IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 29, 2019, accepted September 28, 2019, date of publication October 9, 2019, date of current version November 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2946313

Evaluation of Handwritten Urdu Text by Integration of MNIST Dataset Learning Experience

SAAD BIN AHMED[1,2], IBRAHIM A. HAMEED[3,4,5], SAEEDA NAZ[4], MUHAMMAD IMRAN RAZZAK[5], AND RUBIYAH YUSOF[1]

[1]Center of Artificial Intelligence and RObotics (CAIRO), Malaysia-Japan International Institute of Technology (M-JIIT), Universiti Teknologi Malaysia, Kuala Lumpur 54100, Malaysia
[2]Department of Health Informatics, King Saud bin Abdulaziz University for Health Sciences, Riyadh 11481, Saudi Arabia
[3]Department of ICT and Natural Sciences, Faculty of Information Technology and Electrical Engineering, Norges Teknisk-Naturvitenskapelige Universitet (NTNU), 6025 Alesund, Norway
[4]GGPGC No.1 Abbottabad, Higher Education Department, Abbottabad 22010, Pakistan
[5]School of Computer Science, University of Technology Sydney, Ultimo, NSW 2007, Australia

Corresponding author: Saad Bin Ahmed (isaadahmed@gmail.com)

ABSTRACT The similar nature of patterns may enhance the learning if the experience they attained during training is utilized to achieve maximum accuracy. This paper presents a novel way to exploit the transfer learning experience of similar patterns on handwritten Urdu text analysis. The MNIST pre-trained network is employed by transferring it's learning experience on Urdu Nastaliq Handwritten Dataset (UNHD) samples. The convolutional neural network is used for feature extraction. The experiments were performed using deep multidimensional long short term (MDLSTM) memory networks. The obtained result shows immaculate performance on number of experiments distinguished on the basis of handwritten complexity. The result of demonstrated experiments show that pre-trained network outperforms on subsequent target networks which enable them to focus on a particular feature learning. The conducted experiments presented astonishingly good accuracy on UNHD dataset.

INDEX TERMS Transfer learning, deep analysis, cursive handwritten data, MDLSTM.

I. INTRODUCTION

IT becomes easier for computer scientists to understand the mathematical complexities involves in realization of specialized but complicated systems, because mathematics prepare students to understand and make them learn the equations involves during implementation of computational problems. In general, having an experience to face a specific nature problem in particular situation helps to solve the similar problems if occurs in future. In academical terms the learning is possible by transferring the experience having similar problems for the purpose to learn quickly by exploiting the past experience. The machine learners are trained on single type of problem whereas human learning is broad which span over the lifetime. However, in neural networks, transfer learning phenomenon defines in a way that helps in learning related but new problems through pre-trained network. The goal is to learn the related problems more quickly. Some of reported research outperforms experimental analysis by using this said approach [1]. The focus of this study is to investigate the Urdu handwritten samples by deep learning approach. The proposed assumption is to transfer the learning experience of MNIST dataset [2] and use it's pre-trained network for training the Urdu handwritten (UNHD) text. In order to realize this hypothesis, the implementation is divided into two halves. The details about each half is explained in later sections below.

The interest is gradually witnessed in pattern recognition and natural language processing community to conduct research on cursive scripts like Chinese and Japanese. The research on aforementioned scripts obtained encouraging results in terms of accuracy as described in [3], [4]. As far as Arabic script is concerned, recent years work are investigated as published by researchers [5]–[8], but the inherent complexities which are part of Arabic scripts make it difficult

The associate editor coordinating the review of this manuscript and approving it for publication was Kok Lim Alvin Yau.

to recognize the text even with segmentation based classifiers. In such script's recognition, there is a need to have implicit segmentation techniques as reported by [9]. The context of cursive script is investigated where characters are difficult to segment. The appearance of same character with four variations with respect to its position portrays a great challenge in segmentation of Arabic cursive script. Other variation of Arabic script where context is important are Urdu and Persian. All Arabic script languages are written from right to left and have various representation regarding each character depending on its position in a word. This paper proposed a transfer learning approach designed for handwritten Urdu character recognition and take a benefit of MNIST pre-trained network to apply its learning experience as both datasets encompass handwritten strokes.

In this proposed work, the performance of conducted experiments is analyzed by using pre-trained network and by keeping few hidden layers freeze while tune the rest. Another purpose of choosing handwritten Urdu text is to apply transfer learning idea because of cursive and complex nature of handwritten Urdu text which is yet to be exposed to state-of-the-art techniques. In contrary, the work on synthetic Urdu has been extensively researched and reported since recent years as witnessed in [5], [6], [10], [11], but handwritten Urdu character recognition still in it's very beginning. The handwritten Urdu text usually written in Nasta'liq style. Being a cursive language, it has same issues as it's ancestor languages suffered. Unlike it's ancestor Arabic, there is not prominent OCR available for Urdu script commercially. Although OCR for printed Arabic with higher accuracies have been reported. Hence, any successful research effort in this direction have higher commercial value and pave a path for suggesting Urdu OCR [5], [6]. The complexities of Urdu script have not attracted attention from pattern recognition community in recent past, as it is emerging nowadays.

This paper primarily focuses on an adaptive Multi Dimensional Long Short Term Memory networks (MDLSTM)'s structure that present powerful training algorithm by inclusion of transfer learning. The proposed model is largely motivated by recently presented research [6] to solve printed Urdu text recognition. The proposed system extracts the lower level features from MNIST large database by Convolutional Neural Networks (ConvNets). Unlike the architecture presented in [6], the training of Urdu text lines initiated by convolution with pre-trained network of MNIST database. The idea is to have well trained model equipped with rich knowledge about target tasks that can re-use to train other models. In this manuscript, the convolution of MNIST features is performed with proposed dataset images. In this way the learning experience of MNIST dataset is transferred to learn the patterns of handwritten Urdu samples.

As mentioned earlier, to address the problem of segmentation by considering the complex Urdu script,the implicit segmentation based MDLSTM approach is used. Two different target tasks are identified in presented work as follows,

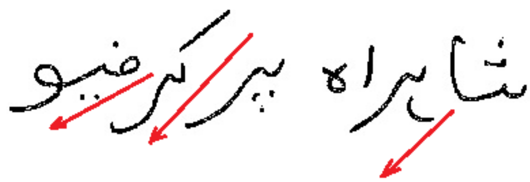


FIGURE 1. Handwritten Urdu text, the arrow head indicates direction and diagonal characteristics of Nasta'liq script.

1) The first task is to use pre-trained network and reconstruct or learn the new network designed for handwritten Urdu text.
2) By freezing few hidden layers, the target is to search for optimal learning for the purpose to explore and note down transfer learning performance empirically.

The prime concern is to explore the transfer-ability effect in learning the cursive script. By transferring the knowledge, the significant performance improvement is achieved by using UNHD dataset. The main contributions of this study encompasses as below,

- This paper is an attempt to demonstrate the transfer learning ability and explores that how one handwritten sample MNIST) can effect the learning of another handwritten but cursive script (UNHD).
- The proposed study presented first time for handwritten Urdu samples to evaluate the performance of presented idea by manipulating the hidden layers. The hidden layers intend to learn the process by finding an optimal layer.
- The conducted experiments demonstrated that proposed architecture outperforms the deep neural network approach based on Convnets and MDLSTM networks.

A. CHARACTERISTICS OF URDU

As mentioned in [12], Urdu has 39 identified letters. Being a descendant of Arabic script, it shares similarity with Arabic script in terms of complexities. The Urdu text is written from right to left having joined adjacent characters within a word. However, characters may appear in four different positions in a word.

The Urdu character positions are identified as first, middle, final or an isolated character. The appearance of same character will change depending on its position with reference to single word. These positions dictate the character about its pattern according to it's position in a word. The main characteristic of Urdu is its diagonally written text which does not appear on fixed base line as depicted in Figure 1. The Urdu text is disintegrated into words, ligatures and characters as shown in Figure 2, the three words are identified, that were segmented into ligature as clearly depicted in the figure without diacritical marks. In Urdu, combination of characters make ligature and if join ligatures together a word can form. The explicit tools are not suitable for segmentation of this complex script correctly. Most of recently published

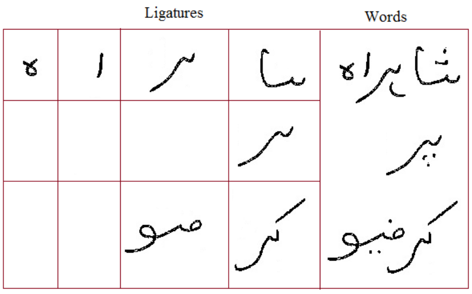


FIGURE 2. **Disintegrated handwritten Urdu text into words and ligatures without diacritical marks on left side.**

work used implicit segmentation methods for Urdu script as reported by [5], [10], [11]. The appearance of character depends on the context as it comes in a word. Although Urdu text is written from right to left but Urdu numerals are written from left to right which makes the Urdu script as bidirectional language. The glyph shape in Urdu is always context sensitive. The detailed characteristics about Urdu language can be seen in [12].

This paper is organized into various sections. The literature review is summarized in section 2, while transfer learning based recognition system is presented and discussed in section 3. This section further explains the evaluation process of MNIST dataset, it also highlights the process of using ConvNets as a feature extractor, while training on pre-trained network architecture is also detailed in this section followed by dataset description. Experimental study is explained in detail in section 4. The analysis and discussion followed by conclusion are presented in section 5 and 6 respectively.

II. PREPARATORY STUDY

This section exhibits a brief overview of few related work presented in recent years. The initial discussion is started from such articles that extend the usage of MNIST dataset for evaluation of their proposed techniques followed by description of RNN based techniques applied on printed Urdu database. The details about how transfer learning based architecture supports quick learning and generate stat-of-the-art results, also provided in this section.

The MNIST database is considered as a benchmark and use for validating the various proposed techniques. The handwritten digit recognition is represented by Lauer *et al.* in [13]. They proposed ConvNets to extract features without taking prior knowledge of data into consideration. The extracted features from ConvNets were merged with the help of support vector machine (SVM). By considering the elastic distortion on the MNIST database, the authors reported 99.46% accuracy. In another identical study, Niu and Suen [14] proposed ConvNets architecture as feature extractor from raw images and use SVM to classify the MNIST digit database. They achieved 94.40% accuracy by using their proposed hybrid model.

In cursive script analysis either provided in handwritten or in printed form, the context plays a crucial role in learning the sequence. The sequence learning does not only mean to learn the long sequences but also to assign different labels to similar patterns. The number of proposed research employed sequence learner like Hidden Markov Model (HMM) [15], Conditional Random Field (CRF) [16], and Recurrent Neural Network (RNN) approaches as a sequence learner [10], [12], [17]. The prime motive is to obtain accurate sequence labeling. The aforementioned approaches have successfully been applied on different nature of sequence learning problems.

ConvNets is usually used to extract the lower features from given input image. In a very interesting work presented by Socher *et al.* [18], their proposed architecture is a combination of ConvNets and RNN for object recognition. ConvNets was used for feature selection followed by RNN for classification. Similar technique used for audio data is presented by Anand and Verma in [19]. They exploit the LSTM feature of sequence learning and applied it on emotion classification from image and audio data and obtained 70% accuracy in predicting emotions. The multidimensional model of LSTM was implemented by Bezerra *et al.* [20] to improve character recognition accuracy.

A transfer learning architecture based on RNN was presented by Chen *et al.* [1]. The visual features were extracted from ultrasound images using ConvNets that reported very encouraging results by using knowledge transfer. Another very interested work using transfer learning approach is presented by Ciresan *et al.* [21]. They proposed Latin and Chinese character recognition using deep neural network. They used pre-trained network of Chinese database by keeping n-layers freeze to recognize upper case Latin characters. They obtained good results on their proposed system.

III. TRANSFER LEARNING BASED RECOGNITION SYSTEM

By following the notion of transfer learning, the proposed architecture is divided into two halves. In first half, MNIST database is evaluated as source function. The proposed idea is implemented first on MNIST database. In second half, the emerged features as extracted by the earlier process were convolved with the features extracted by the same process on Urdu text line. In the following, the proposed architecture is discussed in detail.

A. EVALUATION OF MNIST DATABASE

The MNIST database believes to be a large database having handwritten numeral which is commonly used for evaluation of various pattern recognition tasks. Since it's inception it achieved remarkable accuracy rate [2], [22], [23]. One common thing in MNIST and UNHD databases are handwritten strokes. In order to realize proposed assumption, the high level transient features are extracted from 60, 000 numeral samples by ConvNets. The extracted features are feed-forwarded to MDLSTM system. Figure 3 shows detail depiction of MNIST database evaluation system.

B. CONVNETS AS A FEATURE EXTRACTOR

The ConvNet is widely proposed for feature extraction as it extracts detailed features from input images [4], [6], [14].

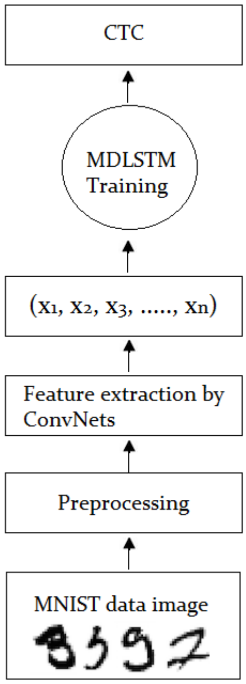


FIGURE 3. Basic architecture of MNIST's samples learning.

The potential of ConvNets for feature extraction is exploited. By keeping in view the implied complexity attached to Urdu script, the idea of implicit segmentation emphasized. The ConvNet requires large amount of labeled data for training which seems not to be handled manually. The entire hypothesis is to consider MNIST and UNHD databases because both database contain handwritten strokes that shares the same complexities. The five layer architecture of ConvNets is proposed. Initially, preprocessing was done by removing noise from given image and standardize the representation of an image after converting it into grayscale. The standardization of an image is performed by dividing the current pixel value with total number of image pixels as represented in equation 1.

$$S = \frac{X_{CurrentValue}}{T_p} \tag{1}$$

whereas, T_p is a total number of involved pixels of an image.

In this way all values will be represented in a range (0-1). The Urdu word's images were cropped from Urdu text and later to standardize the image into 96×96 size. The image was resize using built-in python library. The standardized image size is obtained by taking an average of Urdu word representation according to involved pixels. In each feature map, every neuron is mapped according to small 5×5 region of input image. The 5×5 window size is selected to cover maximum pixels values. The connection from input image to hidden layer is established through local receptive field called a filter size. Each neuron in a layer shares a same bias value. As single feature map does not cover the intensive features, so the process is further delegated for an intention to have variety of features against each given image. A feature map is defined by it's share weight and a bias value, mathematically this relation can be represented as,

$$\alpha\left(d + \sum_{e=0}^{4}\sum_{f=0}^{4} W_{e,f} A_{j+e,k+f}\right) \tag{2}$$

α is neural activation sigmoid function while d is a shared value of bias. $W_{e,f}$ represents filter or kernal weight which depends on filter size whereas, A represents the input activation at point (x,y). In proposed architecture, 96 feature maps are defined by 5×5 set of shared weights with single bias value which make the kernal size 26. From each feature map, 25 shared shared weights are required. The convolutional layer is represented by 96 feature maps that make the total of $96 \times 26 = 2496$ parameters that define the convolutional layer.

As a result a network can detect 96 different kind of features at convolutional layer 1 as represented in Figure 6. The aforementioned process will continue for the next four hidden layers.

In order to condense the extracted features, pooling is an essential step in ConvNets. Here, L2 pooling strategy is used, which takes square root of 5×5 region's sum. The extracted condensed features F_1 from MNIST database images are convolved with Urdu handwritten input image I_w of UNHD database.

The last layer is fully connected network, this layer connects every neuron from L2 pool to output neurons. The prime goal is to find the optimal performance on intrinsic Arabic cursive script.

C. TRAINING ON PRE-TRAINED NETWORK USING MDLSTM

The pre-trained network on MNIST datset is used to access the potential of Urdu handwritten samples.

1) MULTIDIMENSIONAL LSTM

For every Handwritten Urdu word w represented as input I_w, the processing is initiated by standardizing the image skeleton and computed feature maps M, by using ConvNet as described in previous section. Every feature map includes series of distinctive features. The MDLSTM is a variant of RNN approach [17]. As in RNN, the hidden layer neurons are recursively connected to itself and also with subsequent neurons in next layer hence to have an effect of memory. But when the problem grows complex then memory becomes fade gradually which accelerate to handle the problem with long sequences. This problem advised by solution in a shape of LSTM which replaced hidden neurons with memory block and it's multiplicative units. The memory is regulated through these blocks for a specific period of time and then make it forget. The LSTM architecture demonstrated evident success for many sequence learning problems addressed recently [10], [11], [17], [24]. As represented in Figure 5, against the corresponding input image I_w, the extracted image skele-

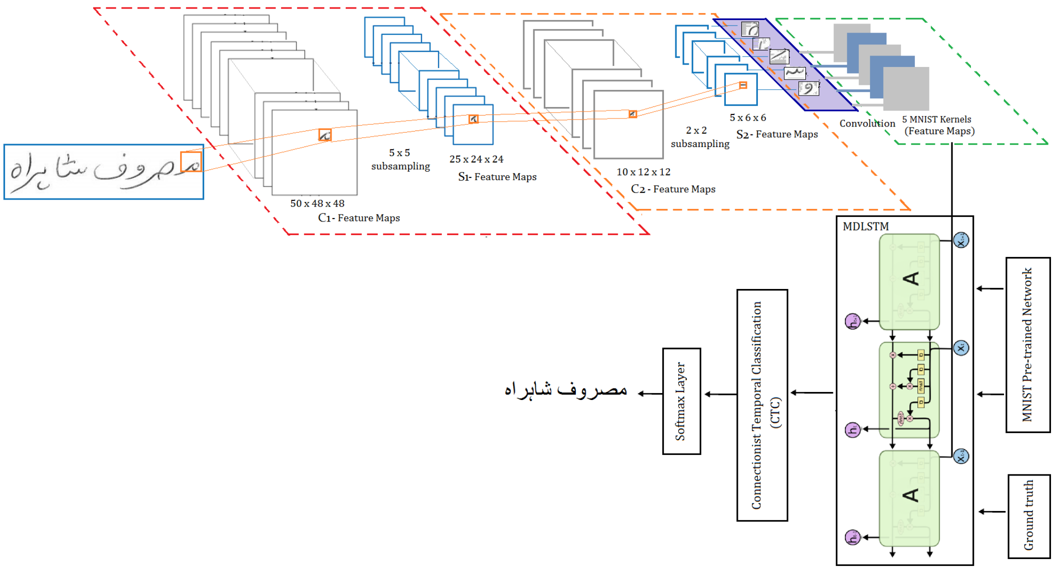


FIGURE 4. Transfer learning based cursive handwritten text recognition.

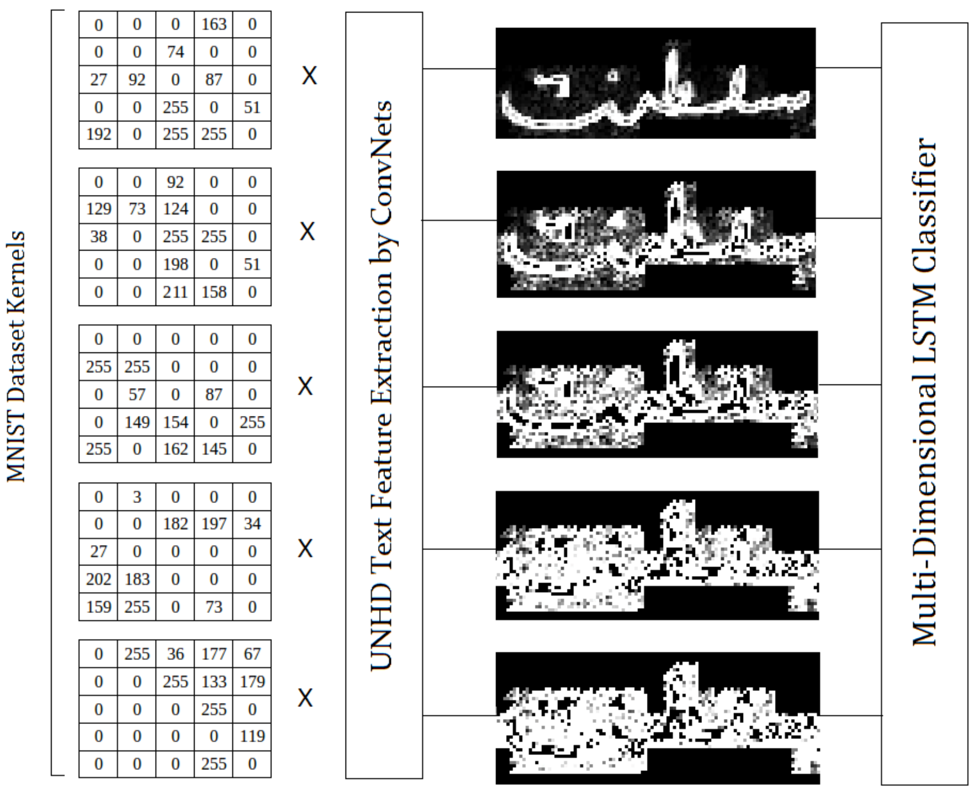


FIGURE 5. MNIST convolutional matrix with UNHD database images.

ton convolved with five kernels i.e., $F_1 - F_5$ (empirically selected during MNIST training) are used as features and pass them to MDLSTM classifier as represented in Figure 6. The 5×5 filter size was move (one at current point in time) over the image for the purpose to extract feature vector and pass them to the classifier along with corresponding ground truth.

2) TRANSFER LEARNING

By initiating the training of UNHD samples from MNIST pre-trained network, the current learning network is taking an advantage of handwritten strokes exist in MNIST dataset that incorporates their learning experience during the training of UNHD samples. To realize this assumption, the MDLSTM network is described by 10, 20, 40, and 60 hidden memory

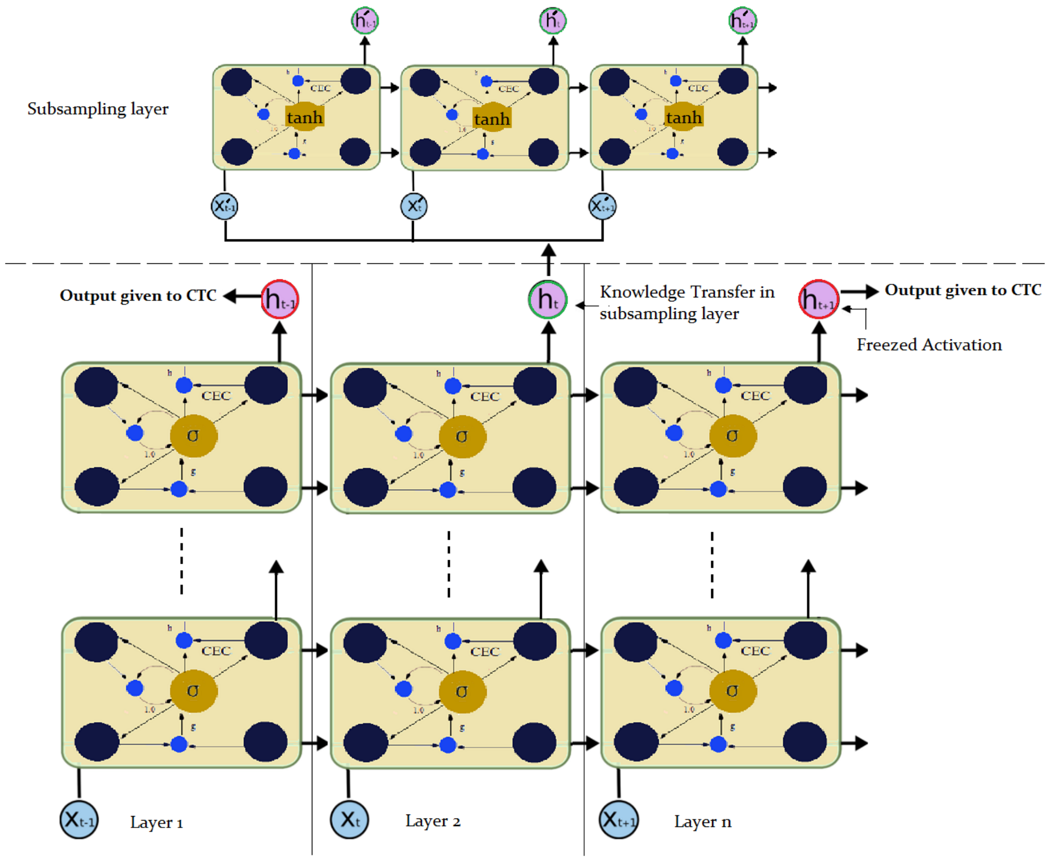


FIGURE 6. Experimental variation in which layer 2 has a subsampling layer. Each neuron in a memory block is activated with sigma. The layer 1 and n is transferring output to CTC directly, whereas memory blocks in sub-sampling layers has the neuron cells which were activated with tanh.

blocks in 5 layer architecture that investigate the given input in all four directions having *tanh* as an activation indicator. Mathematically it can be represented as,

$$tanh(z) = \left(\frac{e^z - e^{-z}}{e^z + e^{-z}}\right) \tag{3}$$

The underlying architecture is fully connected.

By taking the benefit of transfer learning, the experimental study was conducted by dividing the hidden layer 2 and 4 into two sub-sampling layers having size of (10, 20)(20, 30)(30, 40)(40, 50). In this experimental settings, the training will begin with the MNIST pre-trained network. All subsampling layers are feed-forward network that use sigmoid σ as an activation. Mathematically it is represented as,

$$\sigma(z) = \left(\frac{1}{(1 + exp(-z))}\right) \tag{4}$$

Although *tanh* believes to be a good choice to optimize the training because of its faster convergence ability but some time network requires more time to learn the complex data. Therefore, σ is used in sub-sampled layers. Mathematically this relation can be summed up as,

$$A_m = \left(\frac{e^z - e^{-z}}{e^z + e^{-z}}\right) + \left(\frac{1}{(1 + exp(-z))}\right)^n \tag{5}$$

A_m shows the mutual activation whereas, n represents size of sub-sampling layers. The features were collected into 4 x 4 hidden blocks which were passed to feed forward layers having *tanh* summation activation units as represented in Figure 6.

Another experiment was performed by keeping hidden layer 1 and 3 freeze. In this experiment, the training starts by using pre-trained MNIST dataset. As MDLSTM network's hidden layer 2 and 4 have sub-sampled layers which intends to convolve the learning of handwritten cursive script in detail with the experience of MNIST trained network. But in this experiment the sub-sample layers have *tanh* feed forward network. All hidden layers of MDLSTM network were disintegrated into single sequence vector. In this paper, the multi-dimensional architecture of LSTM network is employed for proposed work. In the feed forward pass, the activation a of each cell in LSTM memory block is computed in four directions. Each LSTM memory block computes the activation at input gate a^t_λ, forget gates $a^t_{\alpha,dim}$, internal cell activation

a^t_p,and output gates a^t_δ. The LSTM memory block computes activation in all dimensions *dim*. The disintegrated sequence were fed to Connectionist Temporal Classification [25] which is a part of output layer that comprise same labels as defined as target symbols and one extra label to deal with undeclared or space symbols in a given pattern. Each element in the labels have path associated to input sequence I_x, where x is determined by the sequential path. The gradient descent optimizer is used to reduce the loss obtained by CTC loss function.

The LSTM memory block computes the activation as follows, whereas M_b represents the memory block.

$$M_b = f(a^t_\lambda) + f(a^t_{\alpha,dim}) + u^t_p + f(a^t_\delta) \tag{6}$$

The input gate computes the activation with all hidden memory units J in equation 7, the refined form is represented in equation 8.

$$a^t_\lambda = \sum_{i=0}^{k} x^t_i w_i + \sum_{dim=1}^{n} \left(\sum_{j=1}^{J} b^{t_{dim}}_j w^{dim}_j + \sum_{p=1}^{P} w_p q^{t_{dim}}_p \right)^m \tag{7}$$

R is represented as an activation computed at each cell.

$$R^t_{\alpha,dim} = f(a_\lambda)^t \tag{8}$$

The equations 9 and 10 represents the calculation performed at forget gate.

$$a^t_{\alpha,dim} = \sum_{i=1}^{k} x^t_i w_{i(\alpha,dim)} + \sum_{dim'=1}^{n} \sum_{j=1}^{J} b^{t_{dim'}}_j w^{dim'}_{j(\alpha,dim)} \tag{9}$$

$$R^t_{\alpha,dim} = f(a^t_{\alpha,dim}) \tag{10}$$

The computation at cell is represented by equation 11 with recurrent connection as shown in equation 12

$$a^t_p = \sum_{i=1}^{k} x^t_i w_{ip} + \sum_{dim=1}^{n} \sum_{j=1}^{J} R^{p_{dim}}_j w^{dim}_{jp} \tag{11}$$

$$u^t_p = R^t_{\alpha,dim} g(a^t_p) + \sum_{dim=1:t_{dim}>0} u^{t_{dim}}_p \sum R^t_{\alpha,dim} \tag{12}$$

The calculation performed at output gates is presented in equation 13 and 14

$$a^t_\delta = \sum_{i=1}^{k} x^t_i w_{i,\delta} + \sum_{dim=1}^{n} \sum_{j=1}^{J} R^{p_{dim}}_j w^{dim}_{j\delta} + \sum_{p=1}^{P} w_{p\delta} s^t_p \tag{13}$$

$$R^t_\delta = f(a^t_\delta) \tag{14}$$

The LSTM memory blocks in sub-sampling layers are the product of all LSTM memory blocks involved in feedforward and backward direction as represented in equation 15, where n is the number of LSTM blocks in subsmpling layers.

$$\prod_{s=0}^{n} \left(f(a^t_\lambda) + f(a^t_{\alpha,dim}) + u^t_p + f(a^t_\delta) \right)^n \tag{15}$$

The proposed network was trained by using gradient descent optimizer with a learning rate of 1 x 10^{-4} *and* 1×10^{-3} but the best training network was reported on prior learning rate which is selected empirically to report the best training accuracy. The total network weights are 2, 75, 391. When there is no improvement reported in error rate after consecutive 30 iterations then the training stops.

TABLE 1. Extended UNHD dataset description.

Detail	Description
No. of Authors	500
Textlines	10000
No. of words	312000

TABLE 2. MDLSTM parameter details used for training UNHD using ConvNets feature selection.

Parameter's detail	Values
Hidden Size	20, 60, 100
Sub-sample size	10, 40
Hidden block size	2×6, 2×6
Input Block Size	10×1
Learning rate	1e-4, 1e-3
Momentum	0.9
Total network weight	2,75,391

D. DATASET

The handwritten Urdu text was collected from various segments of people without imposing any special constraint, the people including school going children, college or universities students and office going professionals. Figure 7 represents Urdu handwritten samples having variations in writing styles which is categorized as an exceptional problem for recognition systems. In this scenario, the recognition accuracy relies more on how perfectly it can segment the images. More variations have been captured in an extension of UNHD dataset as writers were asked to write in their natural handwriting without baselines.

The taken samples also include scrambled and overwritten Urdu text.

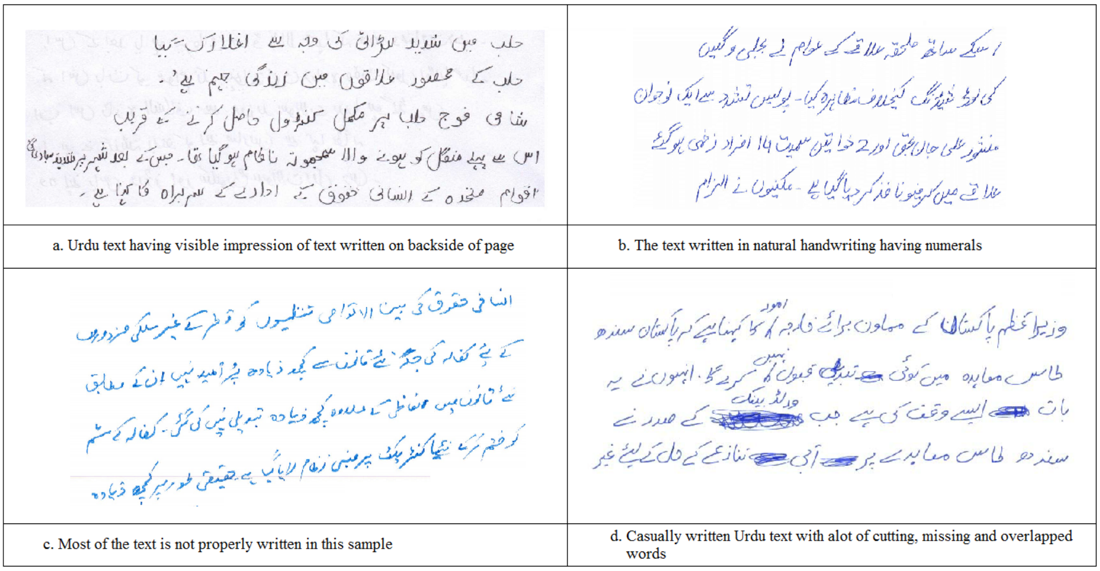
In UNHD dataset, the Urdu text is taken from 500 authors and each author wrote 5 text lines with 10 words per line. In this way almost 10, 000 text lines are collected having 312000 words as mentioned in Table 1. The experiments are conducted on 150 Urdu text lines written by 30 authors having 1509 words and 2851 ligatures. The dataset is divided into train set and test set with same ratio as specified by graves [26].

IV. EXPERIMENTAL STUDY

To vindicate the proposed system, the potential of demonstrated experiments is assessed by tuning the parameters that could influence the training curve. Another experiment is conducted by freezing the activation of few layers while on the other hand, having a sub-sampled layers of non-freezed layer in MDLSTM network architecture. The brief description about conducted experiment is discussed below.

A. BY TUNING PARAMETERS

The training curve is observed by employing different parameter values. The parameter values of MDLSTM classifier are represented in Table 2.



a. Urdu text having visible impression of text written on backside of page

b. The text written in natural handwriting having numerals

c. Most of the text is not properly written in this sample

d. Casually written Urdu text with alot of cutting, missing and overlapped words

FIGURE 7. UNHD data samples having implicit variations.

TABLE 3. Tuning the network by keeping focus on selected parameter's values.

Input block	Learning Rate	Hidden Layer Size	Training Error	Testing Error
10×1	1e-4	20	0.23	0.25
10×1	1e-4	60	0.17	0.20
10×1	1e-4	100	0.05	0.71
10×1	1e-3	20	0.31	0.39
10×1	1e-3	60	0.29	0.32
10×1	1e-3	100	0.13	0.19
16×2	1e-4	20	0.18	0.21
16×2	1e-4	60	0.11	0.16
16×2	1e-4	100	0.03	0.07
16×2	1e-3	20	0.17	0.20
16×2	1e-3	60	0.15	0.18
16×2	1e-3	100	0.09	0.11

The advantage of providing distinct values to training network is to assess the learning of network by experimenting with different values given empirically. The images are handwritten cursive samples which aim to be trained on pre-trained network hence to transfer the learning of MNIST samples to UNHD data. The word transcription is divided into 174720 trainset, 74880 validation set and 62, 400 test set.

The input block size, sub-sample block size, hidden layer size and learning rate are the parameters which are examined by providing different range of values during the experiments as explained in Table 3.

As observed from the table that best accuracy was reported on learning rate 1×10^{-4} while the hidden layer was 100. As noticed that learning rates and number of hidden layers impacted on recognition accuracies. The input block is constant for all performed experiments. The training network starts training on pre-trained MNIST dataset by using gradient descent optimizer which exploit the learning by UNHD handwritten data.

TABLE 4. Error rate on clean and complex data.

Details	Clean Test data	Complex Test data
Number of samples	58827	3573
Error rate	0.063%	0.087%

Several parameters were contributed to investigate the best performance influenced by given values. In this experiment, UNHD dataset is considered without its complex data. The complex data contains Urdu scrambled text appeared with cutting and overwritten words. The *tanh* was used for activation functions. The best accuracy was 91.3% on test set after 370 epochs while hidden layer size was 100. Table 4 depicts the accuracy rate achieved on complex and clean handwritten data.

B. BY INCORPORATING FREEZE HIDDEN LAYERS

Another variation is to investigate the performance by incorporating sub-sampling layers while to freeze the other hidden layer's activation.

TABLE 5. Experimental results with different models and training methods.

Details	Sub-sampling size	Input Block size	Precision	Recall	F-Measure
TL-MDLSTM-1	10,20	10×1	0.63	0.57	0.59
TL-MDLSTM-1	20,30	10×1	0.69	0.58	0.63
TL-MDLSTM-1	30,40	10×1	0.73	0.65	0.69
TL-MDLSTM-1	40,50	10×1	0.68	0.56	0.61
TL-Deep-MDLSTM	10,20	10×1	0.74	0.68	0.71
TL-Deep-MDLSTM	20,30	10×1	0.73	0.69	0.71
TL-Deep-MDLSTM	30,40	10×1	0.85	0.78	0.81
TL-Deep-MDLSTM	40,50	10×1	0.78	0.75	0.76
TL-MDLSTM (Complex)	10,20	10×1	0.77	0.71	0.74
TL-MDLSTM (Complex)	20,30	10×1	0.79	0.70	0.74
TL-MDLSTM (Complex)	30,40	10×1	0.82	0.72	0.77
TL-MDLSTM (Complex)	40,50	10×1	0.79	0.72	0.75
TL-MDLSTM-1	10,20	16×2	0.82	0.78	0.80
TL-MDLSTM-1	20,30	16×2	0.84	0.78	0.81
TL-MDLSTM-1	30,40	16×2	0.90	0.84	0.86
TL-MDLSTM-1	40,50	16×2	0.88	0.83	0.85
TL-Deep-MDLSTM	10,20	16×2	0.79	0.75	0.77
TL-Deep-MDLSTM	20,30	16×2	0.88	0.82	0.84
TL-Deep-MDLSTM	30,40	16×2	0.93	0.84	0.88
TL-Deep-MDLSTM	40,50	16×2	0.85	0.80	0.82
TL-MDLSTM (Complex)	10,20	16×2	0.66	0.59	0.62
TL-MDLSTM (Complex)	20,30	16×2	0.65	0.49	0.56
TL-MDLSTM (Complex)	30,40	16×2	0.67	0.47	0.55
TL-MDLSTM (Complex)	40,50	16×2	0.63	0.45	0.52

As explained earlier that 5 layer architecture is employed which comprise 20, 60 and 100 hidden memory units. As MNIST pre-trained network is used to get the benefit of transfer learning. In this experimental variation, the training starts from pre-trained network but the gradients computed at layer 1,3 and 5 kept freeze. The training of proposed architecture is observed by inclusion of sub-sampling layers implemented at layer 2 and 4.

The detail about number of sub-sampled hidden memory units are presented in Table 5. The dataset evaluation is categorized into three sections according to the difficulty level of acquired samples. TL-MDLSTM-1 is a first variation of MDLSTM transfer learning architecture which evaluates the UNHD dataset. Every experimental variation considers four sets of sub-sampling size i.e., (10, 20)(20, 30)(30, 40)(40, 50) using input block sizes 10×1 and 16×2.

The second variation is TL-Deep-MDLSTM that covers large amount of UNHD dataset samples and also handwritten samples of UCOM dataset [27] take in account for evaluation purpose. The limited number of complicated handwritten Urdu text was evaluated as third experimental study. The challenging text was contributed by 65 authors. The experiments are performed by considering input block size and sub-sampling size as presented in Table 5.

The accuracy is measured by computing precision and recall of three experimental variations according to the detail provided in Table 5. The ROC curves of each experimental study carried out when input block size was 10×1, as represented in Figure 8. Each curve is determined by considering input block size. As observed in Figure 8, the best result was obtained when sub-sample size was (30, 40).

In Figure 9, the precision and recall were mapped by keeping in view the three experimental studies. In each experiment, it is witnessed that good performance was measured when sub-sampling size was (30, 40), but for the third experiment as it contains complex Urdu handwritten data that's why the overall accuracy is comparatively low whereas, the best curve was obtained when sub-sampling size was (10, 20).

The comparison of obtained results using two input block sizes represent comparatively good performance especially when input block size was 16×2. But for the third experimental study, the good accuracy is reported on 10×1.

V. ANALYSIS AND DISCUSSION

In this paper, the proposed idea is evaluated on two main architectures. By following first architecture, the experiment was conducted by using various parameters as indicated in previous section. In second proposed architecture the experiment is performed by freezing the activation of few hidden layers while do the sub-sampling of rest as indicated in Table 5. With this constraint, the experiments were further divided into three different experimental studies. Each experiment was performed by using two different input block size. The input block size is described through division of input image given to MDLSTM network before classification. The assumption is that if input image is divided into more smaller chunks then it will be easier for classifier to learn the complex pattern efficiently. The assumption is investigated by defining the input block size 10×1 and 16×2.

Another very relevant and important aspect is regarding the training on each experimental study originates from MNIST pre-trained network by assuming the fact as mentioned earlier, i.e., MNIST and UNHD contains handwritten strokes.

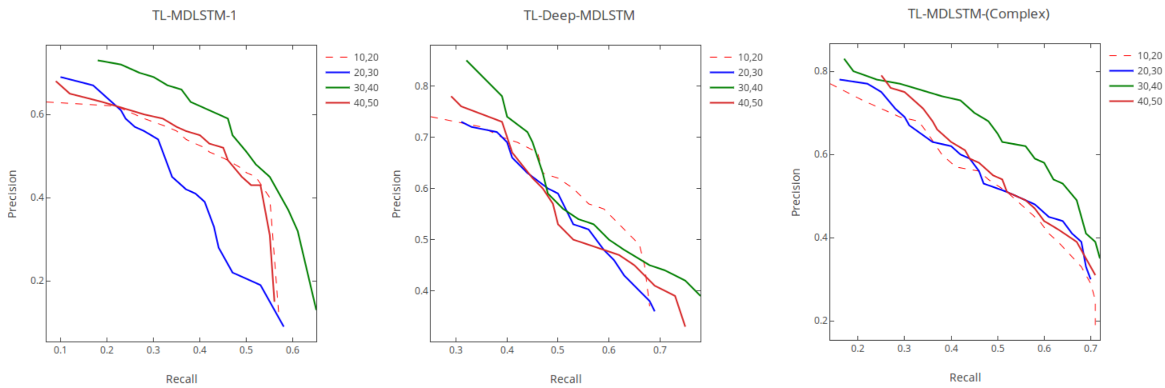


FIGURE 8. The ROC curves obtained from three experimental study on four different sub-sampling sizes when input block size was 10 × 1.

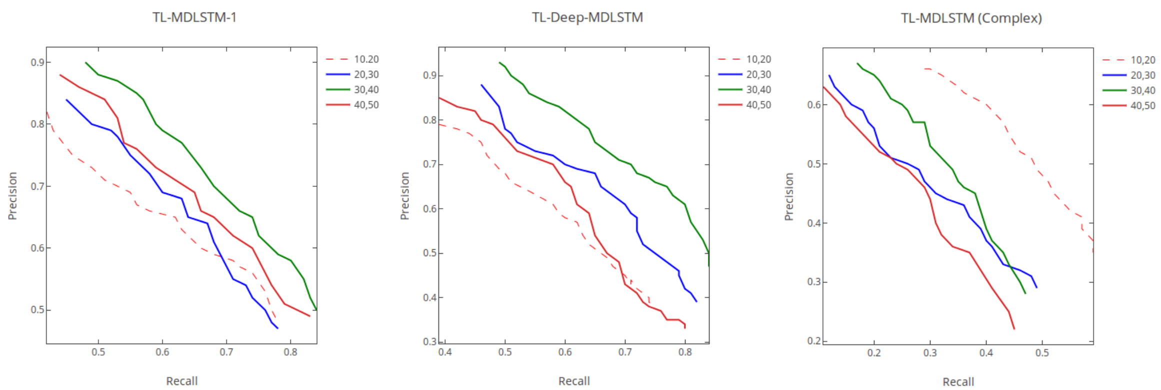


FIGURE 9. The ROC curves obtained from three experimental study on four different sub-sampling sizes when input block size was 16 × 2.

The performance of MNIST using UNHD dataset is assessed by presented work. By implementing the proposed assumption, the achieved performance is very encouraging.

The UNHD is an extension of UCOM dataset [27] which covers broad range of Urdu words and ligatures. The experiments are performed on the extension of UCOM dataset and also on whole UNHD dataset including UCOM data samples. The extension of UCOM dataset also includes complicated Urdu text that includes scrambled Urdu text. Figure 10, shows various samples of Urdu words that are evaluated by the proposed architecture.

It is observed that Urdu research gained popularity during recent years in particular [6], [10], [12]. If comparison takes place of presented work with some significant published Urdu research that have been reported to date, then researcher may know the fact that work on handwritten Urdu is very minimum in comparison to printed or synthetic data. Although, presented work can be treated as a benchmark work as far as Urdu handwritten text recognition is concerned, but the comparison may be possible with the performance of recognition systems devised for printed or synthetic Urdu data.

As reviewed from various available Urdu work, most of the Urdu research work focused on implicit segmentation due to complex nature of Urdu script and used LSTM based classifier which is suitable for context learning as the cursive pattern of Urdu script suggests. This paper also used deep-MDLSTM architecture to learn the context of complex data.

The publications on Urdu text recognition are divided into handwritten and synthetic data as summarized in Table 6. The most prominent initial work on synthetic Urdu text recognition was proposed by [5] and [28]. They extracted x-height features from given input and trained them by LSTM classifier. They reported results by using 1-dimensional bidirectional LSTM architecture. Reference [28] reported 95.4% character level accuracy on UPTI dataset, whereas [5] evaluated their proposed method on Urdu-Jang dataset and achieved 88.94% and 88.79% accuracy on without position and with position details of characters. Another interesting work on synthetic Urdu is reported by [10]. They evaluated their proposed methods on UPTI data set with multidimesional LSTM architecture. The features they considered are number of horizontal and vertical edge pixels, number of foreground pixels, sum of intensity values, mean and variance of horizontal and vertical projection, energy contrast, correlation and homogeneity. They also reported variations in

TABLE 6. Comparison of printed and handwritten Urdu text research.

Research Work	Proposed Technique	Dataset Detail	Reported Accuracy (%)
Ul-Hassan et al [28]	1-Dimensional BLSTM	UPTI (synthetic)	95.4
Ahmed et al [5]	1-Dimensional BLSTM	Urdu-Jang (synthetic)	88.94
Naz et al [10]	Multidimensional LSTM	UPTI (synthetic)	96.4
Naz et al [6]	Recursive CNN-MDLSTM	UPTI (synthetic)	98.12
Naz et al [11]	Zoning Features + 2D-LSTM	UPTI (synthetic)	93.38
Ahmed et al [27]	1-Dimensional LSTM	UCOM (handwritten)	73.02
Naz et al [17]	Statistical Features + MDLSTM	UPTI (synthetic)	94.97
Ahmed et al [24]	1-Dimensional LSTM	UNHD (handwritten)	92.03
Proposed method	Transfer Learning and MDLSTM	UNHD (handwritten)	93.0

FIGURE 10. The original images are presented with their corresponding pixelized gray-scale images. The misclassified characters are represented in red color whereas, the target words are represented in green text.

parameter value to assess the impact on learning network. The parameters they considered are learning rate, sub sampling size, and hidden layer sizes. The reported accuracy at character level on UPTI dataset which is 96.4%. The covolutional neural network based multidimensional LSTM architecture is proposed by [6]. The network was trained with convolutional features using LSTM architecture on 44 different classes and got highest accuracy which is 98.12% at character level. The printed Urdu recognition by zoning feature based 2-DLSTM is proposed by [11]. The sliding window approach was used which slide 3×30 size window over an image and segment the window in 10 zones. They performed experiments by considering different zones. They achieved 93.38% accuracy on largest zone. The statistical feature extraction approach was proposed by [17] with multidimensional LSTM as a classification technique. The statistical features included density, intensity, means and variance of horizontal projection, center of gravity calculated for x and y direction. They reported 94.97% recognition rate.

In most of the presented work, proposed feature extraction and classification methods were evaluated on Urdu printed UPTI dataset. The LSTM has also been applied on handwritten Urdu text. One such work is reported by [27]. They proposed handwritten Urdu dataset called UCOM. Initially, data was collected from 100 students on 6 papers having 8 text lines on each paper. They removed base lines and performed text line segmentation on collected data manually. The dataset was divided into 53248 characters with 62000 words that appears in repetition with different writing styles. They also proposed usage recommendations and possible research scenarios in which UCOM dataset can be evaluated. They applied 1-DLSTM classifier on small subset of UCOM data to assess it's applicability. They reported 73.02% accuracy on handwritten Urdu characters. The extension of UCOM dataset is UNHD which is an abbreviation of Urdu Nastaliq Handwritten Dataset as reported in [24]. The acquired samples have been increased from 48 unique text lines to 700 unique text lines including Urdu numerals and Urdu constraint handwritten samples. They have compiled more than 4000 handwritten Urdu text lines. They proposed 1-DLSTM architecture to perform experiments and reported 92.03% accuracy on character level.

This paper helps in investigating the potential of MNIST handwritten strokes with UNHD handwritten data by considering that both have strokes information which can be the common point of interest during training which lead to best performance by keeping Urdu text in focus. The extended version of UNHD dataset is more challenging handwritten Urdu text. Due to variations in writing styles, different writers were asked to write the same contents but with little variations. This added more variety in terms of ligatures and displayed very encouraging results in terms of recognition accuracy in comparison to other work as reported in Table 6. In printed data, the implicit variations imposed by writer through handwriting is not considered as observed in handwritten Urdu dataset.

This constraint can over come with large number of training samples that requires more concentration and time.

VI. CONCLUSION

This paper presented a significant research in the area of handwritten Urdu text analysis that yields benchmark results on proposed architecture. The recognition rate is considered as a benchmark by keeping in view the complex structure of handwritten Urdu script. The proposed notion is enormously noteworthy in terms of handwritten Urdu text research and it's practical applications.

The strength of MNIST database pre-trained network is exploited which transfer the learning experience of handwritten strokes to learn Urdu handwritten data. The MDLSTM architecture is proposed as a classifier for transfer learning to recognize Urdu handwritten text.

The experiments have been performed in detail to alienate the fact that handwritten strokes in one database may help to improve the recognition rate or learn the handwritten strokes appeared in another database. The inclusion of transfer learning enhances the capability of sequence learner LSTM networks which provides state-of-the-art results on complicated data analysis.

REFERENCES

[1] H. Chen, Q. Dou, D. Ni, J.-Z. Cheng, J. Qin, S. Li, and P.-A. Heng, "Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI-2015* (Lecture Notes in Computer Science), vol. 9349. Munich, Germany: Springer, Oct. 2015, pp. 507–514.

[2] L. Deng, "The MNIST database of handwritten digit images for machine learning research," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, Nov. 2012. [Online]. Available: http://dblp.uni-trier.de/db/journals/spm/spm29.html#Deng12

[3] B. Zhu and M. Nakagawa, "Segmentation of on-line freely written Japanese text using SVM for improving text recognition," *IEICE Trans. Inf. Syst.*, vol. E91-D, no. 1, pp. 105–113, 2008. [Online]. Available: http://dblp.uni-trier.de/db/journals/ieicet/ieicet91d.html#ZhuN08

[4] X. Ren, K. Chen, and J. Sun, "A CNN based scene chinese text recognition algorithm with synthetic data engine," 2016, *arXiv:1604.01891*. [Online]. Available: https://arxiv.org/abs/1604.01891

[5] S. B. Ahmed, S. Naz, M. I. Razzak, S. F. Rashid, M. Z. Afzal, and T. M. Breuel, "Evaluation of cursive and non-cursive scripts using recurrent neural networks," *Neural Comput. Appl.*, vol. 27, no. 3, pp. 603–613, 2016. doi: 10.1007/s00521-015-1881-4.

[6] S. Naz, A. I. Umar, R. Ahmad, I. Siddiqi, S. B. Ahmed, M. I. Razzak, and F. Shafait, "Urdu Nastaliq recognition using convolutional–recursive deep learning," *Neurocomputing*, vol. 243, pp. 80–87, Jun. 2017. [Online]. Available: http://dblp.uni-trier.de/db/journals/ijon/ijon243.html#NazUASARS17

[7] S. B. Ahmed, S. Naz, M. I. Razzak, and R. Yusof, "Arabic cursive text recognition from natural scene images," *Appl. Sci.*, vol. 9, no. 2, p. 236, 2019. [Online]. Available: http://www.mdpi.com/2076-3417/9/2/236

[8] S. B. Ahmed, S. Naz, M. I. Razzak, and R. Yusof, "Cursive scene text analysis by deep convolutional linear pyramids," in *Proc. ICONIP*, 2018, pp. 1–9. [Online]. Available: https://arxiv.org/abs/1809.10792

[9] S. B. Ahmed, Z. Malik, M. I. Razzak, and R. Yusof, "Sub-sampling approach for unconstrained arabic scene text analysis by implicit segmentation based deep learning classifier," *Global J. Comput. Sci. Technol.*, vol. 19, no. 1-D, Mar. 2019. [Online]. Available: https://computerresearch.org/index.php/computer/article/view/1803

[10] S. Naz, A. I. Umar, R. Ahmad, S. B. Ahmed, S. H. Shirazi, I. Siddiqi, and M. I. Razzak, "Offline cursive Urdu-Nastaliq script recognition using multidimensional recurrent neural networks," *Neurocomputing*, vol. 177, pp. 228–241, Feb. 2016. [Online]. Available: http://dblp.uni-trier.de/db/journals/ijon/ijon177.html#NazUAASSR16

[11] S. Naz, S. B. Ahmed, R. Ahmad, and M. I. Razzak, "Zoning features and 2DLSTM for Urdu text-line recognition," *Procedia Comput. Sci.*, vol. 96, pp. 16–22, Jan. 2016. [Online]. Available: http://dblp.uni-trier.de/db/conf/kes/kes2016.html#NazAAR16

[12] S. Naz, K. Hayat, M. I. Razzak, M. W. Anwar, S. A. Madani, and S. U. Khan, "The optical character recognition of Urdu-like cursive scripts," *Pattern Recognit.*, vol. 47, no. 3, pp. 1229–1248, 2014. [Online]. Available: http://dblp.uni-trier.de/db/journals/pr/pr47.html#NazHRAMK14

[13] F. Lauer, C. Y. Suen, and G. Bloch, "A trainable feature extractor for handwritten digit recognition," *Pattern Recognit.*, vol. 40, no. 6, pp. 1816–1824, 2007. [Online]. Available: http://www.visionbib.com/bibliography/char1024.html#TT102356

[14] X.-X. Niu and C. Y. Suen, "A novel hybrid CNN–SVM classifier for recognizing handwritten digits," *Pattern Recognit.*, vol. 45, no. 4, pp. 1318–1325, 2012. [Online]. Available: http://dblp.uni-trier.de/db/journals/pr/pr45.html#NiuS12

[15] A. Maqqor, A. Halli, and K. Satori, "A multi-stream HMM approach to offline handwritten Arabic word recognition," *Int. J. Natural Lang. Comput.*, vol. 2, no. 4, pp. 21–33, Aug. 2013. [Online]. Available: http://arxiv.org/abs/1309.2506

[16] F. J. Huang and Y. LeCun, "Large-scale learning with SVM and convolutional for generic object categorization," in *Proc. CVPR*, Jun. 2006, pp. 284–291. doi: 10.1109/CVPR.2006.164.

[17] S. Naz, A. I. Umar, R. Ahmad, S. B. Ahmed, S. H. Shirazi, and M. I. Razzak, "Urdu Nasta'liq text recognition system based on multi-dimensional recurrent neural network and statistical features," *Neural Comput. Appl.*, vol. 28, no. 2, pp. 219–231, 2017. [Online]. Available: http://dblp.uni-trier.de/db/journals/nca/nca28.html#NazUAASR17

[18] R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, "Convolutional-recursive deep learning for 3D object classification," in *Proc. 26th Annu. Conf. Neural Inf. Process. Syst.*, 2012, pp. 656–664. [Online]. Available: https://dl.acm.org/citation.cfm?id=2999208

[19] N. Anand and P. Verma, "Convoluted feelings convolutional and recurrent nets for detecting emotion from audio data," Stanford Univ., Stanford, CA, USA, Tech. Rep. CS231n, 2015, pp. 2–7. [Online]. Available: http://cs231n.stanford.edu/reports/2015/pdfs/Cs_231n_paper.pdf

[20] B. L. D. Bezerra, C. Zanchettin, and V. B. de Andrade, "A MDRNN-SVM hybrid model for cursive offline handwriting recognition," in *Artificial Neural Networks and Machine Learning—ICANN 2012* (Lecture Notes in Computer Science), vol. 7553. Lausanne, Switzerland: Springer, Sep. 2012, pp. 246–254.

[21] D. C. Cireşan, U. Meier, and J. Schmidhuber, "Transfer learning for Latin and Chinese characters with deep neural networks," in *Proc. IJCNN*, Jun. 2012, pp. 1–6.

[22] E. Kussul and T. Baidyk, "Improved method of handwritten digit recognition tested on MNIST database," *Image Vis. Comput.*, vol. 22, no. 12, pp. 971–981, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885604000721

[23] M. Fatahi, M. Ahmadi, M. Shahsavari, A. Ahmadi, and P. Devienne, "evt_MNIST: A spike based version of traditional MNIST," 2016, *arXiv:1604.06751*. [Online]. Available: https://arxiv.org/abs/1604.06751

[24] S. B. Ahmed, S. Naz, S. Swati, and M. I. Razzak, "Handwritten Urdu character recognition using one-dimensional BLSTM classifier," *Neural Comput. Appl.*, vol. 31, no. 4, pp. 1143–1151, 2017. [Online]. Available: http://arxiv.org/abs/1705.05455

[25] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks," in *Proc. 23rd Int. Conf. Mach. Learn. (ICML)*, 2006, pp. 369–376. doi: 10.1145/1143844.1143891.

[26] A. Graves, "Supervised sequence labelling with recurrent neural networks," in *Studies in Computational Intelligence*, vol. 385. Berlin, Germany: Springer, 2012, pp. 1–131.

[27] S. B. Ahmed, S. Naz, S. Swati, M. I. Razzak, and A. I. Umar, "UCOM offline dataset—An Urdu handwritten dataset generation," *Int. Arab J. Inf. Technol.*, vol. 14, no. 2, pp. 239–245, 2017. [Online]. Available: http://ccis2k.org/iajit/?option=com_content&task=blogcategory

[28] A. Ul-Hasan, S. B. Ahmed, F. Rashid, F. Shafait, and T. M. Breuel, "Offline printed Urdu Nastaleeq script recognition with bidirectional LSTM networks," in *Proc. ICDAR*, Aug. 2013, pp. 1061–1065. [Online]. Available: http://www.computer.org/csdl/proceedings/icdar/2013/4999/00/index.html

SAAD BIN AHMED received the master's degree in computer sciences and intelligent systems from Technische Universitaet, Kaiserslautern, Germany, in 2012, and the Ph.D. degree in intelligent systems from the Universiti Teknologi Malaysia, in 2019. He has been served as a Research Assistant with the Image Understanding and Pattern Recognition Research Group, University of Technology, Kasierslautern. He is currently a Lecturer with King Saud bin Abdulaziz University for Health Sciences (KSAU-HS), Riyadh, Saudi Arabia. He is also a Faculty Member with the Center of Artificial Intelligence and RObotics (CAIRO), Malaysia-Japan Institute of Information Technology (M-JIIT), Universiti Teknologi Malaysia, Kuala-Lumpur, Malaysia. He authored more than 25 articles in impact factor journals, conferences, and book chapters. His current research interests include document image analysis, machine learning, computer vision, and optical character recognition. He is in the image analysis and pattern recognition field for 15 years and has been involved in various pioneer researches like collection of handwritten Urdu data and used it for Urdu character recognition. He is also providing his expertise in capturing Arabic scene text images and performing research on collected samples by machine learning and pattern recognition techniques.

IBRAHIM A. HAMEED was an Assistant Professor with the Department of Industrial Electronics and Control Engineering, Menofia University, Egypt, from 2011 to 2012. He then joined the Department of Electronic Systems, Aalborg University, Denmark, as a Postdoctoral Researcher with the Section of Automation and Control, from 2013 to 2014, and with the Section of Signal and Information Processing, from 2014 to 2015. Since 2015, he has been an Associate Professor with the Department of ICT and Natural Sciences, Norges Teknisk-Naturvitenskapelige Universitet (NTNU), Alesund, Norway.

SAEEDA NAZ received the B.S. degree from the University of Peshawar (UoP), Peshawar, Pakistan, in 2006, the M.S. degree in computer science from the COMSATS Institute of Information Technology (CIIT), Pakistan, in 2012, and the Ph.D. degree (Hons.) in computer science from the Department of Information Technology, Hazara University, Mansehra, Pakistan, in 2016. She has been an Assistant Professor and the Head of the Computer Science Department, GGPGC No.1 Abbottabad, Higher Education Department of Government of Khyber-Pakhtunkhwa, Pakistan, since 2008. She has published two book chapters and more than 30 articles in peer-reviewed national and international conferences and journals. Her current research interests include optical character recognition, pattern recognition, machine learning, medical imaging, and natural language processing.

MUHAMMAD IMRAN RAZZAK was an Associate Professor in health informatics with the College of Public Health and Health Informatics. He is currently a Senior Researcher with the University of Technology Sydney. He has published more than 70 articles in well reputed journals and conferences. He is an inventor of one Patent and the author of more than 60 articles in well reputed journals and conferences. He secured research grants of more than \$1.3 million. His current research interests include machine learning and health informatics. During the research career, he has successfully developed and delivered several research projects. He received Young Researcher 2015 NGHA, Saudi Arabia, based on his research contributions and Best Researcher during his stay at CoEIA.

RUBIYAH YUSOF received the B.Sc. degree (Hons.) in electrical and electronics engineering from the University of Loughborough, U.K., in 1983, the master's degree in control systems from the Cranfield Institute of Technology, U.K., in 1986, and the Ph.D. degree in control systems from the University of Tokushima, Japan, in 1994. Throughout her career as a Senior Lecturer and a Researcher with UTM, she has been acknowledged for her many contributions in artificial intelligence, process control, and instrumentation design. She is currently the Director of the Center of Artificial Intelligence and Robotics Research Group, Malaysia-Japan Institute of Technology (M-JIIT), Universiti Teknologi Malaysia (UTM), Kuala Lumpur. She is recognized for her work in biometrics systems, such as KenalMuka (face recognition system) and signature verification system which received national and international awards. She is the author of the book *Neuro-Control and its Applications* (Springer Verlag, 1995), which was translated to Russian, in 2001. Dr. R. Yusof is a member of the AI Society Malaysia, the Instrumentation and Control Society Malaysia, and the Institute of Electrical and Electronics Engineers Malaysia.