INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

∗**Corresponding author**.
 Asghar Ali Chandio

School of Engineering and Information Technology, University of New South Wales, Australia. Tel.: +92-3343138628

Department of Information Technology Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan
a.chandio@student.adfa.edu.au

Deep learning-based isolated handwritten Sindhi character recognition

Asghar Ali Chandio[1,2]∗, Mehwish Leghari[2,3], Ali Orangzeb Panhwar[4], Shah Zaman Nizamani[2], Mehjabeen Leghari[3]

1 School of Engineering and Information Technology, University of New South Wales, Australia. Tel.: +92-3343138628
2 Department of Information Technology Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Pakistan
3 Department of Information Technology, University of Sindh, Pakistan
4 Benazir Bhutto Shaheed University, Lyari, Pakistan

Abstract

Motivation : The problem of handwritten text recognition is vastly studied since last few decades. Many innovative ideas have been developed, where state-of-the-art accuracy is achieved for the English, Chinese or Indian scripts. The recent developments for the cursive scripts such as Arabic and Urdu handwritten text recognition have achieved remarkable accuracy. However, for the Sindhi script, existing systems have not shown significant results and the problem is still an open challenge. Several challenges such as variations in writing styles, joined text, ligature overlapping, and others associated to the handwritten Sindhi text make the problem more complex. **Objectives:** In this study, a deep residual network with shortcut connections and summation fusion method using convolutional neural network (CNN) is proposed for automatic feature extraction and classification of handwritten Sindhi characters. **Method:** To increase the powerful feature representation ability of the network, the features of the convolutional layers in the residual block are fused together and combined with the output of the previous residual block. The proposed network is trained on a custom developed handwritten Sindhi character dataset. To tackle the problem of small data, a data augmentation with rotation, flipping and image enhancement techniques have been used. **Findings:** The experimental results show that the proposed model outperforms than the best results previously published for the handwritten Sindhi character recognition. **Novelty:** This is the first research that proposes deep residual network with summation fusion for the Sindhi handwritten text recognition.

Keywords: Handwritten Sindhi character recognition; Sindhi text recognition; cursive text recognition; deep learning; ResNet; convolutional neural network

1 Introduction

Despite advances in the offline and online document text recognition, Sindhi handwritten text recognition still remains an unsolved problem. This is mainly due to the language complexities, complex document layout and different unique characteristics associated to the Sindhi language. Handwritten Sindhi character recognition is more challenging than the printed Sindhi character recognition due to: (1) handwritten Sindhi characters have more variations in terms of aspect ratio when written by different writers or the same writer, (2) handwritten Sindhi text has no defined patterns and depends upon the quality of the writer's writing, (3) different shapes of the same character such as isolated, initial, medial and final make the recognition problem further complex (4) ligature overlapping makes the segmentation of characters more difficult (5) several characters have similar basic shape but they differ either by the number of dots or their positions around the shape, (6) it is cursive in nature and is written in the right to left direction (7) interconnections of two or more characters and several other challenges further reduce the recognition accuracy of handwritten Sindhi characters. Figure 1 shows different shapes of the same character used within a word, while Figure 2 shows the two groups of Sindhi characters with same baseline shape but different number of dots, orientations or their positions around the shape.

Fig 1. *Sindhi handwritten characters with different shapes (isolated, initial, medial and final)*
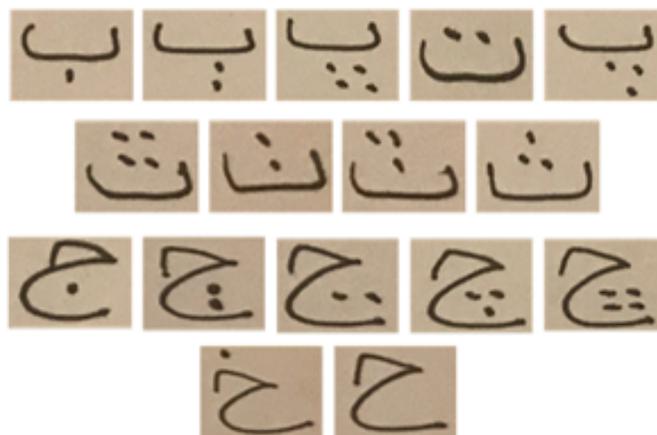
Fig 2. *Two different Sindhi character groups with same baseline shape but different number, orientation or position of dots above, below or inside the shape*

The Sindhi is one of the ancient Indo-Aryan language and is spoken by more than forty million people in the Sindh province, Pakistan and some states of India[1]. It is a type of bidirectional cursive script, where the text is written in the right to left

direction and the numerals are written in the left to right direction. In Pakistan, it is written in Perso-Arabic style, while in India, it is written either in Devanagari or Perso-Arabic scripts[1]. The alphabet of the Sindhi language is mostly derived from the Arabic and Persian scripts with some additional letters which are neither present in the Arabic or Persian scripts. The alphabet of the Sindhi language consists of 52 letters, while Arabic, Persian, Urdu and Pashto scripts have 28, 32, 39 and 44 letters, respectively[2]. Table 1 shows the alphabets of Sindhi language. The letters with red color and bold are only present in the Sindhi language. The letter with green color is present in the Arabic and Persian scripts, however, it has completely different meaning, context and produces different sound when used in the Sindhi script. A detailed review of issues and challenges associated to the handwritten Sindhi text recognition is presented in[3].

Table 1. *Sindhi alphabet letters. The letters with red color and bold are the additional letters that are not present in the Arabic or Persian scripts. The letter with green color is present in Urdu script, how everit has different context.*

No.	Letter	No.	Letter	No.	Letter
1	ا	19	د	37	ڦ
2	ب	20	ڊ	38	ق
3	ٻ	21	ڍ	39	ڪ
4	پ	22	ڌ	40	ک
5	ت	23	ڏ	41	گ
6	ٽ	24	ذ	42	ھگ
7	ٺ	25	ر	43	ڳ
8	ٿ	26	ڙ	44	ڱ
9	ث	27	ز	45	ل
10	ڀ	28	س	46	م
11	ج	29	ش	47	ن
12	ھج	30	ص	48	ڻ
13	ڄ	31	ض	49	و
14	ج	32	ط	50	ھ
15	ڃ	33	ظ	51	ء
16	چ	34	ع	52	ي
17	ح	35	غ		
18	خ	36	ف		

In recent years, deep learning networks particularly CNNs have become most common used methods to solve image processing, pattern recognition and several other computer vision problems. These networks have demonstrated state-of-the-art performance for the Arabic and Urdu handwritten character recognition[4–6] than other methods. Further, CNNs are capable to classify and recognize text at word or character levels without prior information about the structure of the language.

In this paper a state-of-the-art deep learning method using shortcut connection and summation fusion with CNN is proposed to recognize handwritten Sindhi characters. To extract more powerful features, the outputs of convolutional layers in the residual block are further fused together and added with the output of the previous residual block. Generally, conventional methods based on the handcrafted feature extraction algorithms have been used for offline handwritten Sindhi character recognition. The character recognition rate of these methods is not satisfactory yet. To the best of our knowledge, this paper is the pioneer that presents deep learning-based method particularly shortcut connections and summation fusion with CNN to classify and recognize handwritten Sindhi characters.

OCR is one of the important real-world application of automatic pattern recognition systems and is an active research area. A significant research work has been performed for the Latin, Indian, Chinse, Urdu or Arabic scripts[7,8], however the development of Sindhi OCR is still in a preliminary stage and has not shown much improvements. Although, some research has been reported for the Sindhi handwritten character recognition[9–13], but the recognition accuracy is not state-of-the-art.

Awan et al.[9] proposed a neural network-based method to recognize handwritten Sindhi characters. Handwritten character samples collected were scanned and converted into binary images. A horizontal projection method was applied to segment the lines, while a vertical projection was used to segment each character from the lines. A zoning method was used to extract the features from the segmented characters. The average character recognition accuracy reported is 85%. Nizamani and Janjua[10] used artificial neural network (ANN) to recognize isolated handwritten Sindhi characters. The dataset of handwritten Sindhi characters was collected by the native and non-native writers. A dynamic link library was used to fix the input patterns. The

network was trained with backpropagation method. The model was evaluated on the native and non-native handwritten Sindhi characters. The average character recognition accuracy achieved for the native and non-native writers is 91.00% and 79.00% respectively, while the overall accuracy of the model is 85.75%. Similarly, Kumari et al.[11] used a feed-forward neural network to recognize handwritten Sindhi characters. They collected a dataset of only 304 handwritten characters written by 16 different native Sindhi writers. To improve the quality of the handwritten character images, some morphological operations were applied. The network was trained using backpropagation with momentum and adaptive learning rate. They evaluated the model on isolated, two and three handwritten characters. Further, the model was tested on the handwritten characters written by the same and different writers. The average character recognition accuracy for the same and different writers achieved is 85.20% and 81.00% respectively.

Shaikh et al.[12] proposed a sub-word segmentation method for the printed Sindhi text. A height profile vector based on the thinning of a sub-word strokes was calculated and analyzed for the possible individual character segmentation. To get the estimation of possible characters in a sub-word, the location and the number of likely segmentation points were determined. Finally, the possible ending segmentation points in a sub-word were further analyzed to determine the actual number of characters. Memon et al.[13] used character geometry-based feature extraction method with feed-forward neural network to identify glyphs and recognize handwritten Sindhi characters. A horizontal and vertical projection based on the space between two characters were applied to segment the scanned handwritten Sindhi character images into lines and individual characters. Sanjrani et al.[14] and Ali et al.[15] applied machine learning techniques to recognize handwritten Sindhi numerals. A detailed review of the methods proposed for the handwritten Sindhi character recognition is presented in[16].

Some research studies on the online handwritten Sindhi text and numbers recognition are presented in[17,18]. One of the recent works used CNN to recognize multi-size and multi-font printed Sindhi characters[19]. Three different CNN models were implemented, and the best character recognition accuracy reported is 99.96%.

2 Proposed Methodology

The block diagram of the proposed deep learning-based Sindhi handwritten character recognition model is illustrated in Figure 3. The proposed model is based on the residual networks presented in[20,21]. The input images are converted to grayscale before passing to the network. The input data of the model are MxNxD images where M is the width, N is the height of the image and D is the image channel size. In the proposed model, width, height and channel size of the images are 48, 48 and 1 respectively. Different to the model in[20], the proposed model uses 3x3 convolutional layer with 32 output units without following a max pooling layer. Moreover, the max pooling layer are replaced with the average pooling layers and are used in the residual block. The model uses 4 residual blocks with 64, 128, 256 and 512 output units. Each residual block uses three convolutional layers with 1x1, 3x3 and 1x1 kernel sizes. An average pooling layer with a window size of 2x2 is followed by the last residual. Two fully connected layers with 512 and 52 output neurons are used to extract high-level features and classify the characters. The last fully connected layer is followed by a Softmax activation function to perform multi-class classification.

The residual networks are based on several stacked residual blocks, where each block consists of either two or three convolutional layers. Several residual networks with different organization of residual blocks have been developed. The operations between residual units vary depending upon the architecture of the network. Figure 4 (a) shows the residual block proposed in[21] and Figure 4(b) shows the modified residual block with summation fusion proposed for the handwritten Sindhi character recognition. The analysis of different identity mappings in residual network is explained in[21]. The general form of the residual block is expressed as:

$$R_l = h(x_l) + F(x_l, W_l) \tag{1}$$

$$x_{l+1} = f(R_l) \tag{2}$$

where x_l and x_{l+1} are the input and output feature vectors of l-th residual block, F is a residual function, $h(x_l)$ is an identity mapping, W_l is the set of convolutional weights and biases in the l-th residual block, f is an activation function, which is a rectified linear unit (ReLU) in this paper. The identify mapping is an addition operation that adds the output of the previous residual block with the output of the block ahead. When the feature dimensions of both residual blocks are equal, the identity mapping does not add additional network parameters. However, when the dimensions of both blocks are not same, the identity mapping can be performed in two ways: (1) to increase the feature vector dimensions with extra zero padding or (2) to perform a linear projection such as W_s for increasing the feature vector dimensions of the shortcut connections when $F(x_l)$ and x_l have

Load Image

	Input Layer 48x48x1
3x3 conv, 32	Convolutional Layer 48x48x32
ResNet Block-1 /2	24x24x64
ResNet Block-2 /2	12x12x128
ResNet Block-3 /2	6x6x256
ResNet Block-4 /2	3x3x512
Average Pooling	1x1x512
Fully Connected	Fully Connected, 512
Fully Connected	Output Layer, 52
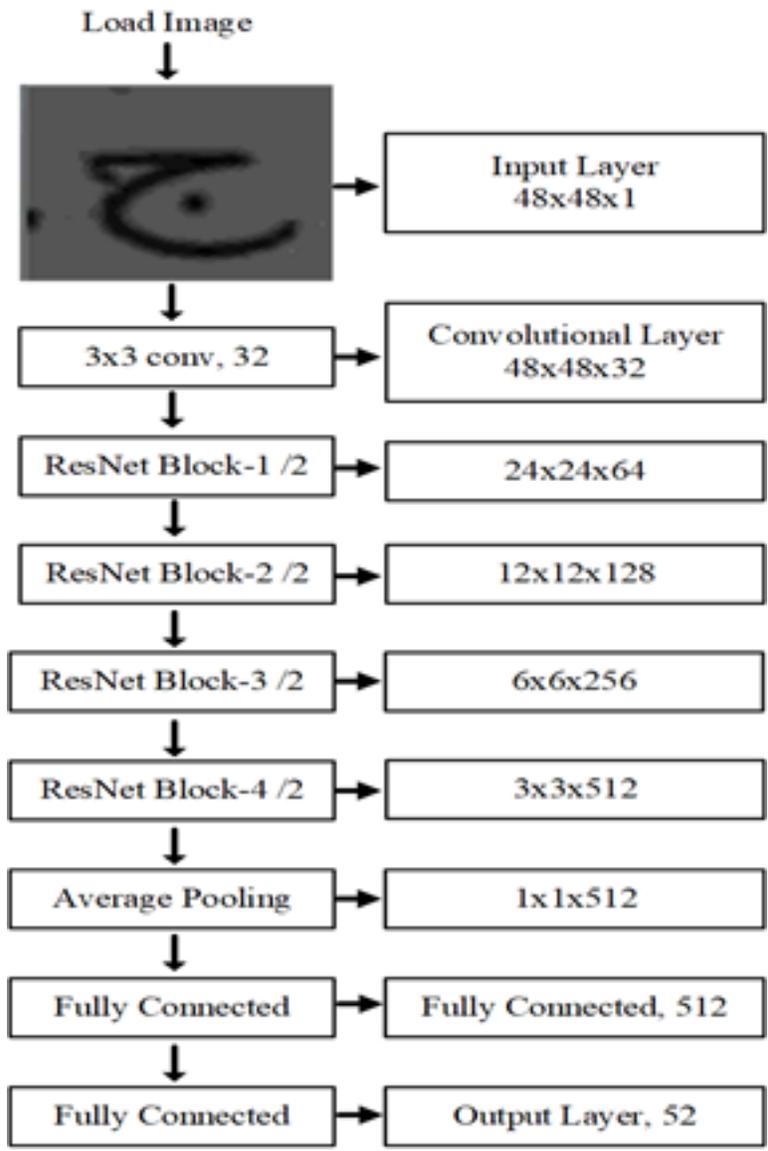
Fig 3. *Proposed Methodology*

different dimensions as:

$$R_l = F(x_l, \{W_l\}) + W_s x_l \tag{3}$$

This linear projection W_s can be implemented with 1x1 convolutional layer. However, this will include additional trainable parameters in the model.

In the proposed residual block as shown in Figure 4(b), an element-wise addition operation is performed to add the output of convolutional layers. The output of two convolutional layers at x and y locations is added together when the feature vectors in both layers have the same dimensions as:

$$FV^{sum} = f^{sum}(m^a, m^b) \tag{4}$$

where m^a and m^b are the two feature vectors, wherein $m^a \, \varepsilon \, R^{HxWxD}$ and $m^b \, \varepsilon \, R^{HxWxD}$. The summation fusion does not increase the feature vector dimensions and adds no additional parameters in the network, which helps the network to converge fast. The other form of the fusion called concatenation will increase the feature vector dimensions and delays the network to reach its
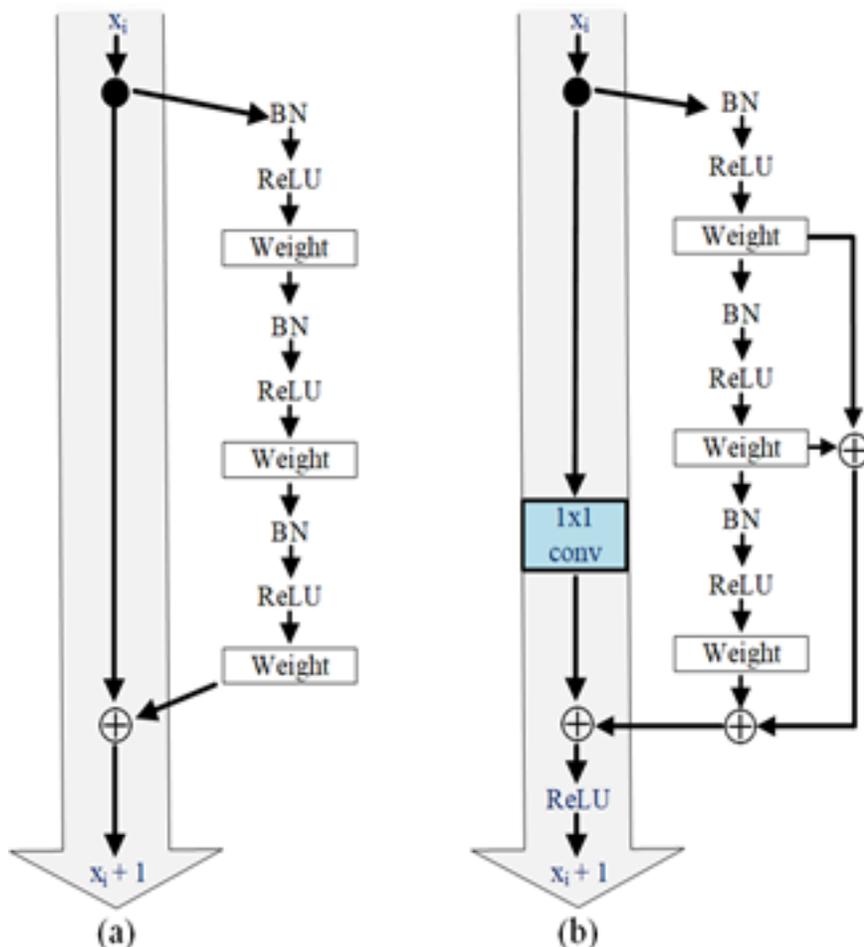
Fig 4. *Residual blocks. (a) a residual block proposed in* [21]. *(b) proposed residual block with summation fusion. Proposed residual block implements 1x1 convolutional layer to add the features of two layers with different feature vector dimensions. The proposed layer uses ReLU after the shortcut connection.*

convergence. Hence, the concatenation fusion is not implemented in the proposed model. Further, the proposed summation fusion with shortcut connection improved the recognition accuracy of handwritten Sindhi characters.

2.1 Network Training

Handwritten Sindhi character recognition is a multi-class classification problem, therefore, a sparse_categorical_crossentropy was selected as a loss function. To minimize the loss value, the model was trained using stochastic gradient descent (SGD) optimizer with a momentum of 0.9, and a weight decay of 0.83 exp -4. Different learning rates were trialed, and the lowest loss value was achieved with a learning rate of 0.005. The network was trained up to 60 epochs with a batch size of 64.

3 Experimental Setup and Results

The experiments were performed on an Intel Core i7 CPU @ 3.60GHz with 16Gb of random-access memory (RAM) and 4Gb of NVIDIA graphical processing unit (GPU). The proposed model was implemented using Keras[1] open source deep learning library with a Tensorflow[2] as backend.

---

[1] https://github.com/keras-team/keras
[2] https://github.com/tensorflow/tensorflow

3.1 Dataset

The dataset samples were collected from 130 native Sindhi text writers on the white plain pages. The Sindhi characters were written in different colors such as blue, red, green and black. The character data collected has variations in terms of writing styles, and aspect ratios. The collected data was photographed into images with 16MP mobile camera. The handwritten characters were manually segmented from the photographed images and saved with 48x48x3 dimensions. The total number of samples for 52 unique character classes are 6760 with 130 samples per class. The dataset was split into training and testing samples with a ratio of 80:20. To tackle the problem of small data while training deep CNN model, a data augmentation method with angle rotation, flipping, image enhancement techniques was used to increase the number of training samples. Figure 5 illustrates some examples of segmented character images in the proposed dataset. The dataset has sufficient number of handwritten character samples and can be used as a benchmark for the Sindhi handwritten text.

Fig 5. *Some sample images of handwritten Sindhi characters in the dataset.*

3.2 Evaluation Protocols

The proposed model was evaluated using three most standard evaluation protocols such as precision, recall and f-score as used in different character recognition problems. Precision is the number of true predictions by the classifier that belong to the positive classes. Recall is the number of true predictions by the classifier that belong to the all positive samples in the dataset. The precision and recall do not give optimal accuracy. Therefore, the overall performance of the model was measured in terms of f-score, which is a weighted average of precision and recall defined as:

$$f - score = \left(\frac{2 * precision * recall}{precision + recall} \right) \times 100 \tag{5}$$

3.3 Evaluation on the handwritten Sindhi character dataset

Handwritten Sindhi character recognition results using the proposed model with residual block and summation fusion and the residual block as proposed in[21] are shown in Table 2 . The precision recall and f-score achieved with the proposed method is 95.00%, 94.00% and 94.00% respectively, whereas with residual block as proposed in[21] these results are 92.00%, 92.00% and 92.00%. This shows that the residual blocks with summation fusion outperform than standard residual blocks. The confusion matrix with test data of handwritten Sindhi characters as illustrated in Figure 6 shows that most of characters have recognition rate of more than 90%.

Table 2. *Handwritten Sindhi character recognition results with the proposed model and residual block presented in* [21]

| Method	Precision (%)	Recall (%)	F-Score (%)
Residual Block [21]	92.00	92.00	92.00
Proposed Method	95.00	94.00	94.00
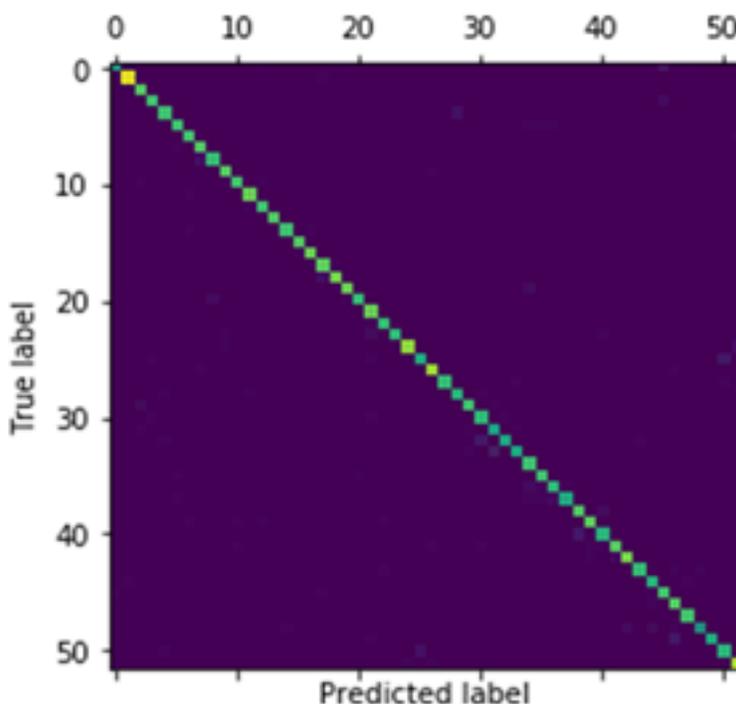
Fig 6. *Confusion matrix*

3.4 Performance comparison with previously published results

A limited research using conventional machine learning methods has been reported for the handwritten Sindhi character recognition. To the best of authors' knowledge, this is the first research proposing deep learning method for the handwritten Sindhi character recognition. The performance of the proposed method is hence compared with existing methods as reported in[9–11,13]. In[9] a zoning method was used to extract the features from the segmented Sindhi characters and an artificial neural network was applied for the classification. The number of collected, training and testing samples are not provided. The average accuracy achieved is 85%. In[10] the handwritten Sindhi character data was collected only from five native and five non-native writers to train the model. Total number of training samples collected were 520. The model was evaluated on 208 samples collected from two native and two non-native writers. The average character recognition accuracy obtained is 85.75%. In[11] a dataset of 304 handwritten Sindhi characters from the native 16 writers on a plain paper was collected. Only 19 characters were written by each user. pixel-level features from each character images were extracted, and a feed-forward neural network

was applied to recognize the characters. Compared to the above methods and datasets, the proposed model implements a deep learning-based technique to recognize handwritten Sindhi characters. The number of samples in the proposed dataset are much more than existing datasets. Table 3 shows the performance comparison of the proposed method with the previously published results.

Table 3. *Performance comparison of the proposed model with the previously published handwritten Sindhi characters.*

Method	Accuracy (%)
Awan et al. [9]	85.00
Nizamani and Janjua [10]	85.75
Kumari et al. [11]	85.20
Proposed Model	94.00

The results in Table 3 show that proposed model with deep learning-based shortcut connection and summation fusion outperforms than conventional machine learning methods.

4 Conclusion and Future Work

A large portion of research has been carried for handwritten text recognition, where state-of-the-art accuracy is achieved for Latin, Indian, Chinse and Arabic scripts. However, very few research works are reported for the Sindhi handwritten text recognition. This study proposed the handwritten Sindhi character recognition using deep learning-based method. A shallow residual network with shortcut connections and summation fusion was proposed. The summation fusion method extracted more powerful features from the handwritten character images and outperformed than original residual network with shortcut connections. To evaluate the model a new handwritten Sindhi character dataset was developed. The data was collected from 130 native Sindhi writers. Each writer was allowed to write one sample of each character class. The results obtained show the proposed model outperformed than conventional machine learning methods on the handwritten Sindhi character recognition. In future, word and line-level data samples will be collected. Further, a whole word and text line-based Sindhi recognition system will be implemented.

Acknowledgment

References

1) Leghari M and Rahman MU. Towards Transliteration between Sindhi Scripts Using Roman Script. *Linguistics and Literature Review*. 2015;1(2):101–110. Available from: https://dx.doi.org/10.32350/llr.12.03.

2) Hakro DN and Talib AZ. Printed Text Image Database for Sindhi OCR. *ACM Transactions on Asian and Low-Resource Language Information Processing*. 2016;15(4):1–18. Available from: https://dx.doi.org/10.1145/2846093.

3) Hakro DN, Ismaili IA, Talib AZ, Bhatti Z and Mojai GN. Issues and challenges in Sindhi OCR. *Sindh University Research Journal-SURJ*. 2014;46(2):143–152.

4) Boufenar C, Kerboua A and Batouche M. Investigation on deep learning for off-line handwritten Arabic character recognition. *Cognitive Systems Research*. 2018;50:180–195. Available from: https://dx.doi.org/10.1016/j.cogsys.2017.11.002.

5) Ahmed SB, Naz S, Swati S and Razzak MI. Handwritten Urdu character recognition using one-dimensional BLSTM classifier. *Neural Computing and Applications*. 2019;31(4):1143–51.

6) Ghanim TM, Khalil MI and Abbas HM. Comparative Study on Deep Convolution Neural Networks DCNN-Based Offline Arabic Handwriting Recognition. *IEEE Access*. 2020;8:95465–95482. Available from: https://dx.doi.org/10.1109/access.2020.2994290.

7) Mathew M, Singh AK and Jawahar CV. Multilingual OCR for indic scripts. *proceedings of 12th IEEE IAPR Workshop on Document Analysis Systems (DAS)*. 2016;p. 186–191.

8) Islam N, Islam Z and Noor N . 2017,. Available from: https://arxiv.org/ftp/arxiv/papers/1710/1710.05703.pdf.

9) Awan SA, Abro ZH, Jalbani AH and Hameed M. Handwritten Sindhi Character Recognition Using Neural Networks. *Mehran University Research Journal of Engineering and Technology*. 2018;37(1):6–6.

10) Nizamani AM and Janjua NU. Isolated Handwritten Character Recognition in Sindhi Language using Artificial Neural Network. *Journal of Independent Studies and Research*. 2012;10(1).

11) Kumari A, Sangrasi DM, Bhatti S, Chowdhry BS and Kumari S. Off-line Sindhi Handwritten Character Identification. *International Journal of Information Technology and Computer Science*. 2019;11(6):9–17. Available from: https://dx.doi.org/10.5815/ijitcs.2019.06.02.

12) Shaikh NA, Mallah GA and Shaikh ZA. Character segmentation of Sindhi, an Arabic style scripting language, using height profile vector. *Australian Journal of Basic and Applied Sciences*. 2009;3(4):4160–4169.

13) Memon NA, Abbasi F and Zardari S. Glyph Identification and Character Recognition for Sindhi OCR. *Mehran University Research Journal of Engineering and Technology*. 2017;36(4).
14) Sanjrani AA, Baber J, Bakhtyar M, Noor W and Khalid M. Handwritten optical character recognition system for Sindhi numerals. *Proceedings of IEEE International Conference on Computing, Electronic and Electrical Engineering*. 2016;p. 262–267.
15) Ali I, Ali I, Subhash AK, Raza SA, Hassan B and Bhatti P . 2019,. Available from: http://paper.ijcsns.org/07_book/201905/20190526.pdf.
16) Solangi YA, Solangi ZA, Raza A, Shaikh NA, Mallah GA and Shah A. Offline-printed Sindhi Optical Text Recognition: Survey. *Proceedings of IEEE 5th International Conference on Engineering Technologies and Applied Sciences*. 2018;p. 1–5.
17) Chandio AA, Leghari M, Hakro D, S A and Jalbani AH. A Novel Approach for Online Sindhi Handwritten Word Recognition using Neural Network. *Sindh University Research Journal-SURJ*. 2016;48(1):213–216.
18) Chandio AA, Jalbani AH, Laghari M and Awan SA. Multi-Digit Handwritten Sindhi Numerals Recognition using SOM Neural Network. *Mehran University Research Journal of Engineering and Technology*. 2017;36(4).
19) Chandio AA, Leghari M, Leghari M and Jalbani AH. Multi-Font and Multi-Size Printed Sindhi Character Recognition using Convolutional Neural Networks. *Pakistan Journal of Engineering and Applied Sciences*. 2019;24(1):36–42.
20) He K, Zhang X, Ren S and Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016;p. 770–778.
21) He K, Zhang X, Ren S and Sun J. Identity mappings in deep residual networks. In: Proceedings of European conference on computer vision . 2016,. p. 630–645.