to a ground truth provided by measuring the bruise in GIMP [130].

3.7.4 Evaluation of shape estimation

There is no well defined ground truth for the shape of the bruise, other than the perceived shape provided by the participant in the Nettskjema questionnaire. The precision of the shape estimation will be evaluated by computing the standard deviation of the roundness estimate for each bruise. The accuracy of the shape estimation will be evaluated by sampling some round and less round bruises and comparing their estimated roundness values.

3.7.5 Evaluation of color estimation

The robustness with respect to illumination will be evaluated by comparing the average bruise color with and without flash. The color difference ΔC_{fb} is found by calculating the Euclidean distance in RGB space between the average color of the bruise in the flash images (R_f, G_f, B_f) and the average color of the bruise in the no flash images (R_n, G_n, B_n).

$$\Delta C_{fb} = \sqrt{(R_f - R_n)^2 + (G_f - G_n)^2 + (B_f - B_n)^2} \tag{3.5}$$

There is no well defined ground truth for the color of the bruise, other than the perceived color provided by the participant in the Nettskjema questionnaire. The precision of the color estimation will be evaluated by computing the standard deviation of the two color measures: contrast and average bruise color for each bruise. Contrast is a scalar, so the standard deviation will be computed using the built-in standard deviation method from the math library in Python [109]. In the case of RGB color, the standard deviation must be calculated explicitly by using the color difference measure defined by (3.5). The standard deviation will be computed within each illumination setting (flash and no flash), to separate the effects of illumination difference and random variation. The color robustness measure encompasses the variation in color estimate caused by change in illumination, and the color precision measure encompasses the variation in color estimate caused by random error.

3.7.6 Evaluation of segmentation

Since there is no segmentation ground truth, the accuracy of the segmentation methods cannot be evaluated directly. Instead, the four selected segmentation methods (Otsu thresholding, k-means clustering without histogram equalization, k-means clustering with histogram equalization and active contours) will be evaluated indirectly based on the performance of the feature extraction. Each segmentation method will be ranked from 1 (best) to 4 (worst) in each of the following categories:

- Area precision
- Length accuracy
- Width accuracy
- Roundness precision
- Roundness accuracy
- Contrast precision
- Color robustness
- Color precision
- Segmentation speed

The scores will be weighted such that size, shape, color estimation and segmentation speed are each weighted equally. An average ranking will be calculated, resulting in a final ranking of the four segmentation methods.

Chapter 4

Results and Discussion

This chapter presents and discusses the results of the image acquisition and image processing, following the evaluation method covered in Section 3.7 . The acquired data set is covered in Section 4.1. The region of interest (ROI) extraction is covered in Section 4.2. The estimation of the three features - size, shape and color - is covered in Sections 4.3 - 4.5. Based on these results, the bruise image segmentation will be evaluated in Section 4.6. The acquisition method will be discussed in Section 4.7. Finally, the results are summarized and discussed in Section 4.8.

4.1 Data set

4.1.1 Subjects

24 participants were recruited through local sports groups and social media. The distribution of participant ages is shown in Figure 4.1.

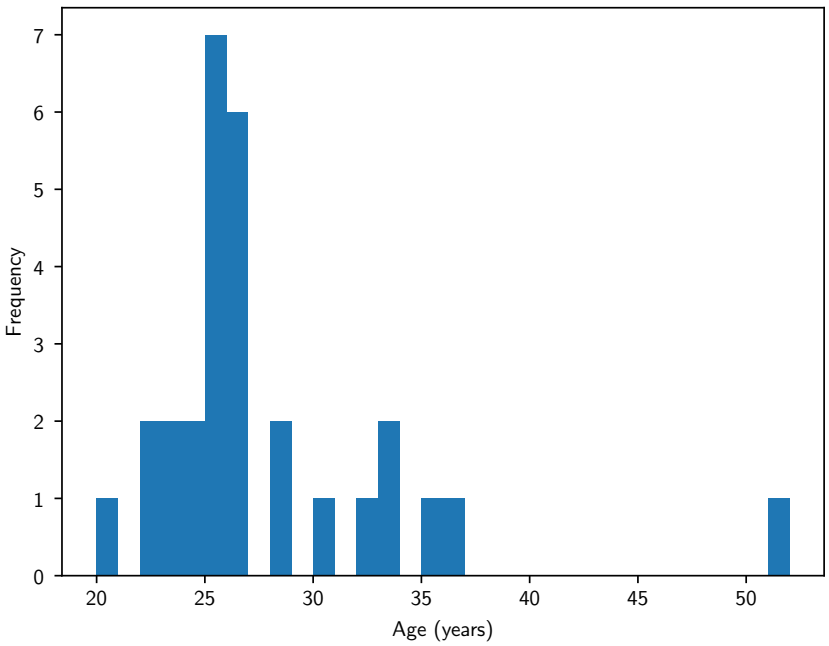


FIGURE 4.1: Distribution of participant ages.

The ratio of female and male participants is shown in Figure 4.2

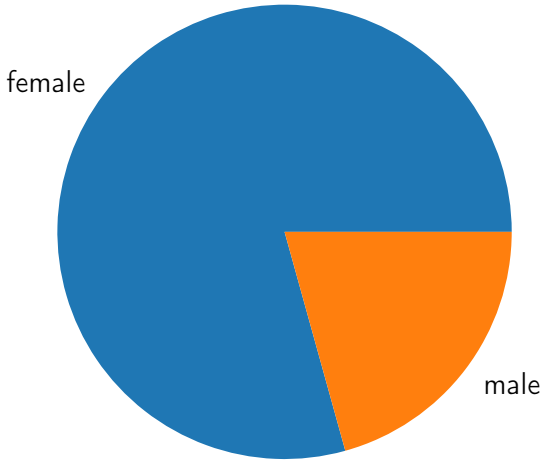


FIGURE 4.2: Ratio of female and male participants.

There is a clear bias towards participants of a similar demographic as the author, the author being female and 24 years old. The age distribution makes sense, since participants were recruited through social media (where friends have a higher chance of seeing the post) and university sports (university students tend to be aged 20-30). It is not clear why the vast majority of the participants are female.

All the participants that were recruited have light skin color. This is explained by the fact that primarily Norwegian participants were recruited. Since the appearance of skin bruises depends on the underlying skin color (see Section 2.1.2), it would be preferable to have more skin colors represented in the data set.

4.1.2 Bruises

Two of the participants submitted two bruises, so 26 bruises were gathered in total. For bruises 5, 20 and 26, the participants submitted follow-up images of their bruises - e.g. answered the questionnaire at two different times.

An overview of the bruises is shown in Table 4.1. Bruises acquired with the original instructions are marked with orange, and the bruises documented with the simplified instructions are marked with blue. See Section 3.1.2 for a description of the two versions of the instructions.

TABLE 4.1: Overview of the bruises: bruise ID (BID), features (size, shape, color and location) and additional information (age and cause).

BID	Size	Shape	Color	Location	Age	Cause
1	$R = 2$ cm	Round or star-shaped	Red, blue, purple, yellow	Knee	4 d	Falling
2	2 x 2 cm	Trapeze-shaped	Blue-green, yellow, brown	Elbow	5 d	Snowboard
3	2-3 cm	Round	Brown	Knee	7 d	Cause
4	$D = 4$ cm	2 circles	Blue-gray	Leg	3 d	N/A
5	6 x 2-3 cm, 4 x 2 cm	Oval	Purple, pink	Hip	7 h, 2 d	Exercise
6	7 x 2.5 cm	Flat oval	Yellow, green with purple spots	Knee	3 d	Hit knee against table
7	2.5 x 2 cm	Triangle	Purple	Chin	4 d	Acrobatics
8	Large grape	Oval	Green, purple	Thigh	5 d	N/A
9	$D = 3$ cm	Round	Brown, red	Calf	2 d	Poledance
10	< 1 cm	Several round	Yellow, green, blue	Arm	18 h	Poledance
11	$D \approx 0.2 - 1$ cm	Several round	Blue, yellow, purple	Thigh	9 d	Poledance
12	$D \approx 3$ cm	Round, oval	Yellow	Elbow	7 d	Cat bite through clothes
13	16 x 9 cm	Spotted	Blue	Thigh	2 d	Rugby
14	6 x 4 cm	Round	Purple	Knee	3 d	Rugby
15	10 x 5-6 cm	Oval	Yellow, green	Elbow	5 d	Acrobatics
16	$D \approx 3$ cm	Round, diffuse edges	Purple, blue	Arm	2 d	Weight training
17	$D \approx 4$ cm	Round with protusion	Dark blue-green, purple	Thigh	3 d	Collision with oven
18	$D = 2$ cm	Round	Blue and yellow	Thigh	3 d	Biking
19	Small	Round	Purple, green, blue	Arm	4 d	Child bite through clothes
20	≈ 5 x 3 cm	Ellipse with hole	Purple	Calf	8 d; 13 d	Hit table
21	≈ 2.5 cm	Circle	Brown	Knee	2 d	N/A
22	≈ 14 x 6 cm	Oval	Blue and red	Foot	20 h	Dance
23	≈ 1 x 3 cm	Oblong	Brown	Calf	3 d	Volleyball
24	≈ 3 cm	Oblong	Yellow, purple, blue, green	Arm	5 d	Hit car door
25	20 x 4-5 cm	Elongated	Dark blue	Thigh	3 d	Trampoline
26	≈ 7 cm	Diffuse	Green-yellow; some blue	Knee	3 d; 5 d	N/A

The participants were given no instructions about how to describe the size, shape and color of their bruises. Thus there is a large variation in how they have chosen to describe these features. Most of the participants describe the size of their bruise using one or more measurements, often radius, diameter or length and width. One participant described the size of their bruise using a comparison with fruit ("large grape") and another one simply described it as "small". When describing the shape of the bruise, most participants compared the outer contour of the bruise to a geometric shape such as circle or oval. Some participants also included other factors such as diffuseness and the existence of a central clearing. When describing color, all participants listed one or more colors present in their bruise.

The colors of the bruises were compared with their ages. The connection between bruise color and age is discussed in Section 2.1.2, with emphasis on when the color yellow appears. Figure 4.3 shows the age distribution of yellow bruises, as well as

the general age distribution of the bruises in the data set.

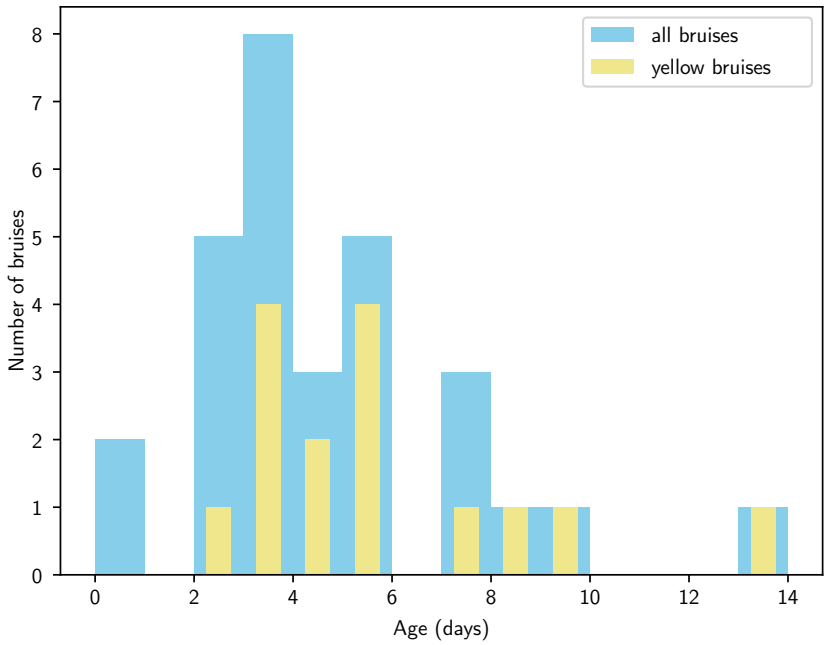
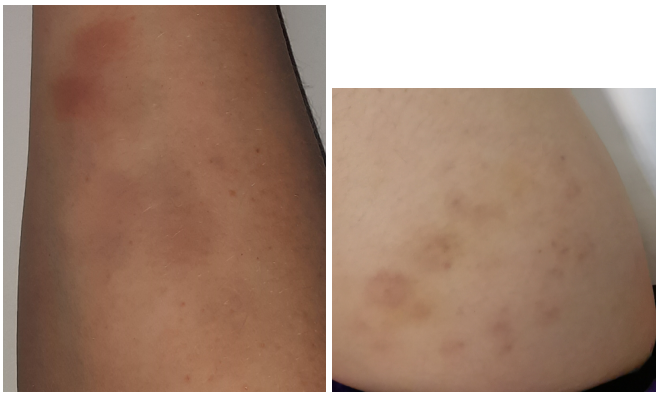


FIGURE 4.3: The age distribution of yellow bruises.

The youngest bruise with observed yellow in it is 2 days old. This fits with the findings of Langlois and Gresham [74] and Randeberg et al. [110] that were discussed in Section 2.1.2. The oldest bruise in the data set, with an age of 13 days, had yellow in it. From the histogram, it is clear that the distribution of yellow bruises is skewed toward higher ages compared to bruises in general.

Upon inspection of the data set, several bruise classes were recurring. First, cluster bruises. These bruises consist of several smaller bruises in a cluster. This type of bruises is interesting, since it may be indicative of abuse. (See Section 2.2.3.) Defining their size and shape can be challenging, since they can both be seen as one large bruise and as many small ones. Two examples of cluster bruises are shown in Figure 4.4.



(A) Bruise 10 (B) Bruise 11

FIGURE 4.4: Two examples of cluster bruises. (Bruise 10 and 11)

The second class is "blob"-like bruises. These bruises are one contiguous shape with well defined edges. They are the most straight-forward to segment. An example of a blob bruise is shown in Figure 4.5.

FIGURE 4.5: An example of a blob bruise. (Bruise 17)

Diffuse bruises are probably the most complicated to segment. They have a faint coloration and the edges are diffuse. Defining their border is difficult both for humans and machines. An example of a diffuse bruise is shown in Figure 4.6.

FIGURE 4.6: An example of a diffuse bruise. (Bruise 9)

The fourth class of bruises are "central clearing" bruises. These bruises are characterized by having a section of lighter or no discoloration within the bruise. This class of bruises is interesting to analyze, since the presence of a central clearing may indicate that the bruise was caused by a violent impact. [105] An example of a central clearing bruise is shown in Figure 4.7.

FIGURE 4.7: An example of a central clearing bruise. (Bruise 20)

The final class, complex, only applies to a single bruise. Its shape its too irregular to fit the description of a blob-like bruise.

FIGURE 4.8: Bruise 25 is classified as a complex bruise due to its irregular shape.

This bruise is significantly more severe than the other bruises in the data set, and is the only bruise that caused hospitalization of the participant.

The classifications of the bruises in the data set are shown in Table 4.2.

TABLE 4.2: The classifications of the bruises in the data set.

BID	Characteristics
1	Blob
2	Diffuse
3	Diffuse
4	Cluster
5	Blob
6	Blob
7	Blob
8	Blob
9	Diffuse
10	Cluster
11	Cluster
12	Faint
13	Cluster
14	Central clearing
15	Diffuse
16	Blob
17	Blob
18	Blob
19	Blob
20	Central clearing
21	Blob
22	Diffuse
23	Blob
24	Blob
25	Complex
26	Cluster

Bruise characteristics exist on a continuous scale between these classes. Some of the bruises exhibit both blob-like and diffuse characteristics, for example.

4.1.3 Image quality

176 bruise images were gathered. The quality of the acquired image data set was evaluated according to the following criteria:

1. Include paper and ruler in image
2. Entire bruise visible
3. Good lighting without prominent shadows
4. Bruise and ruler in focus
5. Ruler parallel to camera and bruise
6. Keep the angle and distance of the camera constant

The results are shown in Table 4.3.

Bruise	Crit. 1	Crit. 2	Crit. 3	Crit. 4	Crit. 5	Crit. 6
1	Yes	Yes	Yes	Yes	No	No
2	Yes	Yes	Yes	No	Yes	Yes
3	Yes	Yes	No	Yes	Yes	Yes
4	Yes	Yes	Yes	Yes	No	Yes
5	Partial	No	No	Partial	Yes	No
6	Yes	Yes	No	Yes	Yes	Yes
7	Yes	Yes	Yes	Yes	Yes	Yes
8	Yes	No	No	Yes	Yes	Yes
9	Partial	Yes	No	Partial	Yes	Yes
10	Yes	Yes	Yes	Yes	No	Yes
11	Partial	Yes	No	No	Yes	No
12	Partial	Yes	No	No	Partial	No
13	Yes	Yes	Yes	Yes	Yes	Yes
14	No	Yes	No	N/A	N/A	No
15	Yes	No	No	Yes	Yes	Yes
16	Yes	Yes	No	Yes	Yes	Yes
17	Yes	Yes	Yes	Yes	No	Yes
18	No	Yes	No	N/A	N/A	No
19	No	Yes	No	N/A	N/A	No
20	Yes	Yes	Yes	No	Yes	Yes
21	Partial	Yes	Yes	Yes	Yes	Yes
22	No	Yes	No	Yes	Yes	Yes
23	Yes	Yes	Yes	No	Yes	No
24	Yes	Yes	Yes	Yes	No	No
25	Yes	Yes	Yes	Yes	No	Yes
26	Yes	Yes	No	Yes	No	Yes

TABLE 4.3: Quality of image set

Bruise 5, 9 and 11 got partial scores for criterion 1: *Include paper and ruler in image*, because measuring tape was used instead of a ruler. Since measuring tape is not rigid, it is less suitable for bruise size extraction. (See Section 2.2.2 In the case of bruise 21, an image of a ruler had apparently been printed on the white paper. In this case, it is uncertain whether the scale of the printed ruler matches real dimensions.

FIGURE 4.9: An image of a ruler printed on paper. (Bruise 21)

Bruise 5 got a partial score for criterion 4: *Bruise and ruler in focus*, because the ruler is in focus but not the bruise. Bruise 9 and 12 got a partial score for criterion 4: *Bruise and ruler in focus*, since the ruler was only in focus in some of the images. Figure 4.10 shows an example of a ruler out of focus. In such cases, extracting the size of the bruise is impossible.

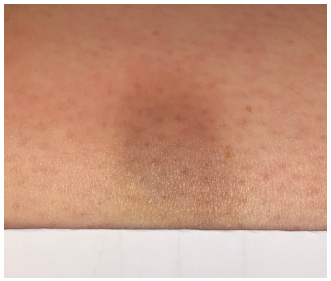
(A) Bruise 2: unfocused ruler (B) Bruise 5: unfocused bruise

FIGURE 4.10: Two examples of poor focus.

Bruises 3, 5 and 8 are examples of bruises where the images have poor lighting. Figure 4.11 shows that it is difficult to separate the darkness caused by the shadow and the darkness caused by the discoloration of the bruise itself, making segmentation of such bruises challenging. In the case of bruise 8, the flash photography has caused significant shine on the skin.

(A) Bruise 3

(B) Bruise 5



(C) Bruise 8

FIGURE 4.11: Three examples of poor lighting.

In the images of bruise 4, both the bruise and the ruler are at an angle from the image plane, distorting the shape of the bruise. This is shown in Figure 4.12.

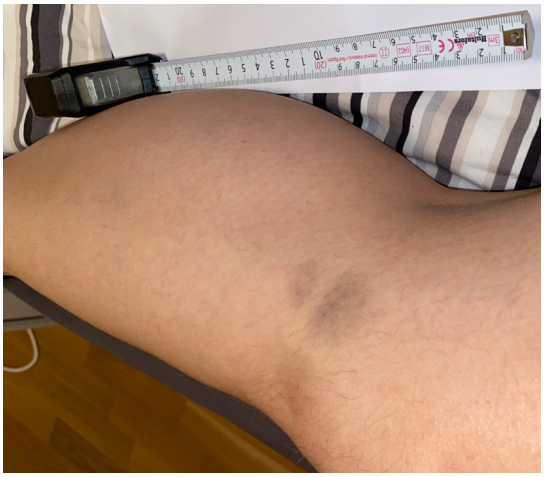


FIGURE 4.12: The bruise and ruler are not parallel with the image plane. (Bruise 4)

The owner of bruise 2 submitted two duplicate images.

4.2 ROI extraction

Region of interest (ROI) was extracted for all bruise images, following the procedure descried in Section 3.2. In some cases, it was impossible to only have bruise and surrounding skin within the ROI. Two such cases are shown in figure 4.13. In these case, it is impossible to extract a rectangular ROI that includes the entire bruise without also including parts of the background.

(A) Bruise 25 (B) Bruise 15

FIGURE 4.13: In these cases, it is impossible to extract a rectangular ROI that contains the entire bruise without also including parts of the background.

In such cases, the three segmentation methods that rely on a rectangular ROI will likely perform worse.

4.3 Size estimation

Two size measures were selected: bruise area and bruise length and width. Before extracting the size of the bruise, ruler detection is performed.

4.3.1 Ruler detection

The automatic ruler detection failed on some images due to the acquisition process and not due to an inherent failure of the method. In the case of some bruises (bruise 2, 11, 20, 23), the ruler was so blurry that it was impossible to determine the spacing between the ruler ticks. In four cases (bruise 14, 18, 19, 21) there was no ruler included in the images, making ruler detection impossible. These bruises are classified as "impossible" bruises in Figure 4.14.

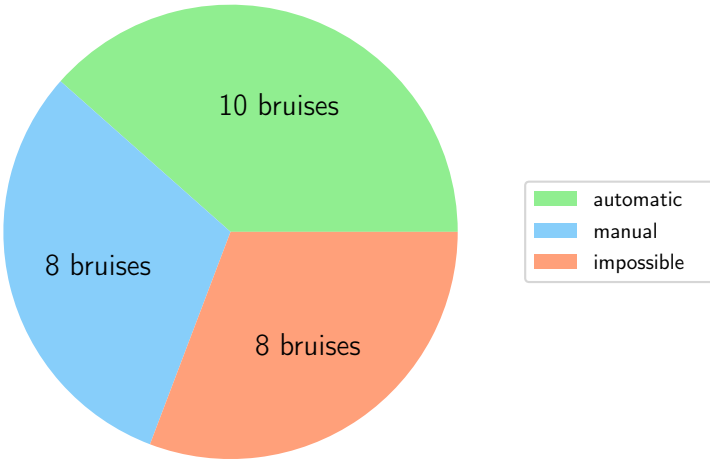


FIGURE 4.14: Performance of ruler detection.

In some cases, the automatic ruler detection failed due to errors that would not prevent a human from correctly determining the ruler spacing. These are marked as "manual" in Figure 4.14, and manual ruler detection was done. In the case of bruise 9, half of the images had a blurry ruler and were thus impossible to perform ruler detection on. In the case of bruises 7, 13 and 24, the ruler included both imperial and metric units. This is a likely source of confusion for the line extraction method. Some bruises had other lined objects in the images, which may have been mistakenly identified as the ruler. The images of bruise 4 were taken on a striped bed-sheet, and the images of bruise 5 included a lined white paper.

Figure 4.15 shows the line detection being confused by the lined paper.

FIGURE 4.15: Lined paper being a source of confusion for the ruler detection. The detected lines in the image are highlighted. (Bruise 5)

Once ruler detection had been performed on all possible images, except the impossible ones, the extracted values of *pixel per millimeter* is compared to the ground truth. The ground truth was defined by measuring the ruler tick separation manually in GIMP [130]. In the images where the ruler is not parallel to the image plane, a minimum and maximum value for the number of pixels per millimeter was extracted.

The following table shows the average error (in pixels) for each bruise.

Bruise	Average error
1	0.5
3	0.8
4	0
5	0
6	0
7	2.75
8	0
9	0.6
10	0
12	0
13	3.2
15	0.25
16	0
17	1.25
22	0
24	4
25	0
26	1.88

TABLE 4.4: Average error, in pixels, of pixel per millimeter extraction.

The average error is quite low, all the values being between 0 and 3.2 pixels.

4.3.2 Area

The bruise area was estimated for each bruise image and segmentation method. The average bruise area was computed by calculating the mean of all area estimates for the same bruise and the same segmentation method. The results are shown in Table 4.5.

TABLE 4.5: Average area (mm^2) for each segmentation method.

BID	OT	KM	KH	AC
1	1902.15	1006.69	640.21	743.63
3	2129.17	1203.15	1304.86	677.97
4	4392.31	1234.35	1901.18	1648.69
5	11495.15	3823.85	4446.50	1815.59
5.2	7425.00	2480.07	3047.64	1321.84
6	3533.60	1268.29	1099.38	1020.21
7	501.82	227.85	306.82	231.75
8	948.36	373.39	237.28	125.15
9	1417.69	753.24	847.58	212.67
10	5935.48	1521.37	2797.38	935.15
12	362.13	159.17	231.82	42.97
13	7312.37	2348.60	2927.04	3380.80
15	14656.20	4955.78	5972.20	4514.92
16	1186.00	680.51	526.33	297.58
17	1225.39	924.55	611.12	377.92
22	8457.39	2829.86	3401.35	2521.98
24	519.96	478.39	306.98;	141.08
25	21000.81	8545.52	9722.26	N/A
26	2776.55	925.76	1087.32	605.48
26.2	2811.95	917.47	1086.04	820.72

Active contours was not done for bruise 25, which is why the estimate is *N/A*. The reason for this is covered in Section 4.6.3.

Defining a ground truth area for each bruise would be subjective and very time-consuming, so it has not been done. Thus the accuracy of these estimates cannot be evaluated. Still, one can analyze the trends. Otsu thresholding produced the largest average area estimate for every single bruise in the data set.

(A) The region of interest (B) Otsu thresholding

FIGURE 4.16: An example of a too large bruise area produced by Otsu thresholding. (Bruise 5)

In this case, the segmentation method identified the shadow and not the bruise. Since the segmentation methods assume that the bruise is in the middle of the region of interest, the white part of Figure 4.16 (B) is classified as the bruise. The area of this segment is significantly larger than the bruise itself. This is a general problem for Otsu thresholding, and will be discussed further in Section 4.6.3.

The precision of the area estimates can be evaluated by computing the standard deviation of the bruise area estimates for each bruise and segmentation method. The results are shown in Table 4.6.

TABLE 4.6: Standard deviation of area (mm^2) for each segmentation method. The lowest standard deviation for each bruise is marked with green.

BID	OT	KM	KH	AC
1	926.68	328.07	307.88	776.99
3	1696.30	930.78	1038.66	420.16
4	1216.86	264.82	476.99	307.21
5	3980.21	826.14	537.70	883.76
5.2	1880.39	396.56	1102.31	655.33
6	970.94	68.24	472.65	121.25
7	475.63	244.23	429.89	171.88
8	15.86	68.16	8.82	14.11
9	249.56	92.86	58.51	18.66
10	829.70	307.15	948.12	145.70
12	485.17	213.75	314.36	60.05
13	5765.97	1454.19	1896.07	2265.07
15	5704.02	1738.90	3941.24	968.20
16	492.14	328.28	140.90	209.88
22	3735.45	1145.55	2402.48	476.66
24	27.2	18.4	12.8	83.77
25	5832.01	2666.71	4415.01	N/A
26	127.74	118.60	128.14	23.78
26.2	1266.49	351.00	433.62	395.26

Otsu thresholding did not result in the lowest standard deviation for any of the bruises. It makes sense that Otsu thresholding caused high values of standard deviation, since the area estimates were also largest for Otsu thresholding.

The other three segmentation methods have equal performances in terms of standard deviation.

4.3.3 Length and width

The ground truth width and length was extracted for one flash image and one image without flash for each bruise. This was done to save time, as manually measuring the length and width of the bruises in all images in the data set would be too time-consuming. The average of the two ground truth measurements is defined as the ground truth. Any difference between the two ground truth estimates indicates possible ambiguity in defining the length and width of the bruise, as well as possible variation of the camera angle. It is also likely that the change in illumination (flash no flash) modifies the apparent size of the bruise by changing which parts of it is visible in the image. The ground truth lengths and widths are presented in Table 4.7.

TABLE 4.7: Ground truth lengths and widths in millimeters, to one decimal point precision.

BID	Length 1	Length 2	Width 1	Width 2	Length	Width
1	27.4	27.5	23.3	24.1	27.5	23.7
3	39.0	41.8	14.7	16.4	40.4	15.6
4	88.0	89.3	53.6	48.1	88.6	50.9
5	117.5	118.6	67.4	62.1	118.1	64.8
5.2	120.9	129.5	58.1	52.5	125.2	55.3
6	98.1	97.4	24.1	24.2	97.8	24.1
7	21.2	25.5	8.1	9.4	23.3	8.7
8	31.6	31.5	21.9	20.4	31.6	21.2
9	32.4	32.4	21.8	18.7	32.4	20.3
10	100.8	101.1	35.6	34.2	100.9	34.9
12	35.9	35.6	21.2	16	35.8	18.6
13	99.5	98.0	39.9	37.6	98.8	38.8
15	80.0	90.8	52.2	49.4	85.4	50.8
16	42.8	37.7	24.0	32.1	40.3	23.5
17	54.5	58.6	36.6	42.3	56.5	39.4
22	149.6	154.8	64.1	64.9	152.2	64.5
24	39.2	26.3	16.7	18.0	37.7	17.3
25	215.1	247.6	72.9	76.1	231.3	74.5
26	109.5	110.2	54.5	47.1	109.9	50.8
26.2	103.6	103.7	48	44.8	103.7	46.4

Some bruises, specifically "blob"-like bruises, have shapes that are fairly straightforward to analyze. One example is bruise 6, shown in Figure 4.17. The manually extracted width and length are marked by red lines.

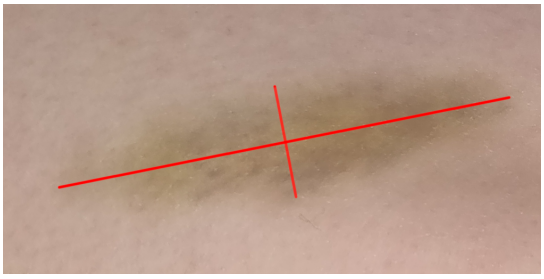


FIGURE 4.17: Manually extracted width and length. (Bruise 6)

Some bruises have a shape that makes it difficult to define an unambiguous width and length. For example, bruise 7, shown in Figure 4.18, has at least two viable candidates for width.

FIGURE 4.18: The width of some bruises is ambiguous. (Bruise 7)

In such cases, an average width between the two candidates was calculated. In some cases, the bruise is so diffuse that it is challenging to determine where it begins and ends. Two such bruises are shown in Figure 4.19.

(A) Bruise 3 (B) Bruise 15

FIGURE 4.19: It is difficult to define the border of diffuse bruises.

The average estimated width and length for each bruise and segmentation method was computed. The four segmentation methods are abbreviated as follows:

- Otsu = OT
- K means = KM
- K means with histogram equalization = KH
- Active contours = AC

Further, ground truth is abbreviated as GT. The estimated length and width, as well as the ground truth, for each bruise are presented in tables 4.8 and 4.9. The most accurate estimate is highlighted for each bruise.

TABLE 4.8: Estimated length for each segmentation method.

BID	OT	KM	KH	AC	GT
1	71.7	45.5	33.9	43.2	27.5
3	72.7	46.4	68.3	64.0	40.4
4	111.0	57.0	89.5	101.9	88.6
5	134.4	125.5	146.8	75.6	118.1
5.2	144.1	104.3	113.8	84.1	125.2
6	121.3	55.3	77.8	93.8	97.8
7	37.6	20.9	29.9	34.7	23.3
8	43.4	25.1	22.6	28.0	31.6
9	53.62	34.7	55.5	38.6	32.4
10	117.5	65.1	93.3	76.9	100.9
12	25.06	12.1	17.3	12.5	35.8
13	161.5	67.5	133.0	144.2	98.8
15	86.3	57.5	79.2	82.3	85.4
16	50.4	38.0	38.5	42.1	40.3
17	3.5	40.2	42.8	48.5	56.5
22	164.3	90.3	106.2	136.2	152.2
24	48.1	34.8	31.0	29.1	37.7
25	253.4	138.7	201.8	N/A	231.3
26	91.5	45.3	66.2	59.3	109.9
26.2	102.9	40.4	54.1	74.6	103.7

TABLE 4.9: Estimated width for each segmentation method.

BID	OT	KM	KH	AC	GT
1	48.1	30.3	23.0	42.3	23.7
3	42.7	33.6	24.1	61.6	15.6
4	51.1	30.9	40.4	98.9	50.9
5	86.7	19.40	38.9	70.1	64.8
5.2	119.5	55.8	57.9	75.0	55.3
6	95.7	31.7	31.8	84.4	24.1
7	23.1	12.1	7.5	32.0	8.7
8	37.6	19.3	10.8	27.4	21.2
9	50.6	32.6	36.7	38.0	20.3
10	61.8	33.6	43.1	75.6	34.9
12	20.2	9.6	7.6	10.9	18.6
13	63.9	46.9	38.8	128.5	38.8
15	49.4	30.3	26.8	78.5	50.8
16	37.3	21.5	14.0	41.4	23.5
17	49.3	30.5	14.0	46.4	39.4
22	65.9	37.2	36.7	115.2	64.5
24	27.2	18.4	12.8	27.5	17.3
25	230.4	84.5	84.5	N/A	74.5
26	75.8	36.7	26.6	47.4	50.8
26.2	91.8	27.8	25.4	69.0	46.4

Otsu thresholding resulted in a too large width estimate in all bruises except one and a too large length estimates in all bruises except five. It makes sense that

the length and width estimates were often too large, as the area estimates were also consistently larger than the other segmentation methods.

Similarly, active contours over-estimated the width in all bruises except two. However, the length estimates were more evenly spaced around the ground truth. In other words, active contours seems to over-estimate the roundness of the bruise. This will be further addressed in Section 4.4.

The estimates from k-means clustering with and without histogram equalization were more evenly spaced around the ground truth.

In the case of bruise 11, two images with no flash were used to find the ground truth, as the ruler was out of focus in the flash images. Two flash images were used for bruises 9 and 12, for the same reason.

4.4 Shape estimation

The average bruise roundness and standard deviation were computed for each bruise and segmentation method. The average estimated roundness of each bruise is presented in Table 4.10. Roundness ranges from 1 (width equals length) to 0 (length infinitely larger than width). There is no defined ground truth of each bruise, so the accuracy of these estimates cannot be computed exactly.

TABLE 4.10: Average roundness for each bruise and segmentation method.

BID	OT	KM	KH	AC
1	0.67	0.67	0.73	0.98
2	0.33	0.72	0.24	0.83
3	0.62	0.73	0.39	0.96
4	0.46	0.55	0.46	0.97
5	0.63	0.50	0.26	0.93
5.2	0.84	0.51	0.49	0.89
6	0.79	0.67	0.39	0.90
7	0.72	0.56	0.29	0.94
8	0.86	0.77	0.56	0.98
9	0.95	0.94	0.65	0.98
10	0.52	0.52	0.48	0.99
11	0.87	0.78	0.53	0.59
12	0.59	0.74	0.49	0.94
13	0.38	0.71	0.26	0.89
14	0.64	0.65	0.53	0.80
15	0.59	0.55	0.33	0.95
16	0.74	0.56	0.39	0.98
17	0.75	0.76	0.36	0.95
18	0.73	0.81	0.67	0.93
19	0.63	0.65	0.63	0.79
20	0.53	0.53	0.41	0.98
20.2	0.54	0.66	0.43	0.97
21	0.63	0.65	0.63	0.79
21	0.68	0.72	0.51	0.96
22	0.40	0.44	0.35	0.85
23	0.77	0.60	0.40	0.93
24	0.63	0.53	0.44	0.95
25	0.91	0.62	0.69	N/A
26	0.83	0.81	0.42	0.80
26.2	0.90	0.69	0.48	0.92

Three pairs of comparison bruises are chosen to sample the accuracy of the roundness estimation and provide an estimate of the accuracy. In each pair, the first bruise is significantly rounder than the second. This roundness labeling is done by manually observing the bruise. The three comparison pairs are as follows: bruise 8 and bruise 6; bruise 18 and bruise 7; and bruise 19 and bruise 23. It is expected that the estimated roundness should be higher for the first bruise in each pair.

For the first pair and second pair, the only segmentation method that failed to correctly identify the first bruise as the roundest was active contours. For the third pair, both active contours and Otsu thresholding failed. Based on these samples, k-means clustering with and without histogram equalization scored best in roundness estimation accuracy. Active contours scored the worst.

Looking at Table 4.10, active contours consistently provides the highest roundness estimate. This is related to the finding in Section 4.3.3 that active contours overestimates the width of bruises. This bias toward round bruise shapes may be a result of the initial curve being a circle. Future work should investigate the effect of changing this initial curve.

The standard deviation of the bruise roundness is a measure of shape estimation precision. Low standard deviation indicates high precision. The standard deviation of the bruise roundness for each bruise and segmentation method is shown in Table 4.11. The lowest standard deviation for each bruise is marked with bold text.

TABLE 4.11: Standard deviation of roundness for each bruise and segmentation method. The lowest standard deviation for each bruise is highlighted.

BID	OT	KM	KH	AC
1	0.08	0.15	0.20	0.02
2	0.07	0.10	0.12	0.16
3	0.14	0.15	0.23	0.02
4	0.22	0.10	0.11	0.03
5	0.26	0.28	0.10	0.03
5.2	0.17	0.20	0.22	0.03
6	0.13	0.38	0.27	0.06
7	0.36	0.16	0.10	0.06
8	0.13	0.04	0.43	0.02
9	0.06	0.04	0.12	0.01
10	0.04	0.12	0.16	0.02
11	0.20	0.05	0.24	0.17
12	0.25	0.11	0.13	0.11
13	0.05	0.10	0.17	0.10
14	0.08	0.18	0.16	0.16
15	0.14	0.25	0.15	0.04
16	0.19	0.18	0.27	0.01
17	0.30	0.07	0.31	0.06
18	0.04	0.17	0.19	0.08
19	0.05	0.06	0.06	0.20
20	0.08	0.07	0.10	0.01
20.2	0.08	0.28	0.24	0.03
21	0.04	0.19	0.05	0.02
22	0.09	0.32	0.16	0.13
23	0.17	0.07	0.16	0.03
24	0.30	0.04	0.12	0.05
25	0.14	0.16	0.14	N/A
26	0.20	0.12	0.23	0.06
26.2	0.08	0.11	0.03	0.07

Active contours has the lowest standard deviation on average. This may be caused by the methods being biased toward round bruise shapes and thus not varying much.

4.5 Color estimation

4.5.1 Bruise contrast

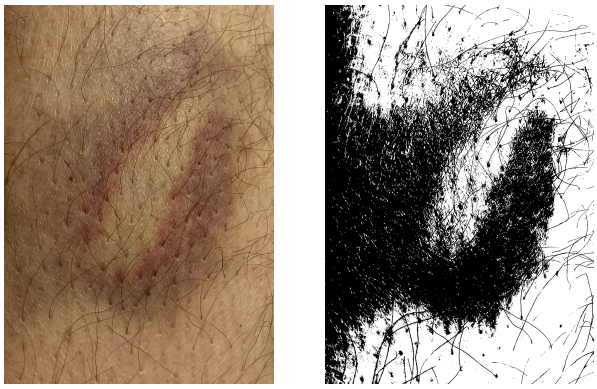
The bruise contrast was computed for each bruise image and each segmentation variation. Bruise contrast is defined as the difference between the luminance of the skin L_s and the luminance of the bruise L_b. See Section 3.6. The average bruise contrast is shown in Table 4.12.

TABLE 4.12: Average bruise contrast

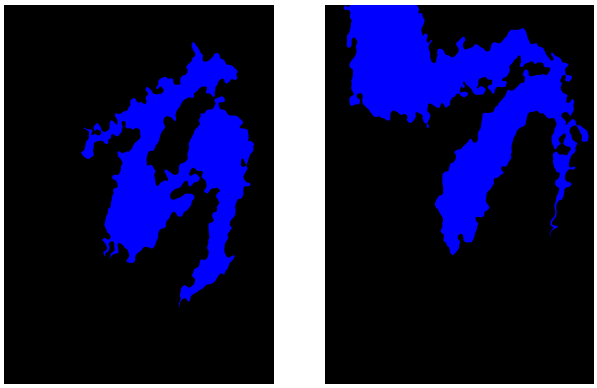
BID	OT	KM	KH	AC
1	14.36	13.80	14.42	-0.11
2	0.99	-0.06	-0.60	0.78
3	-9.77	-4.40	-5.43	0.06
4	-2.87	-0.12	-1.17	0.71
5	5.46	5.54	8.89	0.32
5.2	7.50	0.22	2.90	0.19
6	8.25	5.92	6.22	0.88
7	11.71	9.24	11.22	-0.72
8	14.25	9.80	8.03	-0.32
9	-3.14	-3.18	1.78	0.00
10	-14.51	-1.61	-0.27	0.03
11	-13.72	-6.90	-5.54	-0.12
12	-0.43	1.56	-0.39	0.23
13	-4.25	-3.73	-1.29	0.05
14	-6.64	0.29	-6.93	-0.89
15	-21.44	-8.28	-4.99	-0.46
16	20.33	5.19	3.05	-0.89
17	13.49	5.15	8.40	0.02
18	11.24	7.45	6.15	0.98
19	8.95	6.28	6.14	0.10
20	-22.65	-3.32	-7.41	3.46
20.2	-20.11	-8.52	-9.96	6.87
21	5.43	3.29	3.19	0.39
22	15.81	-1.69	0.61	-0.27
23	0.03	0.06	-1.42	-3.00
24	11.36	10.38	12.36	0.00
25	0.07	1.48	7.24	N/A
26	11.22	6.19	3.87	0.44
26.2	-8.38	-1.01	-0.95	0.13

Positive values are excepted, since normally bruises are darker than the surrounding skin. Sometimes this will not be the case in an image, due to the presence of shadows and other confounding factors.

In the case of bruise 20, a central clearing bruise, all segmentation methods except active contours resulted in significantly negative contrast values. Figure 4.20 shows bruise 20 segmented by three of the four segmentation methods: Otsu thresholding (OT), k-means clustering (KM) and k-means clustering with histogram equalization (KH).



(A) The region of interest. (B) OT

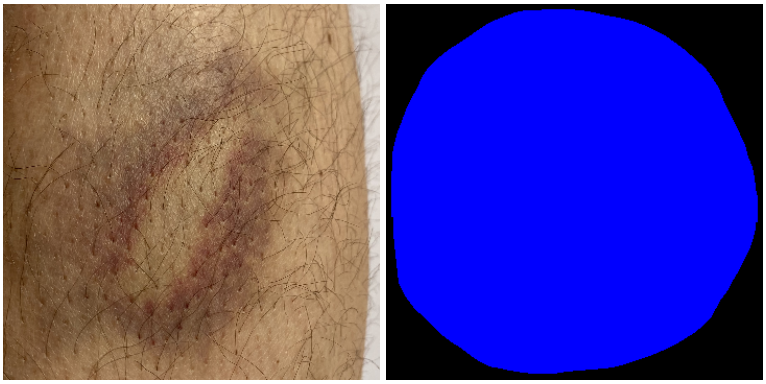


(C) KM (D) KH

FIGURE 4.20: Segmentation results for k-means and Otsu thresholding. (Bruise 20)

It is clear why these segmentation methods resulted in negative contrast values. Since the central clearing is in the middle of the region of interest, and the central clearing has a similar color to the surrounding skin, the central clearing and the skin was classified as the bruise. Since the skin and central clearing are lighter than the bruise, but the classes were switched, the resulting contrast values are negative.

Figure 4.21 shows the segmentation results for the same bruise with active contours.



(A) Bruise 6 (B) Active contours

FIGURE 4.21: Segmentation with active contours. (Bruise 20)

The segmentation result produced by active contours does not correspond to the bruise, but it still results in better contrast values than from the other three segmentation methods. This is by chance, since the circle in Figure 4.21b happens to include more of the dark coloration than the rest of the image. As discussed in Section 4.4, active contours defaults to a circle-like bruise segments in cases where the edge of the bruise is not detected.

TABLE 4.13: Standard deviation of bruise contrast. The lowest standard deviation for each bruise is highlighted green.

BID	OT	KM	KH	AC
1	1.96	1.57	1.14	0.10
2	5.17	1.90	0.61	0.21
3	14.29	5.57	5.64	0.27
4	14.19	2.24	5.51	1.64
5	23.82	10.87	9.20	0.48
5.2	26.09	8.06	10.39	0.32
6	0.85	3.45	4.85	0.73
7	1.91	1.49	2.55	1.15
8	3.16	4.02	0.61	1.30
9	10.06	3.17	5.57	0.05
10	2.23	4.55	3.49	0.13
11	3.45	6.35	2.59	0.27
12	10.19	3.17	2.28	1.90
13	7.05	2.39	1.01	0.12
14	20.13	6.33	9.13	3.77
15	2.80	4.41	2.18	0.16
16	6.56	7.58	3.18	3.91
17	1.44	7.33	4.49	0.55
18	3.53	5.49	2.55	1.87
19	0.63	1.13	0.46	0.32
20	4.37	7.50	2.78	1.30
20.2	4.14	7.76	4.14	3.98
21	0.36	0.50	1.08	0.44
22	26.93	4.77	4.74	2.98
23	9.68	3.23	6.18	5.31
24	1.41	2.38	0.95	0.45
25	25.94	4.09	11.75	N/A
26	4.30	3.24	2.43	0.68
26.2	0.31	6.28	7.46	0.49

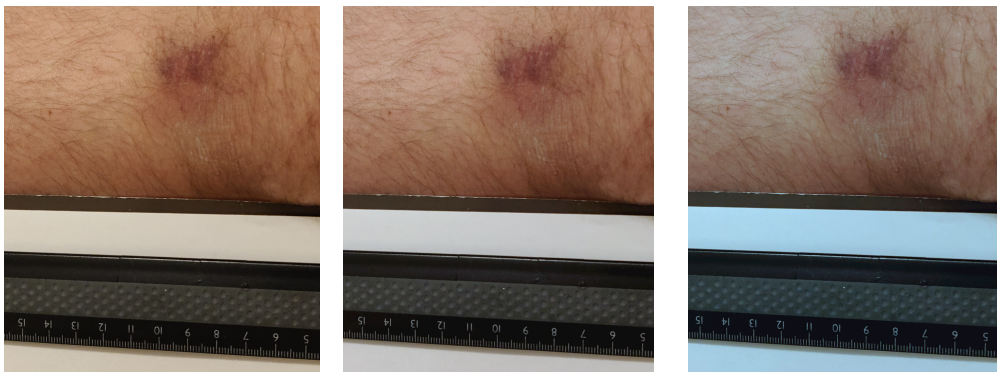
Lowest standard deviation in active contours. This is related to the low standard deviation values for roundness estimation. The segmentation done by active contours seems to vary less between repetitions than the other segmentation methods.

Otsu seems to have the highest standard deviation. This can be related to the bruise segment variation being the largest in Otsu thresholding.

4.5.2 Computational color constancy

White patch (WP) and gray world (GW) color correction was applied to all the bruise images.

Figure 4.22 shows a no flash image of bruise 1 with and without computational color constancy.



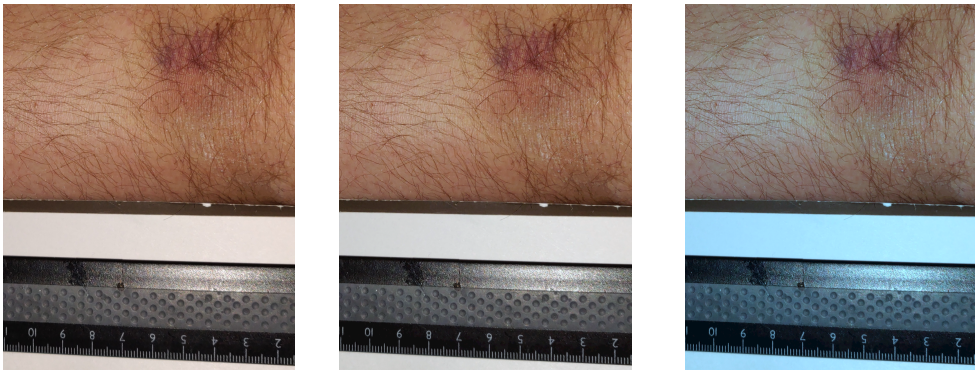
(A) Bruise image with no computational color constancy. (B) Bruise image corrected with white patch (WP). (C) Bruise image corrected with gray world (GW).

FIGURE 4.22: A no flash image of bruise 1, with and without computational color constancy.

Upon visual inspection, the white paper seems the whitest in the image corrected with white patch. This makes sense, since the white patch algorithm is based on identifying the whitest parts of the image and making them white. Using a color picker, the colors of the white paper are as follows:

- No computational color constancy: (205, 196, 181)
- White patch:(199, 200, 201)
- gray world: (165, 206, 230)

From this, it is clear that the paper is the whitest in the image corrected with white patch, since there is the least variation in the RGB values. In the case of gray world, the white paper appears bluish, which is also apparent in the third RGB value being largest. Figure 4.23 shows another example, the same bruise but this time with flash.



(A) Bruise image with no computational color constancy. (B) Bruise image corrected with white patch (WP). (C) Bruise image corrected with gray world (GW).

FIGURE 4.23: A flash image of bruise 1, with and without computational color constancy.

Again, the white paper appears bluish in the image corrected with gray world. However, the difference between the image without computational color constancy and the image corrected with white patch is less noticeable here.

The color of the white paper was manually sampled in one image from each bruise. The results are shown in Table 4.14. No white paper was included in the images of bruises 14, 18, 19, 22, so they have been omitted. The computational color constancy method that resulted in the whitest paper is highlighted for each bruise.

TABLE 4.14: Comparing the color of the white paper (R,G,B) for white patch (WP) and gray world (GW)

BID	WP	GW
1	(199,200,201)	(165,206,230)
2	(208,225,221)	(187,226,255)
3	(218,224,222)	(173,233,255)
4	(202,222,222)	(182,217,253)
5	(220,227,227)	(203,236,250)
5.2	(216,215,213)	(184,216,232)
6	(169,191,198)	(166,196,218)
7	(233,230,229)	(217,235,227)
8	(228,233,237)	(187,244,255)
9	(211,215,223)	(148,231,255)
10	(153,156,163)	(136,151,166)
11	(197,197,199)	(165,203,248)
12	(219,218,219)	(193,215,235)
14	(213,211,214)	(184,219,244)
15	(196,193,192)	(188,198,207)
16	(206,218,220)	(196,220,254)
17	(250,248,246)	(164,255,255)
20	(219,220,221)	(186,209,236)
20.2	(221,227,227)	(201,226,246)
21	(164,168,170)	(151,168,187)
23	(149,156,155)	(127,147,177)
24	(159,173,177)	(161,182,197)
25	(255,254,251)	(228,255,255)
26	(213,225,226)	(190,233,252)
26.2	(219,224,223)	(200,225,246)

White patch outperformed gray world in all the images that were sampled. In all cases except bruise 7, the gray world images are bluish (the B-component is larger than the R and G-components of the color). This effect may be caused by the fact that the average color of the image tends toward yellow/orange due to the abundance of skin color in them. According to the gray world assumption, this deviation from average achromaticity is caused by yellow/orangeish light, which is why the correction creates a bluish tint on the images. In the case of bruise 7, about a fifth of the images consists of a blue shirt worn by the participant. In this case, the image was not yellow/orange on average, and the gray world algorithm did not apply a bluish tint.

4.5.3 Average bruise color

The average bruise color difference between flash and no flash images is given in Table 4.15. This is a measure of the robustness of color estimation with respect to changing illumination. Each column refers to a combination of segmentation method (OT = Otsu, KM = K means, KH = K means with histogram equalization, AC = active contours) and computational color constancy method (WP = white patch, GW = gray world). The lowest average color difference for each bruise is highlighted.

TABLE 4.15: Color robustness: Average bruise color difference between flash and no flash for white patch (WP) and gray world (GW).

BID	WP-OT	GW-OT	WP-KM	GW-KM	WP-KH	GW-KH	WP-AC	GW-AC
1	69.24	74.18	76.56	80.88	90.82	93.20	273.55	279.53
2	27.79	29.43	23.93	27.43	29.83	33.58	45.40	44.91
3	39.78	41.16	64.36	66.78	61.83	64.22	6.24	1.51
4	36.81	38.53	25.70	27.17	41.42	43.73	45.98	41.66
5	84.45	85.89	86.16	87.35	82.23	83.35	62.65	61.90
5.2	67.66	69.25	55.32	56.98	91.23	92.73	67.80	66.00
6	33.37	34.46	49.24	50.94	52.76	54.70	26.89	25.54
7	28.37	26.38	23.79	21.70	29.75	27.81	118.19	219.51
8	55.03	52.41	71.31	68.87	43.03	41.14	196.90	194.38
9	14.79	16.21	25.16	25.76	43.92	45.48	3.82	3.77
10	25.27	34.42	38.39	37.51	46.26	45.45	96.41	95.88
11	29.65	11.59	48.45	37.98	28.91	11.64	69.95	68.57
12	31.90	18.21	29.25	9.38	29.71	10.29	132.17	133.94
13	37.13	42.40	45.54	50.49	65.58	69.78	23.99	29.64
14	88.76	86.34	66.51	63.98	85.88	83.37	45.39	47.32
15	18.95	18.80	7.73	8.00	27.99	27.79	65.46	66.02
16	3.32	1.95	9.28	5.09	18.03	20.79	225.03	229.57
17	21.13	13.75	52.48	51.04	41.29	38.75	41.60	36.11
18	107.40	101.16	123.88	117.72	140.71	134.89	65.07	67.60
19	25.78	25.13	17.82	16.44	19.02	17.94	25.09	26.17
20	65.98	67.44	98.19	100.37	67.71	69.33	58.38	59.41
20.2	75.63	76.31	113.79	115.28	83.62	84.45	16.96	13.73
21	35.64	34.96	30.10	29.14	27.59	26.64	131.50	133.12
22	36.93	19.73	34.98	6.21	27.46	7.58	46.75	167.10
23	65.50	71.70	38.72	43.72	53.89	59.81	58.70	54.09
24	9.63	14.08	4.38	4.65	8.98	13.31	7.88	9.14
25	13.48	13.75	7.63	10.01	11.05	12.66	N/A	N/A
26	25.05	19.37	23.67	17.94	30.24	24.34	37.38	42.65
26.2	22.96	18.45	66.64	63.68	72.64	69.75	33.05	35.92

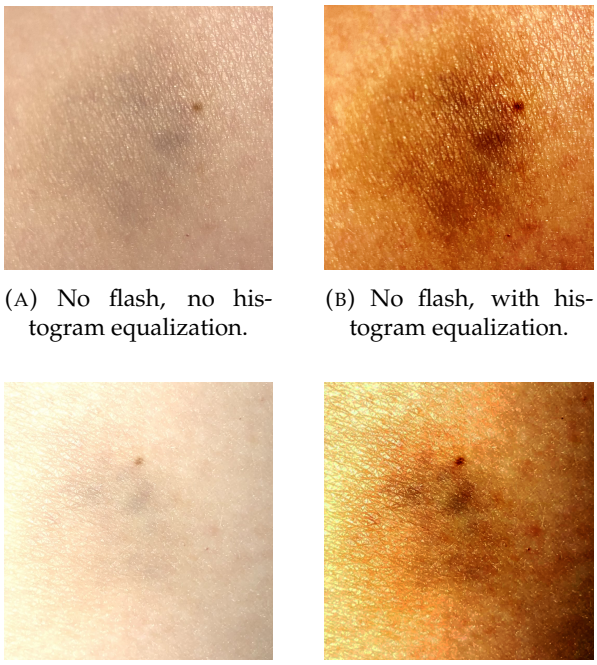
The goal of computational color constancy methods is to correct the unknown illumination to be a canonical light source. If the computational color constancy methods work perfectly as intended, there should be no bruise color difference between the flash and no flash image. There are several reasons why this is not the case. Firstly, the flash and no flash images have been acquired by non-professionals without using a tripod or other means of maintaining the exact same angle and position of the camera. Therefore, the scene will differ by more than just illumination. In addition to this, the segmentation methods may yield different results based on the illumination, which will impact the average color of the bruise segment. Finally, neither white patch nor gray world are perfect methods and their central assumptions may be violated.

White patch and gray world were compared for each bruise and segmentation method to see which has the lowest average bruise color difference. There are 29 bruises (26 bruises + 3 sets of follow up images). Interestingly, for every four of the segmentation methods, white patch performed the best in 15 cases and gray world

performed the best in 14 cases. Based on this, it is impossible to select one of the computational color constancy methods as the best one.

No white paper was included in the images of bruises 14, 18, 19 and 22. Gray world outperformed white patch in all of these bruises. This makes sense, since white patch relies on a white patch (white paper) being present in the image.

Of the four segmentation methods, k-means clustering with histogram equalization performed the worst in terms of color robustness. It is possible that the histogram equalization step had different results for flash and no flash images, causing this discrepancy. In the case of bruise 18, the average bruise color difference between flash and no flash is significantly higher for k-means clustering with histogram equalization than for the other three methods. Figure 4.24 shows the effect of histogram equalization on flash and no flash images of bruise 18.



(A) No flash, no histogram equalization.

(B) No flash, with histogram equalization.

(C) With flash, no histogram equalization.

(D) With flash, with histogram equalization.

FIGURE 4.24: The effect of histogram equalization on flash and no flash images of bruise 18.

For the no flash image, histogram equalization increased the contrast of the bruise. The flash image is over-exposed, causing the bruise to less visible. In this case, histogram equalization did increase the contrast of the bruise, but also heavily increased the contrast of the shadow in the image. This different effect of histogram equalization on flash and no flash images may explain the low color robustness for k-means clustering with histogram equalization.

Table 4.16 shows the standard deviation of the color estimates for each segmentation and computational color constancy method. This is a measure of the precision of the color estimation. The lowest standard deviation for each bruise is highlighted.

TABLE 4.16: Color precision: Standard deviation of bruise color estimate with white patch (WP) and gray world (GW).

BID	WP-OT	GW-OT	WP-KM	GW-KM	WP-KH	GW-KH	WP-AC	GW-AC
1	8.07	8.22	18.94	18.95	20.46	20.69	11.53	11.62
2	11.97	12.54	5.90	6.12	2.37	2.44	2.21	2.69
3	31.66	32.44	15.78	16.28	16.92	17.47	2.46	1.45
4	34.35	23.78	7.17	7.83	9.52	10.06	1.59	1.88
5	23.36	24.11	18.71	19.32	19.93	20.54	18.19	17.24
5.2	32.29	32.50	22.54	22.78	22.90	22.96	3.29	3.35
6	2.61	2.54	4.42	4.40	10.74	11.10	32.72	33.46
7	4.39	4.69	5.72	5.61	6.94	7.04	21.13	36.52
8	0.69	0.89	0.67	0.78	0.83	1.16	1.17	1.26
9	8.78	9.26	5.51	5.97	15.72	16.58	0.73	0.72
10	4.20	4.35	13.20	12.93	6.29	6.18	3.04	2.06
11	6.21	6.09	11.82	12.10	9.69	9.59	4.86	5.61
12	17.27	18.21	11.10	11.80	6.14	6.21	26.63	28.82
13	8.53	8.69	5.90	6.02	2.28	2.19	0.93	0.97
14	28.92	29.14	18.33	18.36	19.10	19.07	61.53	61.96
15	1.67	1.55	2.52	2.79	3.23	3.34	0.61	1.29
16	9.88	9.73	22.41	22.63	15.91	16.06	10.41	10.92
17	2.49	3.07	5.28	5.45	7.34	7.41	2.93	4.73
18	14.23	13.86	11.05	10.32	15.65	15.99	5.81	4.64
19	7.85	7.61	5.78	5.70	5.24	5.36	33.99	34.73
20	0.66	0.71	1.55	1.77	2.91	3.14	4.08	4.41
20.2	0.58	0.86	0.69	1.04	0.81	0.97	2.55	2.46
21	1.08	1.12	1.19	1.13	0.50	0.44	1.13	1.15
22	26.61	2.53	8.14	8.13	11.46	5.41	10.30	70.51
23	5.69	5.70	9.51	9.42	10.56	10.59	21.40	21.56
24	7.46	4.90	4.36	1.26	6.40	3.20	12.46	10.19
25	57.95	57.81	15.55	15.76	38.62	38.37	N/A	N/A
26	1.18	0.62	1.42	1.73	2.06	2.21	2.26	2.16
26.2	0.63	0.26	1.22	1.12	1.12	0.97	0.62	0.61

White patch and gray world were compared for each bruise and segmentation method to see which has the lowest standard deviation of bruise color. White patch outperformed gray world in 17 of the 29 cases for Otsu thresholding, 19 cases for k-means without histogram equalization and 20 cases for k-means with histogram equalization and active contours. This indicates that white patch is slightly better at reducing variation in color between image repetitions.

4.6 Segmentation

4.6.1 Execution time

Execution time in seconds was measured for each of the four segmentation methods for a selection of the bruise images.

TABLE 4.17: Image dimensions in pixels and execution time in seconds for a selection of bruise images.

Image height	Image width	OT	KM	KH	AC
2176	4606	0.01	1.79	2.02	66.58
4032	3024	0.003	2.37	2.66	12.39
4608	3456	0.004	2.47	2.90	7.88

Otsu thresholding is the fastest segmentation method of the four selected. This is to be expected, since (as mentioned in Section 2.3) Otsu thresholding is one of the segmentation methods with the lowest time complexity, The execution time difference between k-means with and without histogram equalization is mostly attributed

to the time taken to perform histogram equalization. Active contours may have such a high execution time due to the ray casting algorithm that was implemented by the author and not optimized.

4.6.2 Segmentation method ranking

The four segmentation are ranked by comparing their feature extraction results. This ranking is shown in Table 4.18. The average ranking is computed by weighting the four following categories equally: size estimation, shape estimation, color estimation and execution time. That means that segmentation speed is weighted by a factor of 1, since there is only one ranking in the category of execution time, but area precision is weighted by a factor of 0.33, since there are three rankings in the category of bruise size.

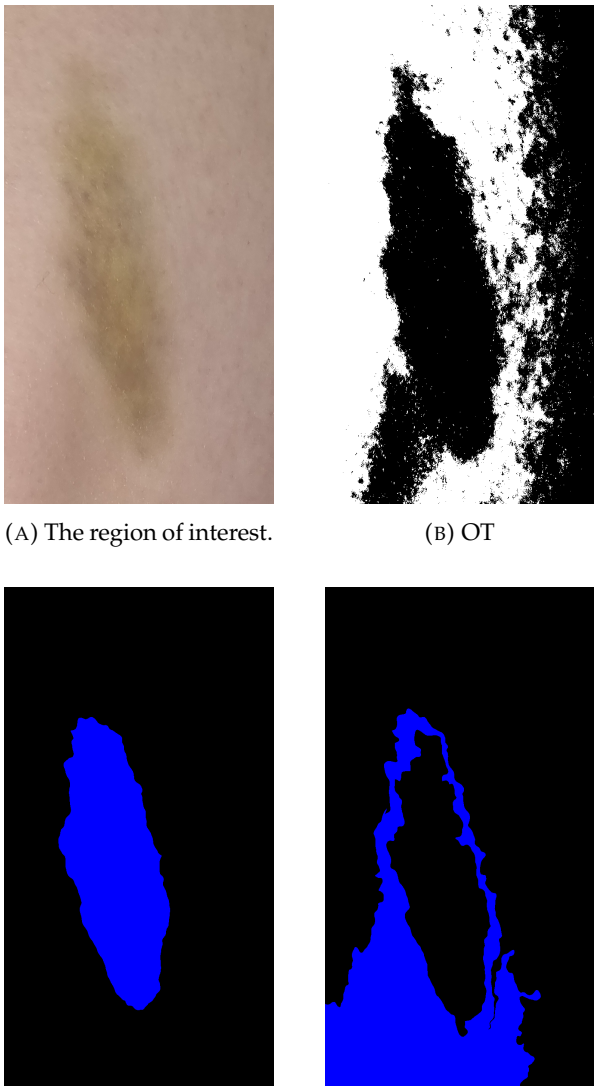
TABLE 4.18: Ranking the segmentation methods according to feature extraction results. 1 is best and 4 is worst.

Parameter	OT	KM	KH	AC
Area precision	4	2	3	1
Length accuracy	1	2	3	4
Width accuracy	3	1	2	4
Roundness precision	2	3	3	1
Roundness accuracy	2	1	1	3
Contrast precision	3	3	2	1
Color robustness	3	1	4	2
Color precision	3	2	4	1
Segmentation speed	1	2	3	4
Average ranking	2.15	1.91	2.74	2.57
Final ranking	2	1	4	3

The final ranking of the segmentation methods, from best to worst, is: k-means clustering without histogram equalization, Otsu thresholding, active contours and k-means clustering with histogram equalization.

4.6.3 Analysis

The segmentation methods were also compared by visual inspection of the bruises and segmentation results. Figure 4.25 shows bruise 6, a blob bruise, segmented by three of the four segmentation methods: Otsu thresholding (OT), k-means clustering (KM) and k-means clustering with histogram equalization (KH).



(A) The region of interest. (B) OT

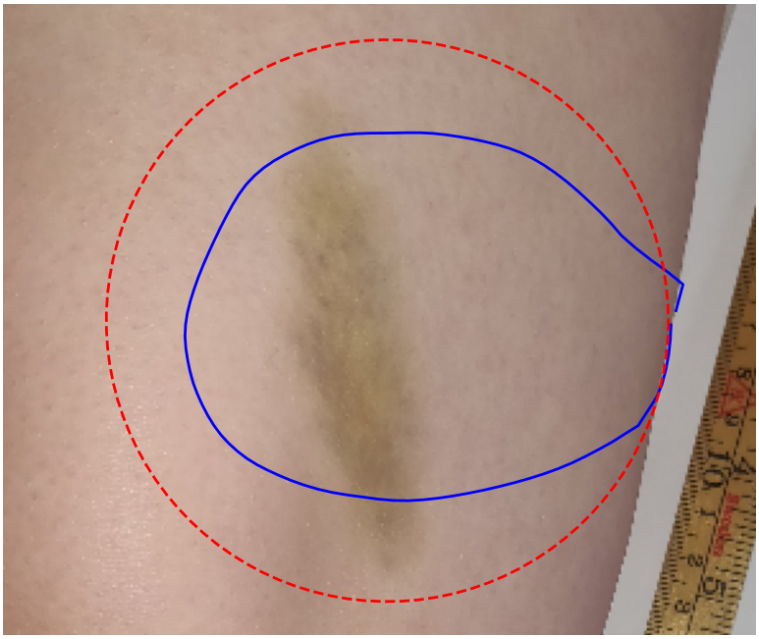
(C) KM (D) KH

FIGURE 4.25: Segmentation results for k-means and Otsu thresholding. (Bruise 6)

Upon visual inspection, k-means clustering was the method among the three that resulted in the segmented region with the most similarity with the bruise. The Otsu thresholding could be improved by using a morphological opening operation to remove the "speckles" around the central segment. The large black area in the right part of the image is likely caused by the shadow, and causes an overestimation of the size of the bruise. Since Otsu thresholding uses the grayscale histogram of the image, and the shadowed skin has a similar graylevel value to the bruise, the shadowed skin is classified as bruise. In the case of the k-means clustering with histogram equalization, the segmentation method seems to have segmented the yellowish border around the bruise and not the central bruise itself.

Figure 4.26 shows the segmentation results for the same bruise with active contours.

(A) Bruise 6 (B) AC

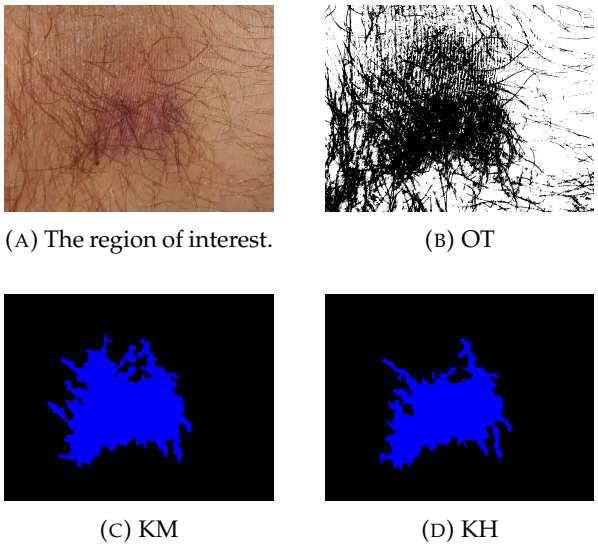


(C) The process.

FIGURE 4.26: Segmentation with active contours. (Bruise 6)

Figure 4.26c shows the process from initial curve (red dashed line) to final snake curve (blue solid line). The final snake curve does not correspond well with the bruise in the image. Looking at the rightmost part of the curve, it seems to follow the edge of the limb. Apparently this edge was the most salient edge within the initial curve, attracting the curve in that direction. Had the initial curve been placed slightly to the left, this may not have happened.

Figure 4.27 shows bruise 1, another blob bruise, segmented by three of the four segmentation methods: Otsu thresholding (OT), k-means clustering (KM) and k-means clustering with histogram equalization (KH).



(A) The region of interest. (B) OT

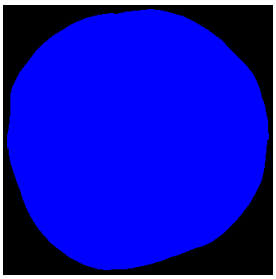
(C) KM (D) KH

FIGURE 4.27: Segmentation results for k-means and Otsu thresholding. (Bruise 1)

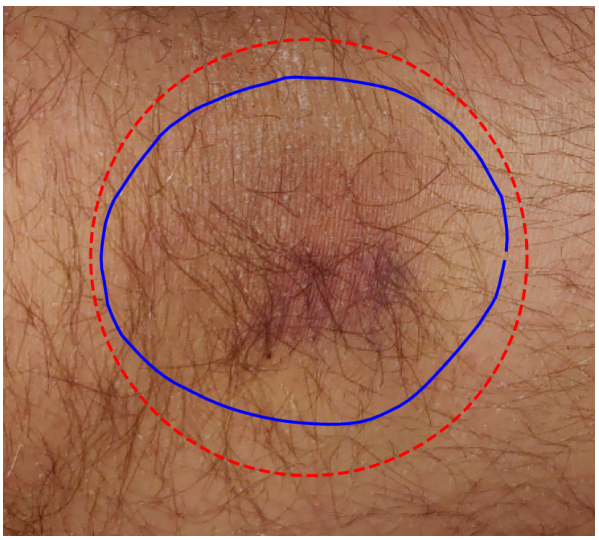
Looking at Figure 4.27b, it is clear that the skin hair of the participant interfered with the segmentation process. Since the skin hair and the bruise have similar gray level values, they are classified as the same class by the thresholding operation. This is less of a problem for k-means clustering, which takes both the color of the pixels (R, G, B) and their location (x, y) into account.

Figure 4.28 shows the segmentation results for the same bruise with active contours.

(A) Bruise 1



(B) Active contours

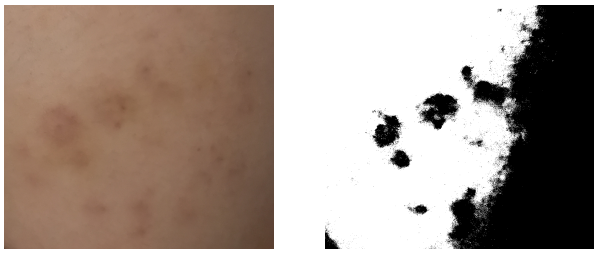


(C) The process.

FIGURE 4.28: Segmentation with active contours. (Bruise 1)

The active contours curve has contracted somewhat from the initial curve, but does not seem to have detected the bruise. The resulting segment is significantly larger than the bruise.

Figure 4.29 shows bruise 11, a cluster bruise, segmented by three of the four segmentation methods: Otsu thresholding (OT), k-means clustering (KM) and k-means clustering with histogram equalization (KH).

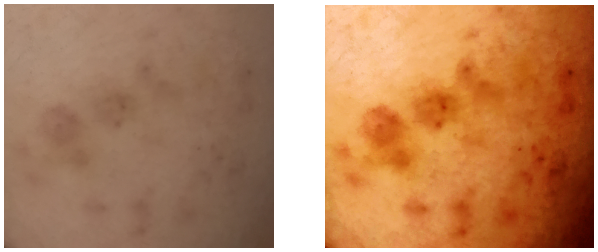


(A) The region of interest. (B) OT

(C) KM (D) KH

FIGURE 4.29: Segmentation results for k-means and Otsu thresholding. (Bruise 11)

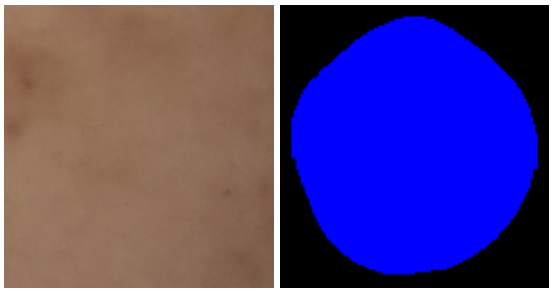
Again, Otsu thresholding suffers from the presence of shadows. It correctly identifies some parts of the bruise, but the entire right side of the image is mistakenly classified as a bruise due to the shadow there. K-means clustering has completely failed to identify the shape of the bruise. K-means clustering with histogram equalization, however, corresponds somewhat with the shape of the bruise. Histogram equalization increases the contrast of the image, which may have improved the recognition of the bruise. The effect of histogram equalization is shown in Figure 4.30.



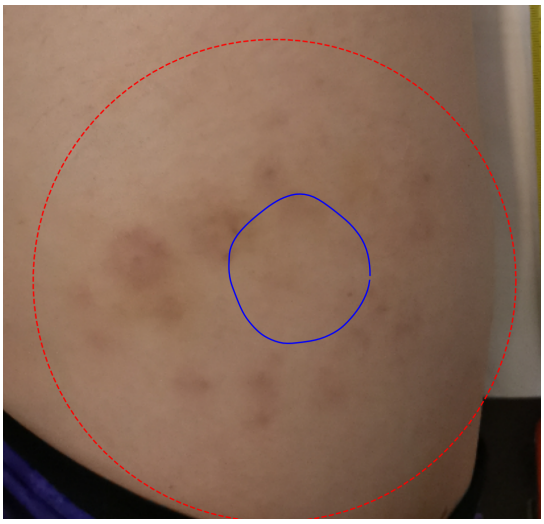
(A) No histogram equalization. (B) With histogram equalization.

FIGURE 4.30: The effect of histogram equalization. (Bruise 11)

Figure 4.35 shows the segmentation results for the same bruise with active contours.



(A) Bruise 6 (B) AC

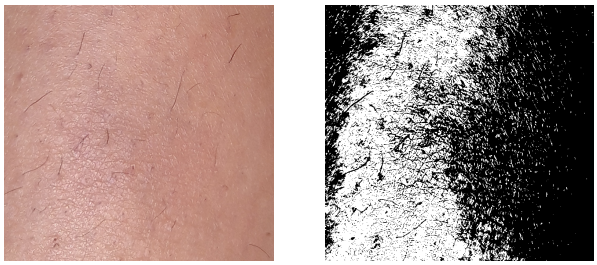


(C) The process.

FIGURE 4.31: Segmentation with active contours. (Bruise 11)

In this case, active contours completely failed to recognize the bruise. In the case of bruise 6, the edge at the end of the limb was identified and attracted the snake curve. In this case, the difference the limb and the background is less pronounced. The curve seems to have simply contracted without being attracted to any edge in particular.

Figure 4.32 shows bruise 9, a diffuse bruise, segmented by three of the four segmentation methods: Otsu thresholding (OT), k-means clustering (KM) and k-means clustering with histogram equalization (KH).



(A) The region of interest. (B) OT

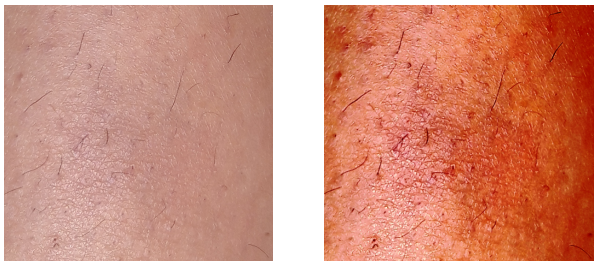
(C) KM (D) KH

FIGURE 4.32: Segmentation results for k-means and Otsu thresholding. (Bruise 9)

In this case, it is difficult to assess the accuracy of the segmentation, since the bruise is so faint. Figure 4.33 shows a larger view of the bruise. It is difficult to precisely define the borders of the bruise.

FIGURE 4.33: Bruise 9, larger view.

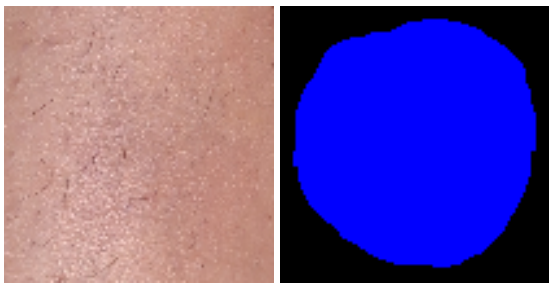
Even though the bruise is faint, it is clear that Otsu thresholding and k-means clustering with histogram equalization have not correctly identified the bruise, which is in the middle of the region of interest. They seem to have been affected by the shadows in the image. K-means clustering without histogram equalization corresponds somewhat better with the bruise, but is still quite inaccurate. Histogram equalization caused a worse segmentation performance in this case, since it enhanced the contrast of the skin hairs and not the bruise. Figure 4.34 shows the effect of histogram equalization.



(A) No histogram equalization. (B) With histogram equalization.

FIGURE 4.34: The effect of histogram equalization. (Bruise 9)

Figure 4.35 shows the segmentation results for the same bruise with active contours.



(A) Bruise 9 (B) AC

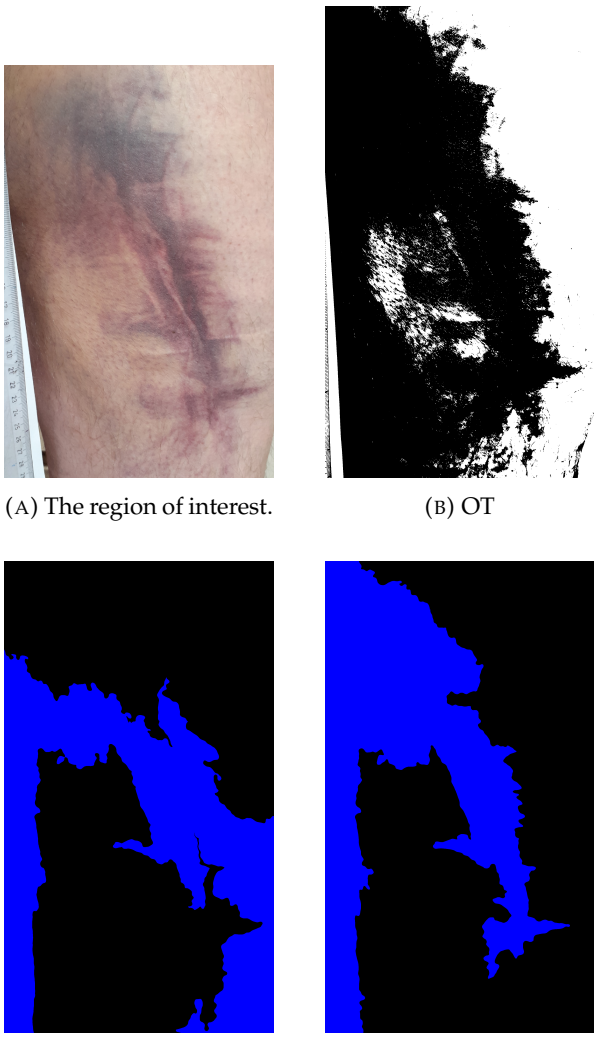
(C) The process.

FIGURE 4.35: Segmentation with active contours. (Bruise 9)

Even though the active contours segmentation resulted in a more plausible bruise segment than the other three methods, it is not clear that the bruise was actually detected by the methods. As is seen in Figures 4.26 and 4.35, active contours segmentation tends to end with a circle-like bruise segment unless a salient image edge is detected.

In the case of bruise 25, the one bruise classified as *complex*, a circle could not be defined that both included the entire bruise and stayed within the boundaries of the image. Since the bruise covered the majority of the image, the region of interest is not much smaller than the original image. Figure 4.32 shows bruise 25, segmented by

three of the four segmentation methods: Otsu thresholding (OT), k-means clustering (KM) and k-means clustering with histogram equalization (KH).



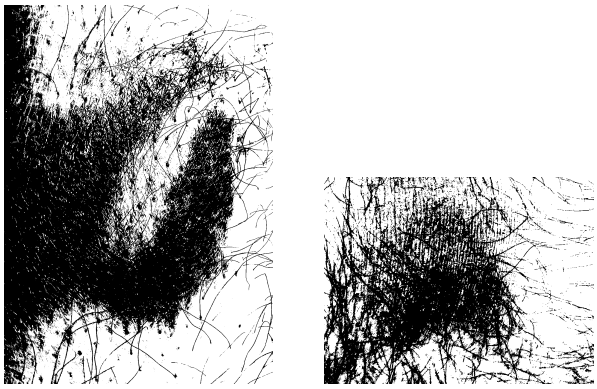
(A) The region of interest. (B) OT

(C) KM (D) KH

FIGURE 4.36: Segmentation results for k-means and Otsu thresholding. (Bruise 25)

As discussed in Section 4.2, it was not possible to extract a rectangular region of interest that included the entire bruise without also including parts of the background. It is clear that the ruler in the left part of the image was detected by all of the segmentation methods. It is also difficult to assess the accuracy of the segmentation, since it is unclear where the borders of the bruise are. It seems like the bruise may continue out of the upper edge of the image, but this may also just be a shadow.

It is clear that the performance of the segmentation methods depends on several factors: the type of bruise, the quality of the image and the presence of confounding factors such as skin hair. Otsu thresholding struggles with images that have skin hair or prominent shadows in them. Figure 4.37 shows examples of this.



(A) Bruise 20: hair (B) Bruise 1: hair

(C) Bruise 11: shadow (D) Bruise 9: shadow

FIGURE 4.37: Examples of shadows and skin hair interfering with Otsu thresholding.

4.7 Acquisition method

The participants were asked to rank the difficulty of doing each of these tasks.

1. Understand the instructions
2. Find a ruler and white paper
3. Measure the distance between the bruise and camera
4. Photograph the bruise
5. Upload the photos

Each task is ranked on a scale from (1) very easy to (5) very difficult. The average difficulty, standard deviation and most common difficulty rating of each task is shown in Table 4.19.

TABLE 4.19: Rating the difficulty of the bruise documentation tasks.

Task	Average difficulty	Standard deviation	Mode
1	1.64	0.85	1
2	1.86	0.74	2
3	3.10	0.94	3
4	2.45	1.16	1
5	1.45	0.81	1

As was expected, Task 3: *Measure the distance between the bruise and camera* was the most difficult task. It was removed from the simplified instructions for this reason.

For tasks 1, 4 and 5, *very easy* was the most common answer. However, the average difficulty of Task 4: *Photograph the bruises* is 2.45, which indicates that some participants found it difficult. Indeed, Task 4 had the highest number (7) of participants answering *difficult*. Only Task 3 was given the highest difficulty rating, *very difficult*, and this only happened once.

Bruises 14, 18, 19 and 22 were documented without a ruler and white paper. Interestingly, the participants who documented bruises 14 and 18 answered that finding a ruler and white paper was *easy* and *very easy*, respectively, even though none of their images included a ruler or white paper. They both also answered that understanding the instructions was *easy* and *very easy*. None of them gave any further feedback. The person who documented bruise 22 included a ruler, but no white paper. They answered that both understanding the instructions and finding a ruler and white paper was *very easy*.

The person who documented bruise 19 did not answer the question about the difficulty of finding a ruler and white paper. They answered that understanding the instructions was difficult. In the feedback text box, they wrote that it was not clear how they should use the paper and ruler.

In section 4.1.3 the acquired image set was evaluated with respect to the following criteria:

1. Include paper and ruler in image
2. Entire bruise visible
3. Good lighting without prominent shadows
4. Bruise and ruler in focus
5. Ruler parallel to camera and bruise
6. Keep the angle and distance of the camera constant

Each criteria can either be completely fulfilled, partially fulfilled or not fulfilled, which is denoted by 1, 0.5 and 0, respectively. The average score, standard deviation and mode of each criteria was computed. The results are shown in Table 4.20.

TABLE 4.20: Average scores for each image quality criterion.

Criterion	Average score	Standard deviation	Mode
1	0.75	0.38	1
2	0.92	0.27	1
3	0.46	0.51	0
4	0.65	0.46	1
5	0.60	0.49	1
6	0.65	0.39	1

Of the six criteria, Criterion 2: *Entire bruise visible* has the highest average score. In fact, only two participants did not fulfill this criterion. Criterion 3: *Good lighting without prominent shadows* has the lowest average score, and is the only criterion that was failed by the majority of the participants. Only 12 participants fulfilled Criterion 3. This indicates that it was difficult to achieve good lighting without prominent shadows.

4.8 Summary and discussion of findings

176 images of 26 bruises were acquired from 24 participants. The bruises in the data set are diverse in terms of size, shape and color. Five bruise classes were identified: blob, cluster, diffuse, central clearing and complex.

Some of the images in the data set are of low quality due to errors in the acquisition process. The acquisition method was evaluated by assessing the quality of the acquired image set and by gathering feedback from the participants. The participants found measuring the distance between the bruise and camera to be the most difficult, which is why it was excluded from the simplified instructions. Some participants failed to include a white paper and ruler, but reported that including a white paper and ruler was easy. This indicates that they both did not read the instructions carefully and did not read the questionnaire questions carefully. Participants found it difficult to keep the ruler parallel to the camera and bruise, as well as achieving good lighting without prominent shadows.

Each bruise image was segmented using the four selected segmentation methods: Otsu thresholding, k-means clustering, k-means clustering with histogram equalization and active contours. Based on the segmented images, the three selected features (size, shape and color) could be estimated.

In some cases, it was impossible to extract a region of interest (ROI) which includes the entire bruise without also including parts of the background. The performance of Otsu thresholding and k-means with and without histogram equalization rely on the assumption that the region of interest only contains two classes: bruise and skin. This assumption is violated in these cases.

Ruler detection is the first step of bruise size estimation. Automatic ruler detection was done for 10 bruises, manual ruler detection was done for 8 bruises, and ruler detection was impossible for 8 bruises. In the case of the impossible bruises, there was either no ruler present in the images, or the ruler was so blurry that determining the spacing between the ruler ticks was impossible. In the cases were manual ruler detection was done, automatic ruler detection failed due to errors that would not prevent a human from correctly determining the spacing between the ruler ticks. Such errors include using a lined white paper and rulers with both imperial and metric units. Good results were obtained for the 18 bruises where ruler detection, automatic or manual, was done. The average error, in pixels, was between 0 and 3.2 pixels for all of the bruises.

Two size measures were calculated for each bruise: bruise area and bruise width and length. There is no ground truth for the bruise area, so the accuracy could not be evaluated. Instead, the standard deviation of the area estimation was calculated. Otsu thresholding provided the largest area estimate for every single bruise in the data set, and had the highest standard deviation. Shadows proved to a problem for Otsu thresholding. Since Otsu thresholding uses the grayscale histogram of the image, and shadowed skin may have a similar gray level value to the bruise, shadows were sometimes mistakenly classified as bruises - resulting in a too large size estimate.

A ground truth length and width was extracted for each bruise, such that the accuracy of the bruise length and width estimation could be evaluated. As in the area estimates, Otsu thresholding often resulted in too large length and width estimates. K-means with and without histogram equalization produced length and width estimates that were more evenly spaced around the ground truth. Active contours over-estimated the width of most bruises, while the length estimates were closer to

the ground truth. This is caused by the fact that the active contours method seems to default to a circle-like shape when it fails to detect a clear bruise edge.

The tendency of active contours to produce circle-like bruise segments is also evident in the roundness estimates. Active contours consistently provided the highest roundness estimates and the lowest standard deviation of roundness. There is no ground truth for the roundness of the bruises, but three pairs of comparison bruises were defined to sample the accuracy of the roundness estimation. Active contours performed the worst in these comparisons. K-means with and without histogram equalization performed the best.

Two color measures were calculated for each bruise: bruise contrast and average bruise color. Since there is no ground truth for either of these measures, the accuracy of the color estimation could not be evaluated. The precision of the contrast estimation was evaluated by calculating the standard deviation. Active contours resulted in the lowest standard deviation for the majority of the bruises. This is likely related to the low standard deviation of shape estimation.

Two computational color constancy methods, white patch and gray world, were applied to all the bruise images. They were evaluated by comparing the whiteness of the paper. White patch outperformed gray world in all cases. This is to be expected, since the assumption of white patch was fulfilled by including a white paper in the scene, but the assumption of gray world was not necessarily fulfilled.

The robustness of the average color estimation was assessed by comparing the average bruise color with and without flash for each segmentation method and computational color constancy method. White patch and gray world had very similar performances. In the four cases where no white paper was included, gray world outperformed white patch. The precision of the average color estimation was assessed by computing the standard deviation of the color estimates. Here, white patch slightly outperformed gray world on average. Overall, white patch seems like a better choice given that a white paper has been included in the image. K-means clustering with histogram equalization produced the worst results in terms of robustness of average color estimation. This may be due to the effect of histogram equalization being different for flash and no flash images.

The four segmentation were ranked based on the aforementioned feature extraction results and their execution time. The final ranking, from best to worst, is: k-means clustering without histogram equalization, Otsu thresholding, active contours and k-means clustering with histogram equalization. As discussed in Section 4.6.3, the performance of the segmentation methods depended on the bruise type. Otsu thresholding struggles with shadows and skin hair, but outperforms the other segmentation methods in recognizing cluster bruises.

Chapter 5

Conclusion

5.1 Objectives

The objectives of this project are as follows:

1. Create a data set of bruise images acquired by self-documentation with smartphone imaging.
2. Estimate the relevant bruise features (size, shape and color)
3. Evaluate the performance of the feature extraction
4. Evaluate the bruise self-documentation method

The first objective was achieved successfully. The data set consists of 26 bruises and 176 bruise images.

The second objective was also achieved. The size, shape and color was estimated for every bruise, using several different image processing techniques. The proposed ruler detection method provided high precision in the images where the ruler was in focus. Two computational color constancy methods, white patch and gray world, were compared for the purpose of improving the robustness of bruise color estimation. Of the two, white patch performed the best in every single bruise.

The performance of the feature extraction was evaluated by computing the precision and accuracy of the estimated size, shape and color of the bruises. Four image segmentation methods, Otsu thresholding, k-means clustering, k-means clustering with histogram equalization and active contours, were compared and ranked. Of the four, k-means clustering performed the best.

Finally, the fourth objective was achieved. The bruise self-documentation method was evaluated by inspecting the quality of the acquired images and by gathering feedback from the participants. Achieving good lighting without significant shadows proved to be the most difficult part of the method. Further, measuring the distance between the bruise and the camera was reported as being difficult.

The main goal of this thesis was to answer the central research question: *Is self-documentation with smartphone imaging a suitable method for documenting skin bruises?* Further research is needed to answer that question with certainty. It is clear that a well-written set of instructions is needed in order for the images to be of high enough quality to extract the bruise features well. Smartphone imaging as a technology is likely suitable, but whether the bruise features can be reliably and accurately extracted regardless of smartphone model, lighting conditions and other factors such as skin color and bruise type is yet to be determined. This will be addressed in future work, see Section 5.2.

5.2 Challenges and future work

The largest challenge to the image segmentation was the poor lighting in the bruise images. More than half of the bruises were photographed in poor lighting conditions. In many images, the shadows on the skin were darker than the bruise, making image segmentation difficult. This could be improved by altering the instructions. The instructions should make the importance of good lighting more clear. The example image included in the instruction document showed a bruise image with poor lighting, providing a poor example for the participants. Another common challenge was rulers being out of focus or having both imperial and metric units, making the ruler detection method difficult. This could again be improved by clearly specifying the type of ruler and that it should be in focus in the instructions.

Another challenge was the existence of confounding factors on the skin, such as body hair and skin pigmentation (moles). This challenge is harder to counteract, since asking the participants to remove such confounding factors would be too invasive. Future work could include skin hair detection to deal with this challenge.

In some cases, a rectangular region of interest which includes the entire bruise without also including the background could not be extracted. These ROI's decreased the quality of the segmentation methods that relied on the assumption that there are only two classes within the ROI. This could be avoided by allowing for parallelogram shaped ROI's. Similarly, the active contours segmentation method could be improved by allowing other initial curves than circles, for example ellipses. Automatic optimal parameter selection could improve the performance of both k-means and active contours. Otsu thresholding could have been improved by including a morphological opening step, in which small specks of the segment are removed, as done by [64], discussed in Section 2.3. Future work should also consider other segmentation methods, such as watershed.

Assessing the performance of the segmentation proved challenging, as there was no true ground truth for the bruise segmentation. In future work, a ground truth will be acquired by having an expert manually segment the bruise images.

The chosen shape measure, roundness, only considers the length and the width of the bruise. Future work should also consider the interior of the bruise, so that recognition of central clearing bruises is possible. It is also of interest to recognize other typical bruise shapes, such as cluster bruises. The chosen method of length and width estimation can also be improved further.

The chosen color measures, bruise contrast and average bruise color, inherently assume that the color of the bruise is homogeneous. Many of the bruises have more than one color present. If a bruise is half yellow, half blue, the average bruise color descriptor will categorize the bruise as green. Future work should identify other color measures. In addition to the bruise contrast, the difference in the Lab-space L component between the bruise and the surrounding skin, one could compute the Lab color difference ΔE^*_{ab} and the chromaticity difference along the two Lab chromaticity axes.

Other computational color constancy methods than white patch and gray world will be tested in the feature, including machine learning and gamut mapping based methods.

The histogram equalization did often not work as intended. Future work could include testing adaptive histogram equalization.

The age and cause of each bruise was gathered. Future work will include analyzing the correlation of the bruise features (size, shape and color) with the age and cause of the bruises.

A future data set will be acquired with a more standardized setup, where the images are taken by the researcher and not the participants. A color calibration card will be used instead of a white paper, and a camera tripod will be used to keep the angle and position constant. The images will have higher quality, since issues such as out of focus rulers and bad lighting can be avoided. In a future project, the reliability of feature extraction will be evaluated by documenting the same bruise with different camera models and lighting environments and see how this affects the extracted features. This will assess whether it is possible to reliably extract bruise features under uncontrolled circumstances. It is also of interest to gather bruise images from a more diverse set of skin colors.

Appendix A

Instructions

NTNU

SELF-DOCUMENTATION OF BRUISES: STEP BY STEP

YOU WILL NEED

- A phone with a functioning camera
- A white paper and a ruler
- A bruise on your arms or legs that you remember **when** and **how** was made

WHAT YOU WILL DO

1. **Distance:** Write down the approximate distance between the skin and the phone
2. **Bruise photos:** Take 10 photos of your bruise. (Fig 1, instructions below)
3. **Reference photos:** Take 10 photos of your skin on the opposite side of your body. (Fig 2)
4. **Upload:** Upload the photos in *Nettskjema.* See link at the bottom.

The reference photos are used to account for varying skin colors.

HOW TO TAKE THE PHOTOS

- 5 photos with flash, 5 photos without flash (Both bruise photos and reference photos)
- Try to keep the distance and angle of your phone constant.
- The entire bruise should be visible in the bruise photos.
- The ruler and the paper should be visible in all photos.

We need many photos to account for varying lighting conditions.

SUGGESTIONS (NOT MANDATORY)

- Avoid strong shadows or fluorescent light.
- The skin should be in focus.
- Store the bruise photos and reference photos in separate folders, to make it easier

WHEN YOU ARE DONE: NETTSKJEMA

- Upload the photos here: https://nettskjema.no/a/bruise-images-e
- Answer some questions here: https://nettskjema.no/a/bruise-questions-e

Figure 1: Example set up. The bruise, paper and ruler are visible in the photo.

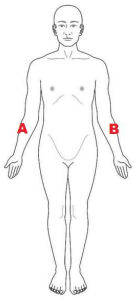


Figure 2: (A) Reference photo should be on the opposite side of the body wrt the (B) bruise

Appendix B

Simplified instructions

Self-documentation of bruises Pauline Hardeberg Zimmermann

NTNU

SELF-DOCUMENTATION OF BRUISES: STEP BY STEP

YOU WILL NEED

- 10 minutes
- A phone with a functioning camera
- A white paper and a ruler

WHAT YOU WILL DO

1. **Bruise photos:** Take 4 photos of your bruise. (Fig 1, instructions below)
2. **Upload:** Upload the photos in *Nettskjema.* See link at the bottom.

HOW TO TAKE THE PHOTOS

- 2 photos **with flash**, 2 photos **without flash**
- Try to keep the distance and angle of your phone constant.
- The entire bruise should be visible.
- The ruler and the paper should be visible.

SUGGESTIONS (NOT MANDATORY)

- Avoid strong shadows or fluorescent light.
- The skin should be in focus.

WHEN YOU ARE DONE: NETTSKJEMA

- Upload photos and answer questions here: https://nettskjema.no/a/bruises-el

Thank you!

Figure 1: Example set up. The bruise, paper and ruler are visible in the photo.

Bibliography

[1] ColorChecker® Classic; X-Rite. https://www.xrite.com/categories/calibration-profiling/colorchecker-classic.

[2] Forensigraph – Forensic Rulers & Bags.

[3] Photomacrographic Scales -ABFO No.2. https://tritechforensics.com/photomacrographic-scales-abfo-no-2/.

[4] State v. Powell, Apr. 1995.

[5] Sweaney v. Ada County, Idaho, July 1997.

[6] State v. Hovig, 202 P. 3d 318 - 2009, Jan. 2009.

[7] Adobe Photoshop Elements. Adobe Inc., 2018.

[8] L. Adelson. *The Pathology of Homicde*. Charles C Thomas, Springfield, 1974. ISBN 978-0-398-03000-1.

[9] M. Afifi, B. Price, S. Cohen, and M. S. Brown. When Color Constancy Goes Wrong: Correcting Improperly White-Balanced Images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1535–1544, 2019.

[10] A. Agarwal, A. Issac, M. K. Dutta, K. Riha, and V. Uher. Automated skin lesion segmentation using K-Means clustering from digital dermoscopic images. In *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*, pages 743–748, July 2017. doi: 10.1109/TSP.2017.8076087.

[11] C. Ahn and R. S. Salcido. Advances in Wound Photography and Assessment Methods. *Advances in Skin & Wound Care*, 21(2):85–93, Feb. 2008. ISSN 1527-7941. doi: 10.1097/01.ASW.0000305411.58350.7d.

[12] D. Alempijevic, S. Savic, S. Pavlekic, and D. Jecmenica. Severity of injuries among sexual assault victims. *Journal of Forensic and Legal Medicine*, 14(5):266–269, July 2007. ISSN 1752-928X. doi: 10.1016/j.jcfm.2006.08.008.

[13] R. R. Anderson. Polarized Light Examination and Photography of the Skin. *Archives of Dermatology*, 127(7):1000–1005, July 1991. ISSN 0003-987X. doi: 10.1001/archderm.1991.01680060074007.

[14] K. Aney. Karly's legacy: Local author tells tale of murdered girl who sparked new law. *East Oregonian*, 2012.

[15] H. C. Baker, N. Marsh, and I. Quinones. Photography of Faded or Concealed Bruises on Human Skin. *Journal of Forensic Identification*, 63(1):103–125, Jan/Feb 2013. ISSN 0895173X. Num Pages: 23.

[16] C. Barata, M. E. Celebi, and J. S. Marques. Improving Dermoscopy Image Classification Using Color Constancy. *IEEE Journal of Biomedical and Health Informatics*, 19(3):1146–1152, May 2015. ISSN 2168-2208. doi: 10.1109/JBHI.2014.2336473. Conference Name: IEEE Journal of Biomedical and Health Informatics.

[17] A. Bhalerao and G. Reynolds. Ruler Detection for Autoscaling Forensic Images. *International Journal of Digital Crime and Forensics*, 9(27), Jan. 2014. doi: 10.4018/ijdcf.2014010102. Journal Abbreviation: IJDCF Pages: 9-27.

[18] H. I. Black, S. Coupaud, N. N. Daéid, and P. E. Riches. On the relationships between applied force, photography technique, and the quantification of bruise appearance. *Forensic Science International*, 305:109998, Dec. 2019. ISSN 0379-0738. doi: 10.1016/j.forsciint.2019.109998.

[19] M. Bohnert, R. Baumgartner, and S. Pollak. Spectrophotometric evaluation of the colour of intra- and subcutaneous bruises. *International Journal of Legal Medicine*, 113(6):343–348, Oct. 2000. ISSN 1437-1596. doi: 10.1007/s004149900107.

[20] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[21] J. E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965. ISSN 0018-8670. doi: 10.1147/sj.41.0025. Conference Name: IBM Systems Journal.

[22] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26, July 1980. ISSN 0016-0032. doi: 10.1016/0016-0032(80)90058-7.

[23] C. Bugaud, G. Ocrisse, F. Salmon, and D. Rinaldo. Bruise susceptibility of banana peel in relation to genotype and post-climacteric storage conditions. *Postharvest Biology and Technology*, 87:113–119, Jan. 2014. ISSN 0925-5214. doi: 10.1016/j.postharvbio.2013.08.009.

[24] C. J. K. Bulstrode, A. W. Goode, and P. J. Scott. Stereophotogrammetry for measuring rates of cutaneous healing: A comparison with conventional techniques. *Clinical Science*, 71(4):437–443, Oct. 1986. ISSN 0143-5221, 1470-8736. doi: 10.1042/cs0710437.

[25] L. Busin, N. Vandenbroucke, and L. Macaire. Color Spaces and Image Segmentation. In *Advances in Imaging and Electron Physics*. Academic Press, 2008. ISBN 978-0-08-056913-0.

[26] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, Nov. 1986. ISSN 1939-3539. doi: 10.1109/TPAMI.1986.4767851. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

[27] C. Capper. The Language of Forensic Medicine: The Meaning of Some Terms Employed. *Medicine, Science and the Law*, 41(3):256–259, July 2001. ISSN 0025-8024, 2042-1818. doi: 10.1177/002580240104100309.

[28] M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, and R. H. Moss. A methodological approach to the classification of

dermoscopy images. *Computerized Medical Imaging and Graphics*, 31(6):362–373, Sept. 2007. ISSN 0895-6111. doi: 10.1016/j.compmedimag.2007.01.003.

[29] M. E. Celebi, T. Mendonca, and J. S. Marques, editors. *Dermoscopy Image Analysis*. CRC Press, 0 edition, Oct. 2015. ISBN 978-0-429-06919-2. doi: 10.1201/b19107-8.

[30] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, Feb. 2001. ISSN 1941-0042. doi: 10.1109/83.902291. Conference Name: IEEE Transactions on Image Processing.

[31] C. W. Christian. The Evaluation of Suspected Child Physical Abuse. *Pediatrics*, 135(5):e1337–e1354, May 2015. ISSN 0031-4005, 1098-4275. doi: 10.1542/peds.2015-0356.

[32] Commission Internationale de l'Eclairage. *Colorimetry: Technical Report*. Central Bureau of the CIE, 2004.

[33] T. L. Cosman, H. M. Arthur, D. Bryant-Lukosius, and P. H. Strachan. Reliability of Vascular Access Site Bruise Measurement. *Journal of Nursing Measurement*, 23(1):179–200, 2015. ISSN 1061-3749, 1945-7049. doi: 10.1891/1061-3749.23.1.179.

[34] L. S. Deutsch, K. Resch, T. Barber, Y. Zuckerman, J. T. Stone, and C. Cerulli. Bruise Documentation, Race and Barriers to Seeking Legal Relief for Intimate Partner Violence Survivors: A Retrospective Qualitative Study. *Journal of Family Violence*, 32(8):767–773, Nov. 2017. ISSN 1573-2851. doi: 10.1007/s10896-017-9917-4.

[35] C.-J. Du and D.-W. Sun. Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science & Technology*, 15(5):230–249, May 2004. ISSN 0924-2244. doi: 10.1016/j.tifs.2003.10.006.

[36] W. G. Eckert. The writings of Sir Bernard Spilsbury: Part I. *The American journal of forensic medicine and pathology*, 5(3):231–238, Sept. 1984. ISSN 1533-404X. doi: 10.1097/00000433-198409000-00012.

[37] Edmund Optics Inc. ColorGauge Nano, Matte. https://www.edmundoptics.com/p/colorgauge-nano-matte/29793/.

[38] N. S. Ellerstein. The Cutaneous Manifestations of Child Abuse and Neglect. *Archives of Pediatrics & Adolescent Medicine*, 133(9):906, Sept. 1979. ISSN 1072-4710. doi: 10.1001/archpedi.1979.02130090034005.

[39] B. Erkol, R. H. Moss, R. J. Stanley, W. V. Stoecker, and E. Hvatum. Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes. *Skin Research and Technology*, 11(1):17–26, 2005. ISSN 1600-0846. doi: https://doi.org/10.1111/j.1600-0846.2005.00092.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1600-0846.2005.00092.x.

[40] S. Evans, S. Baylis, R. Carabott, M. Jones, Z. Kelson, N. Marsh, J. Payne-James, J. Ramadani, P. Vanezis, and A. Kemp. Guidelines for photography of cutaneous marks and injuries: A multi-professional perspective. *Journal of Visual Communication in Medicine*, 37(1-2):3–12, May

2014. ISSN 1745-3054. doi: 10.3109/17453054.2014.911152. _eprint: https://doi.org/10.3109/17453054.2014.911152.

[41] J. W. Feather, M. Hajizadeh-Saffar, G. Leslie, and J. B. Dawson. A portable scanning reflectance spectrophotometer using visible wavelengths for the rapid measurement of skin pigments. *Physics in Medicine and Biology*, 34(7): 807–820, July 1989. ISSN 0031-9155. doi: 10.1088/0031-9155/34/7/002.

[42] E. Fuchs and C. Byrne. The epidermis: Rising to the surface. *Current Opinion in Genetics & Development*, 4(5):725–736, Oct. 1994. ISSN 0959-437X. doi: 10.1016/0959-437X(94)90140-X.

[43] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational Color Constancy: Survey and Experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, Sept. 2011. ISSN 1057-7149, 1941-0042. doi: 10.1109/TIP.2011.2118224.

[44] J. Glaister. *Glaister's Medical Jurisprudence and Toxicology*. Churchill Livingstone, Edinburgh, 13 edition, 1973. ISBN 978-0-443-00894-8.

[45] T. Y. Goh, S. N. Basah, H. Yazid, M. J. Aziz Safar, and F. S. Ahmad Saad. Performance analysis of image thresholding: Otsu technique. *Measurement*, 114: 298–307, Jan. 2018. ISSN 0263-2241. doi: 10.1016/j.measurement.2017.09.052.

[46] R. Gonzales and R. Woods. *Digital Image Processing*. Pearson, 4 edition, 2008. ISBN 978-0-13-335672-4.

[47] J. Gould. Superpowered skin. *Nature*, 563:S84–S85, Nov. 2018. doi: 10.1038/d41586-018-07429-3.

[48] N. Hameed, A. Shabut, and M. A. Hossain. A Computer-aided diagnosis system for classifying prominent skin lesions using machine learning. In *2018 10th Computer Science and Electronic Engineering (CEEC)*, pages 186–191, Colchester, United Kingdom, Sept. 2018. IEEE. ISBN 978-1-5386-7275-4. doi: 10.1109/CEEC.2018.8674183.

[49] K. L. Hanlon. Cross-polarised and parallel-polarised light: Viewing and photography for examination and documentation of biological materials in medicine and forensics. *Journal of Visual Communication in Medicine*, 41(1):3–8, Jan. 2018. ISSN 1745-3054. doi: 10.1080/17453054.2018.1420418. _eprint: https://doi.org/10.1080/17453054.2018.1420418.

[50] J. Y. Hardeberg. *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches*. Universal-Publishers, 2001. ISBN 978-1-58112-135-3.

[51] C. Harris, A. Alcock, L. Trefan, D. Nuttall, S. T. Evans, S. Maguire, and A. M. Kemp. Optimising the measurement of bruises in children across conventional and cross polarized images using segmentation analysis techniques in Image J, Photoshop and circle diameter measurements. *Journal of Forensic and Legal Medicine*, 54:114–120, Feb. 2018. ISSN 1752-928X. doi: 10.1016/j.jflm.2017.12.020.

[52] T. S. Harris. Bruises in Children: Normal or Child Abuse? *Journal of Pediatric Health Care*, 24(4):216–221, July 2010. ISSN 0891-5245. doi: 10.1016/j.pedhc.2009.03.007.

[53] M. K. Hasan, S. Roy, C. Mondal, M. A. Alam, M. T. E. Elahi, A. Dutta, S. M. T. U. Raju, and M. Ahmad. Dermo-DOCTOR: A web application for detection and recognition of the skin lesion using a deep convolutional neural network. *Biomedical Signal Processing and Control*, 68, Feb. 2021.

[54] I. Hernández-Neuta, F. Neumann, J. Brightmeyer, T. B. Tis, N. Madaboosi, Q. Wei, A. Ozcan, and M. Nilsson. Smartphone-based clinical diagnostics: Towards democratization of evidence-based health care. *Journal of Internal Medicine*, 285(1):19–39, 2019. ISSN 1365-2796. doi: 10.1111/joim.12820. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/joim.12820.

[55] G. E. Hill, G. E. Hill, K. J. McGraw, and K. J. McGraw. *Bird Coloration: Mechanisms and Measurements*. Harvard University Press, 2006. ISBN 978-0-674-01893-8.

[56] M. Hixenbaugh. Hundreds of parents say kids wrongly taken from them after doctors misdiagnosed abuse. *NBC News*, Dec. 2019.

[57] P. V. C. Hough. Method and means for recognizing complex patterns, Dec. 1962.

[58] V. K. Hughes and N. E. I. Langlois. Use of reflectance spectrophotometry and colorimetry in a general linear model for the determination of the age of bruises. *Forensic Science, Medicine, and Pathology*, 6(4):275–281, Dec. 2010. ISSN 1556-2891. doi: 10.1007/s12024-010-9171-z.

[59] V. K. Hughes, P. S. Ellis, and N. E. I. Langlois. The perception of yellow in bruises. *Journal of Clinical Forensic Medicine*, 11(5):257–259, Oct. 2004. ISSN 1353-1131. doi: 10.1016/j.jcfm.2004.01.007.

[60] M. Ito and N. Katoh. Three-dimensional gamut mapping using various color difference formulae and color spaces. In *Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts IV*, volume 3648, pages 83–95. International Society for Optics and Photonics, Dec. 1998. doi: 10.1117/12.334547.

[61] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Computing Surveys*, 31(3):264–323, Sept. 1999. ISSN 0360-0300, 1557-7341. doi: 10.1145/331499.331504.

[62] W. D. James, D. Elston, and T. Berger. *Andrew's Diseases of the Skin*. Elsevier Health Sciences, 13 edition, Mar. 2011. ISBN 978-1-4377-3619-9.

[63] G. H. Joblove and D. Greenberg. Color spaces for computer graphics. In *Proceedings of the 5th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '78, pages 20–25, New York, NY, USA, Aug. 1978. Association for Computing Machinery. ISBN 978-1-4503-7908-3. doi: 10.1145/800248.807362.

[64] B. Johnson and R. Fazel-Rezai. Contusion (bruise) segmentation diagnosis: A graphical user interphase approach. In *2016 IEEE International Conference on Electro Information Technology (EIT)*, pages 0744–0749, May 2016. doi: 10.1109/EIT.2016.7535332.

[65] K. Kaczor, M. Clyde Pierce, K. Makoroff, and T. S. Corey. Bruising and Physical Child Abuse. *Clinical Pediatric Emergency Medicine*, 7(3):153–160, Sept. 2006. ISSN 1522-8401. doi: 10.1016/j.cpem.2006.06.007.

[66] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, Jan. 1988. ISSN 0920-5691, 1573-1405. doi: 10.1007/BF00133570.

[67] A. M. Kemp, S. A. Maguire, D. Nuttall, P. Collins, and F. Dunstan. Bruising in children who are assessed for suspected physical abuse. *Archives of Disease in Childhood*, 99(2):108–113, Feb. 2014. ISSN 0003-9888, 1468-2044. doi: 10.1136/archdischild-2013-304339.

[68] A. M. Kemp, F. Dunstan, D. Nuttall, M. Hamilton, P. Collins, and S. Maguire. Patterns of bruising in preschool children—a longitudinal study. *Archives of Disease in Childhood*, 100(5):426–431, May 2015. ISSN 0003-9888, 1468-2044. doi: 10.1136/archdischild-2014-307120.

[69] K. M. Kennedy. The relationship of victim injury to the progression of sexual crimes through the criminal justice system. *Journal of Forensic and Legal Medicine*, 19(6):309–311, Aug. 2012. ISSN 1752-928X. doi: 10.1016/j.jflm.2012.04.033.

[70] Y.-T. Kim. Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Transactions on Consumer Electronics*, 43(1):1–8, Feb. 1997. ISSN 1558-4127. doi: 10.1109/30.580378. Conference Name: IEEE Transactions on Consumer Electronics.

[71] E. H. Land. The Retinex Theory of Color Vision. *Scientific American*, 237(6):108–129, 1977. ISSN 0036-8733.

[72] D. Langemo, J. Anderson, D. Hanson, S. Hunter, and P. Thompson. Measuring Wound Length, Width, and Area: Which Technique? *Advances in Skin & Wound Care*, 21(1):42–45, Jan. 2008. ISSN 1527-7941. doi: 10.1097/01.ASW.0000284967.69863.2f.

[73] N. E. I. Langlois. The science behind the quest to determine the age of bruises—a review of the English language literature. *Forensic Science, Medicine, and Pathology*, 3(4):241–251, Dec. 2007. ISSN 1547-769X. doi: 10.1007/s12024-007-9019-3.

[74] N. E. I. Langlois and G. A. Gresham. The ageing of bruises: A review and study of the colour changes with time. *Forensic Science International*, 50(2):227–238, Sept. 1991. ISSN 0379-0738. doi: 10.1016/0379-0738(91)90154-B.

[75] Z. Lawson, D. Nuttall, S. Young, S. Evans, S. Maguire, F. Dunstan, and A. M. Kemp. Which is the preferred image modality for paediatricians when assessing photographs of bruises in children? *International Journal of Legal Medicine*, 125(6):825–830, Nov. 2011. ISSN 1437-1596. doi: 10.1007/s00414-010-0532-7.

[76] Z. Lawson, F. Dunstan, D. Nuttall, S. Maguire, A. Kemp, S. Young, M. Barker, and L. David. How Consistently Do We Measure Bruises? A Comparison of Manual and Electronic Methods. *Child Abuse Review*, 24(1):28–36, 2015. ISSN 1099-0852. doi: 10.1002/car.2217. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/car.2217.

[77] C.-R. Li, H.-Y. Tsai, C.-C. Yang, M.-Y. Lin, K.-C. Huang, and Y.-H. Lin. Quantifying the color changes in bruised skin using a color-calibrated imaging system. In *2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–5, June 2020. doi: 10.1109/MeMeA49120.2020.9137217.

[78] R. Liesner, I. Hann, and K. Khair. Non-accidental injury and the haematologist: The causes and investigation of easy bruising. *Blood Coagulation & Fibrinolysis*, 15:S41, May 2004. ISSN 0957-5235.

[79] D. Liu and J. Yu. Otsu Method and K-means. In *2009 Ninth International Conference on Hybrid Intelligent Systems*, volume 1, pages 344–349, Aug. 2009. doi: 10.1109/HIS.2009.74.

[80] J. Liu, Z. Geng, Z. Fan, J. Liu, and H. Chen. Point-of-care testing based on smartphone: The current state-of-the-art (2017–2018). *Biosensors and Bioelectronics*, 132:17–37, May 2019. ISSN 0956-5663. doi: 10.1016/j.bios.2019.01.068.

[81] Lovdata. Lov om helsepersonell m.v. (helsepersonelloven). https://lovdata.no/dokument/NL/lov/1999-07-02-64, 1999.

[82] G. Lu and B. Fei. Medical hyperspectral imaging: A review. *Journal of Biomedical Optics*, 19(1):010901, Jan. 2014. ISSN 1083-3668, 1560-2281. doi: 10.1117/1.JBO.19.1.010901.

[83] Y. Lu and R. Lu. Histogram-based automatic thresholding for bruise detection of apples by structured-illumination reflectance imaging. *Biosystems Engineering*, 160:30–41, Aug. 2017. ISSN 1537-5110. doi: 10.1016/j.biosystemseng.2017.05.005.

[84] F. Lyons and L. E. Ousley. *Dermatology for the Advanced Practice Nurse*. Springer Publishing Company, July 2014. ISBN 978-0-8261-3643-5.

[85] J. MacQueen. Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability: Weather modification*, 1(14):281–297, 1967.

[86] S. Maguire, M. K. Mann, J. Sibert, and A. Kemp. Are there patterns of bruising in childhood which are diagnostic or suggestive of abuse? A systematic review. *Archives of Disease in Childhood*, 90(2):182–186, Feb. 2005. ISSN 0003-9888, 1468-2044. doi: 10.1136/adc.2003.044065.

[87] MathWorks. Color constancy toolbox. https://se.mathworks.com/matlabcentral/fileexchange/color-constancy-toolbox, Aug. 2015.

[88] MATLAB. *Version R2021a*. The MathWorks Inc., Natick, Massachusetts, 2021.

[89] M. J. McGregor, J. D. Mont, and T. L. Myhr. Sexual assault forensic medical examination: Is evidence related to successful prosecution? *Annals of Emergency Medicine*, 39(6):639–647, June 2002. ISSN 0196-0644. doi: 10.1067/mem.2002.123694.

[90] J. D. Melville, J. L. Lukefahr, J. Cornell, N. D. Kellogg, and J. L. Lancaster. The Effect of Image Quality on the Assessment of Child Abuse Photographs. *Pediatric Emergency Care*, 29(5):607–611, May 2013. ISSN 0749-5161. doi: 10.1097/PEC.0b013e31828e848e.

[91] R. Montoriol, F. Savall, M. Vergnault, N. Telmon, C. Bartoli, and M. E. Rougé Bugat. Spatial distribution of physical intimate partner violence in the city of Toulouse. *La Revue de Médecine Légale*, 9(4):162–167, Dec. 2018. ISSN 1878-6529. doi: 10.1016/j.medleg.2018.08.001.

[92] A. Morar, F. Moldoveanu, and E. Gröller. Image segmentation based on active contours without edges. In *2012 IEEE 8th International Conference on Intelligent Computer Communication and Processing*, pages 213–220, Aug. 2012. doi: 10.1109/ICCP.2012.6356188.

[93] S. Naidoo. A profile of the oro-facial injuries in child physical abuse at a children's hospital. *Child Abuse & Neglect*, 24(4):521–534, Apr. 2000. ISSN 0145-2134. doi: 10.1016/S0145-2134(00)00114-9.

[94] K. R. Nash and D. J. Sheridan. Can one accurately date a bruise? State of the science:. *Journal of Forensic Nursing*, 5(1):31–37, Mar. 2009. ISSN 1556-3693. doi: 10.1111/j.1939-3938.2009.01028.x.

[95] Nasjonalt Kunnskapssenter om Vold og Traumatisk Stress. Sporsikring. https://voldsveileder.nkvts.no/blog/innhold/dokumentasjon-og-journalforing/sporsikring/.

[96] J. Neitz and G. H. Jacobs. Polymorphism of the long-wavelength cone in normal human colour vision. *Nature*, 323(6089):623–625, Oct. 1986. ISSN 1476-4687. doi: 10.1038/323623a0.

[97] S. W. Oh and S. J. Kim. Approaching the computational color constancy as a classification problem through deep learning. *Pattern Recognition*, 61:405–416, Jan. 2017. ISSN 0031-3203. doi: 10.1016/j.patcog.2016.08.013.

[98] Oregon Department of Justice. Karly's Law. https://www.doj.state.or.us/crime-victims/for-medical-providers/karlys-law/.

[99] N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, Jan. 1979.

[100] J. Payne-James, A. Busuttil, and W. Smock. *Forensic Medicine: Clinical and Pathological Aspects*. Cambridge University Press, 2003. ISBN 978-1-84110-026-5.

[101] J. J. Payne-James, C. Hawkins, S. Baylis, and N. P. Marsh. Quality of photographic images provided for injury interpretation: Room for improvement? *Forensic science, medicine, and pathology*, 8(4):447–450, 2012.

[102] J. J. Payne-James, M. M. Stark, M. Nittis, and D. R. Sheasby. Injury Assessment, Documentation, and Interpretation. In M. M. Stark, editor, *Clinical Forensic Medicine: A Physician's Guide*, pages 143–194. Springer International Publishing, Cham, 2020. ISBN 978-3-030-29462-5. doi: 10.1007/978-3-030-29462-5_4.

[103] Pew Research Center. Demographics of Mobile Device Ownership and Adoption in the United States.

[104] D. T. Pham, S. S. Dimov, and C. D. Nguyen. Selection of K in K -means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 219(1):103–119, Jan. 2005. ISSN 0954-4062, 2041-2983. doi: 10.1243/095440605X8298.

[105] E. Pham and S. Liao. Clinician's Role in the Documentation of Elder Mistreatment. *Geriatrics and Aging*, 12:323–327, 2009.

[106] M. L. Pilling, P. Vanezis, D. Perrett, and A. Johnston. Visual assessment of the timing of bruising by forensic experts. *Journal of Forensic and Legal Medicine*, 17 (3):143–149, Apr. 2010. ISSN 1752-928X. doi: 10.1016/j.jflm.2009.10.002.

[107] C. J. Polson, D. J. Gee, and B. Knight. *The Essentials of Forensic Medicine*. Pergamon Press, Oxford, 4 edition, 1973.

[108] D. M. Pressel. Evaluation of Physical Abuse in Children. *American Family Physician*, 61(10):3057–3064, May 2000. ISSN 0002-838X, 1532-0650.

[109] Python. *Version 3.8.5*. Python Software Foundation, 2020.

[110] L. L. Randeberg, A. M. Winnem, S. Blindheim, O. A. Haugen, and L. O. Svaasand. Optical classification of bruises. In *Lasers in Surgery: Advanced Characterization, Therapeutics, and Systems XIV*, volume 5312, pages 54–64. International Society for Optics and Photonics, July 2004. doi: 10.1117/12.538045.

[111] L. L. Randeberg, O. A. Haugen, R. Haaverstad, and L. O. Svaasand. A novel approach to age determination of traumatic injuries by reflectance spectroscopy. *Lasers in Surgery and Medicine*, 38(4):277–289, 2006. ISSN 1096-9101. doi: 10.1002/lsm.20301.

[112] L. L. Randeberg, A. M. Winnem, N. E. Langlois, E. L. P. Larsen, R. Haaverstad, B. Skallerud, O. A. Haugen, and L. O. Svaasand. Skin changes following minor trauma. *Lasers in Surgery and Medicine*, 39(5):403–413, 2007. ISSN 1096-9101. doi: 10.1002/lsm.20494.

[113] L. L. Randeberg, A. M. Winnem, E. L. P. Larsen, R. Haaverstad, O. A. Haugen, and L. O. Svaasand. In vivo hyperspectral imaging of traumatic skin injuries in a porcine model. In *Photonic Therapeutics and Diagnostics III*, volume 6424, page 642408. International Society for Optics and Photonics, Mar. 2007. doi: 10.1117/12.699380.

[114] L. L. Randeberg, B. Skallerud, N. E. Langlois, O. A. Haugen, and L. O. Svaasand. The Optics of Bruising. In A. J. Welch and M. J. van Gemert, editors, *Optical-Thermal Response of Laser-Irradiated Tissue*, pages 825–858. Springer Netherlands, Dordrecht, 2011. ISBN 978-90-481-8831-4. doi: 10.1007/978-90-481-8831-4_22.

[115] Regionale komiteer for medisinsk og helsefaglig forskningsetikk. Hjem. https://rekportalen.no/.

[116] Rik Henderson. From Galaxy S to Galaxy S10, here's a timeline of Samsung's flagship Android phones in pictures. https://www.pocket-lint.com/phones/news/samsung/136736-timeline-of-samsung-galaxy-flagship-android-phones-in-pictures, Mar. 2019.

[117] S. D. Roy, D. H. Das, M. K. Bhowmik, and A. K. Ghosh. Bruise detection in apples using infrared imaging. In *2016 9th International Conference on Electrical and Computer Engineering (ICECE)*, pages 118–122, Dec. 2016. doi: 10.1109/ICECE.2016.7853870.

[118] D. Rubin, C. W. Christian, L. T. Bilaniuk, K. A. Zazyczny, and D. R. Durbin. Occult Head Injury in High-Risk Abused Children. *Pediatrics*, 111(6), June 2003. doi: https://doi.org/10.1542/peds.111.6.1382.

[119] P. K. Sahoo, S. Soltani, and A. K. C. Wong. A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41(2):233–260, Feb. 1988. ISSN 0734-189X. doi: 10.1016/0734-189X(88)90022-9.

[120] M. Satone, S. Diwakar, and V. Joshi. Automatic Bruise Detection in Fruits Using Thermal Images. *International Journal of Advanced Research in Computer Science and Software Engineering*, 7(5):727–732, May 2017. ISSN 22776451, 2277128X. doi: 10.23956/ijarcsse/SV7I5/0116.

[121] P. Saukko and B. Knight. *Knight's Forensic Pathology*. CRC Press, London, 4 edition, Nov. 2015. ISBN 978-0-429-10235-6. doi: 10.1201/b13266.

[122] G. Schaefer, M. I. Rajab, M. Emre Celebi, and H. Iyatomi. Colour and contrast enhancement for improved skin lesion segmentation. *Computerized Medical Imaging and Graphics*, 35(2):99–104, Mar. 2011. ISSN 0895-6111. doi: 10.1016/j.compmedimag.2010.08.004.

[123] A. J. Schwartz and L. R. Ricci. How Accurately Can Bruises Be Aged in Abused Children? Literature Review and Synthesis. *Pediatrics*, 97(2):254–257, Feb. 1996. ISSN 0031-4005, 1098-4275.

[124] R. Shepherd. *Simpson's Forensic Medicine*. CRC Press, 2003. ISBN 978-0-340-81059-0.

[125] M. Shimrat. Algorithm 112: Position of point relative to polygon. *Communications of the ACM*, 5(8):434, Aug. 1962. ISSN 0001-0782, 1557-7317. doi: 10.1145/368637.368653.

[126] A. Smythe. Parents facing 'unfair child abuse claims' over bruising. *BBC News*, Apr. 2018.

[127] T. Stedman. *Stedman's Medical Dictionary*. Dalcassian Publishing Company, 1979.

[128] T. Stephenson and Y. Bialas. Estimation of the age of bruising. *Archives of Disease in Childhood*, 74(1):53–55, Jan. 1996. ISSN 0003-9888.

[129] C. J. Sully, K. L. Olds, and N. E. I. Langlois. Evaluation of a model of bruising in pigmented skin for investigating the potential for alternate light source illumination to enhance the appearance of bruises by photography of visible and infrared light. *Forensic Science, Medicine and Pathology*, 15(4):555–563, Dec. 2019. ISSN 1556-2891. doi: 10.1007/s12024-019-00135-0.

[130] The GIMP Development Team. GIMP, Mar. 2021.

[131] O. J. Tobias and R. Seara. Image segmentation by histogram thresholding using fuzzy sets. *IEEE Transactions on Image Processing*, 11(12):1457–1465, Dec. 2002. ISSN 1941-0042. doi: 10.1109/TIP.2002.806231. Conference Name: IEEE Transactions on Image Processing.

[132] L. Trefan, C. Harris, S. Evans, D. Nuttall, S. Maguire, and A. Kemp. A comparison of four different imaging modalities – Conventional, cross polarized, infra-red and ultra-violet in the assessment of childhood bruising. *Journal of Forensic and Legal Medicine*, 59:30–35, Oct. 2018. ISSN 1752-928X. doi: 10.1016/j.jflm.2018.07.015.

[133] O. Trujillo, P. Vanezis, and M. Cermignani. Photometric assessment of skin colour and lightness using a tristimulus colorimeter: Reliability of inter and intra-investigator observations in healthy adult volunteers. *Forensic Science International*, 81(1):1–10, July 1996. ISSN 0379-0738. doi: 10.1016/0379-0738(96)01939-1.

[134] M. Tsokos. Diagnostic criteria for cutaneous injuries in child abuse: Classification, findings, and interpretation. *Forensic Science, Medicine, and Pathology*, 11(2):235–242, June 2015. ISSN 1547-769X, 1556-2891. doi: 10.1007/s12024-015-9671-y.

[135] K. Ueda, T. Baba, Y. Nakagawa, and K. Amano. Detection of Scale Intervals in Digital Images. In *21st International Conference on Data Engineering Workshops (ICDEW'05)*, pages 1232–1232, Apr. 2005. doi: 10.1109/ICDE.2005.211.

[136] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu. Scikit-image: Image processing in Python. *PeerJ*, 2:e453, June 2014. ISSN 2167-8359. doi: 10.7717/peerj.453.

[137] P. Vanezis. Interpreting bruises at necropsy. *Journal of Clinical Pathology*, 54(5):348–355, May 2001. ISSN 0021-9746, 1472-4146. doi: 10.1136/jcp.54.5.348.

[138] P. Viktorin. Bresenham: An implementation of Bresenham's line drawing algorithm.

[139] M. G. Ward, A. Ornstein, A. Niec, and C. L. Murray. The medical assessment of bruising in suspected child maltreatment cases: A clinical perspective. *Paediatrics & Child Health*, 18(8):434–438, Oct. 2013. ISSN 1205-7088. doi: 10.1093/pch/18.8.433.

[140] G. F. Webster. Common skin disorders in the elderly. *Clinical Cornerstone*, 4(1):39–44, Jan. 2001. ISSN 1098-3597. doi: 10.1016/S1098-3597(01)90006-7.

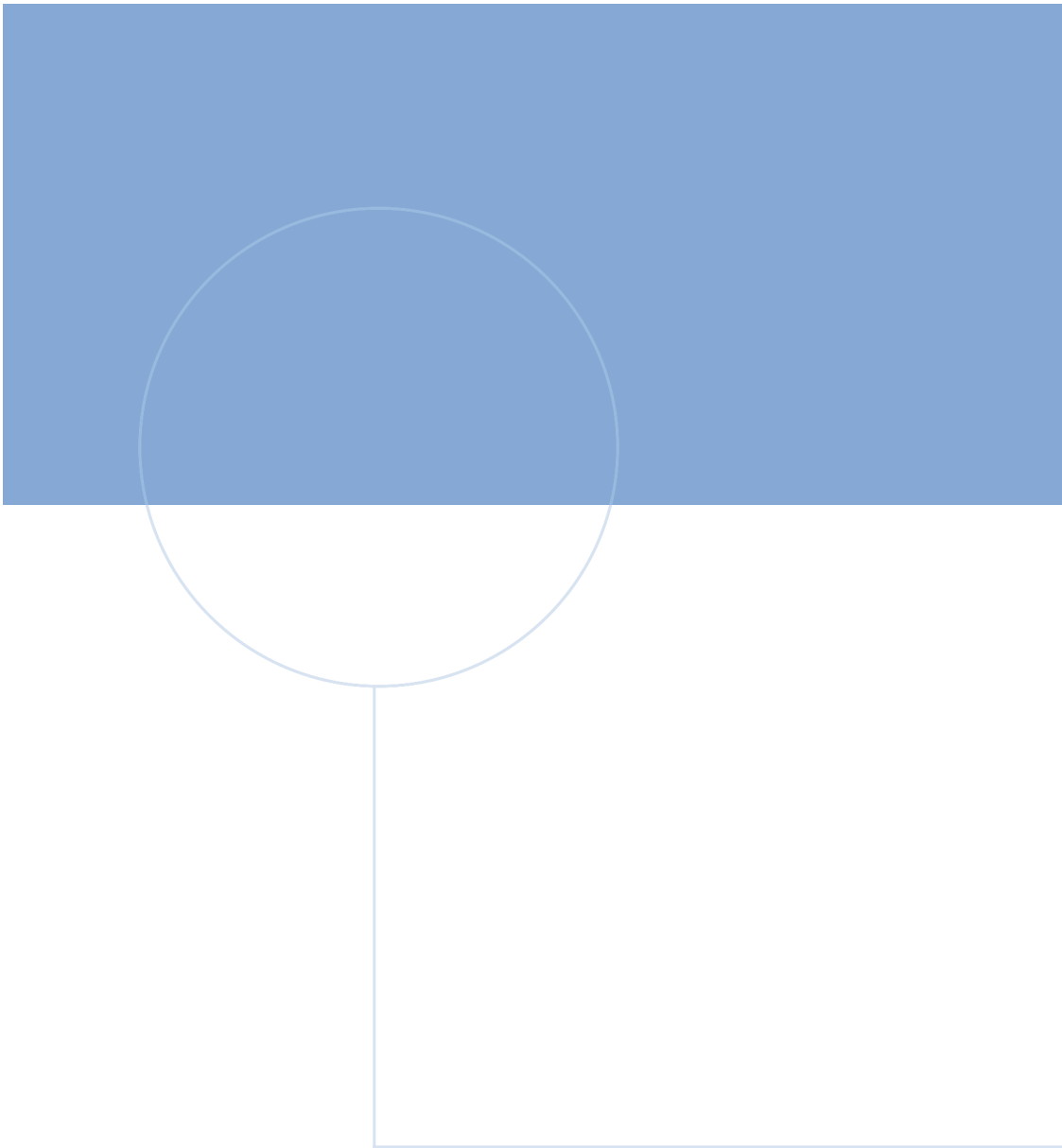
[141] A. J. Welch, M. J. van Gemert, and W. M. Star. Definitions and Overview of Tissue Optics. In A. J. Welch and M. J. van Gemert, editors, *Optical-Thermal Response of Laser-Irradiated Tissue*, pages 27–64. Springer Netherlands, Dordrecht, 2011. ISBN 978-90-481-8831-4. doi: 10.1007/978-90-481-8831-4_3.

[142] A. Wiglesworth, R. Austin, M. Corona, D. Schneider, S. Liao, L. Gibbs, and L. Mosqueda. Bruising as a Marker of Physical Elder Abuse. *Journal of the American Geriatrics Society*, 57(7):1191–1196, 2009. ISSN 1532-5415. doi: 10.1111/j.1532-5415.2009.02330.x.

[143] F. D. Wright and G. S. Golden. The use of full spectrum digital photography for evidence collection and preservation in cases involving forensic odontology. *Forensic Science International*, 201(1):59–67, Sept. 2010. ISSN 0379-0738. doi: 10.1016/j.forsciint.2010.03.013.

[144] Y. Yajima and M. Funayama. Spectrophotometric and tristimulus analysis of the colors of subcutaneous bleeding in living persons. *Forensic Science International*, 156(2):131–137, Jan. 2006. ISSN 0379-0738. doi: 10.1016/j.forsciint.2003.09.022.

[145] N. Zhu, G. Wang, G. Yang, and W. Dai. A Fast 2D Otsu Thresholding Algorithm Based on Improved Histogram. In *2009 Chinese Conference on Pattern Recognition*, pages 1–5, Nov. 2009. doi: 10.1109/CCPR.2009.5344078.



NTNU

Norwegian University of
Science and Technology