

Deep Wavelet Neural Network based Robust Text Recognition for Overlapping Characters

Neha Tripathi^{1*}, Pushpinder Singh Patheja²
School of Computing Science and Engineering
VIT Bhopal, India

Abstract—This paper presents a deep learning based intelligent text recognition system with touching and overlapped characters. The robustness and effectiveness in the proposed model are enhanced through the modified configuration of neural network known as Deep Wavelet Neural Network (DWNN). The capability of deep learning networks to learn efficiently from an unlabeled dataset has attracted the attention of many researchers over the last decade. However, the performance of these networks is subject to the quality of the dataset and invariant image representation. Numerous optical character recognition techniques have also been presented in the recent years, but the overlapped and touching characters have not been addressed much. The nonlinear and uncertain representation of image data in case of overlapped text adds severe complexity in the process of feature extraction and respective learning. The proposed architecture of DWNN uses fast decaying wavelet functions as activation function in place of conventional sigmoid function to cope up with the uncertainties and nonlinearity of the data representation in overlapped text images. It comprises of cascaded layered architecture of translated and dilated versions of wavelets as activation functions for the training and feature extraction at multiple levels. The local transformation and deformation variation in the visual data has also been taken care efficiently through the modified architecture of DWNN. Comprehensive experimental analysis has been performed over various test images to verify the effectiveness of the proposed text recognition system. The performance of the proposed method is assessed with the help of the metrics, namely, estimation error, cost function and accuracy. The proposed approach will be implemented in MATLAB.

Keywords—Text recognition; overlapped characters; deep wavelet neural network; feature extraction; segmentation; basis function; optical character recognition

I. INTRODUCTION

The field of optical character recognition has attracted a lot of attention over the last two decades due to its capability to extract the meaningful information from the printed or handwritten text. It has been used successfully in the applications like automatic language translation, text to speech converters, smart scanning devices, text summarization, automated postal address and ZIP code reading, bank cheque reading, etc. The characters are recognized through the process of conversion of these characters into the machine-readable formats like ASCII code. The intended information is extracted from the images based on a thorough analysis of the text and graphical features of the document. Both these feature sets are processed to deal with the textual and non-textual component

of the image of the document. The non-textual components involve company logos, emoticons, make up line diagrams, delimiting lines between text, etc. The variety and diversity introduced in text extraction through different font sizes, font types, and orientation make the problem of OCR even more challenging [1-3].

A typical framework of OCR involves the process of preprocessing, segmentation, feature extraction and recognition. These processes need to be robust as well as efficient to present an accurate text extraction model. Various techniques have been proposed by numerous researchers to address the problems and challenges faced at these processes. However, segmentation process is the most complicated stage in OCR because of its dependency over large number of factors including the quality of scanner/camera, illumination, font type, font size, orientation, angular features, ink diffusion, etc. [4]. The process of feature extraction and recognition is also an important phase of OCR as it governs the accuracy of the overall process. Various techniques ranging from statistical models to deep learning framework have been proposed for the text recognition based on the characteristics of the features of documents [5-8].

The overall performance of any OCR technique relies majorly upon the quality of image. Presence of noise in the image can greatly degrade the efficiency and accuracy of the process. Although the preprocessing phase is responsible to filter out the noise present in the image, but it is subject to the type of the respective noise. It is also assumed in most of the OCR techniques that the text lines are equitably straight and the distance between neighboring text lines is precise. However, these assumptions could be characterized for text with overlapped characters and slanted orientation. Therefore, the complexity associated with the problem of extracting the text from the images with overlapped or touching characters has attracted the attention over the last few years. But extremely limited work has been done in this field so far. The major problem in designing an efficient OCR framework for overlapped or touched characters is to eliminate the noise from the binary images and smooth them for the feature extraction and recognition. The training algorithm must be intelligent and adaptive enough to deal with the abrupt feature variations generated due to the overlapping [9-12].

Various machine learning and deep learning algorithms have been presented by the researchers for the feature extraction and recognition of text from images. But their accuracy is greatly subject to the textual distribution. Even the best OCR techniques for normal text distribution are found to

*Corresponding Author

perform very poorly for the overlapped characters images. This is because of the reason that the training of the intelligent framework is done through the datasets where the characters are distant and clearly separated from each other [13]. The same network fails to recognize the characters in case of overlapped distribution. Therefore, it requires an adaptive intelligent model which could be able to mitigate the effects of sharp and abrupt changes in the text features distribution.

Some novel configurations of deep neural networks like Convolutional Neural network (CNN), Long Short-term Memory (LSTM)/ Recurrent Neural Network (RNN), Generative Adversarial Network (GAN), etc. have got a lot of appreciation and attention due to their superior learning characteristics and efficient classification performance [14-18]. These networks have also been implemented by various researchers for OCR problems with varying complexity. However, these networks could not be able to present the promising results in highly dynamic and uncertain feature distribution space. The capability of these networks to provide a fast decision making with long term learning and lesser time complexity is limited and therefore the derived model is found to be conservative learning model. The major reason behind this is the basis function used in these networks which is not orthogonal in nature and results into an inefficient and non-unique representation of decision space. Neural networks with non-orthogonal activation functions cannot guarantee the convergence of the learning curve and may get trapped in local minima for certain initial conditions [19].

These limitations of conventional neural network framework have been addressed in some literatures and various novel activation functions have been proposed. However, the replacement of sigmoid functions in DWNN by rapidly decaying functions known as wavelets has generated very promising results for the dynamic and uncertain feature distribution and larger decision space. Due to the time-frequency localization property of wavelet function, the learning characteristics are found to be immensely improved as compared to the conventional DWNN [20]. These modified networks are also named as Wavelet Neural network (WNN)/WaveNets as they augment the learning potential of conventional neural network architecture with the identification and decomposition ability of wavelets.

Deep learning techniques have proved their superiority over traditional approaches for pattern recognition problems. However, the complexities associated with the overlapped and touching characters in a text requires even deeper approaches. This paper presents an intelligent and robust deep learning framework, DWNN for the text extraction from the images with overlapped and touching characters. The textual and image features distribution is very abruptly and randomly distributed in case of overlapped scripts, loosely configured characters, broken characters, connected characters. They are major cause of segmentation errors and result in inaccurate recognition. The application of high performance DWNN to learn and recognize the characters can greatly enhance the overall OCR performance. The major contributions of this research work can be mentioned as follows:

- Deriving the mathematical framework of the proposed DWNN using multiple layers.
- Exploiting the features distribution from even the local patches of the images through the localized spectral nature of activation function.
- Employing high dimensional deep feature representation.
- Analyzing the performance of the proposed framework for different challenging character variations.
- Attaining the best possible accuracy for the noisy, overlapped, and touching characters.

The paper is organized as follows: Section II deals with the literature survey through the analysis of related work. The mathematical framework of the proposed DWNN is given in Section III. The proposed DWNN based text extraction strategy for overlapped characters is discussed in Section IV. Effectiveness of the proposed strategy is illustrated through the simulation analysis in Section V while Section VI concludes the paper.

II. RELATED WORK

The potential applications of OCR in various fields have attracted a lot of researchers to pursue research in this field. Numerous algorithms have been presented over the last decade for various stages of OCR viz. preprocessing, segmentation, feature extraction and recognition. Various techniques for the preprocessing of the images are presented by the researchers depending upon the type of images. Some commonly used techniques are noise removal, skew removal, thinning, morphological operations, etc. [21-23]. However, the most challenging aspect of OCR is the segmentation of images because of the diversity in the characteristics of text. It is also the most dominating phase of the OCR technique as the overall performance is depending upon the quality of segmentation. Because a single segmentation technique cannot be suitable for all type of textual distribution, many segmentation techniques have been proposed over the last decade [24].

Most of the segmentation techniques presented in the literature are based on the assumptions that the textual distribution is uniform, equidistant, and straight. But segmentation of overlapped and touching characters has not been addressed vastly and remained an open-ended problem. Farulla et al. [25] addressed this problem and proposed a fuzzy logic-based approach by combining various segmentation techniques altogether. Fuzzy rule base was prepared to derive a combined segmentation methodology which is robust to the noisy data. Garain and Chaudhuri [26] have implemented a multiple factors-based approach for the segmentation of touching characters in a printed text. The factors taken in the respective analysis were middleness, transitions and blob thickness. An algorithm was derived to facilitate the segmentation using these parameters. Nomura et al. [27] presented a histogram-based method for the primary character segmentation followed by morphological thickening and thinning operation to segment the overlapped characters. However, the accuracy was subject to the variation of textual data. A novel Harrow space filter was proposed by Tian et al.

[28] for the license plate character recognition. They have augmented weighted map algorithm to add robustness to the segmentation approach. Zheng et al. [29] have presented a segmentation technique for Arabic characters using the structural properties. They have utilized the vertical projections and some heuristics of these properties to differentiate between background and foreground regions to detect isolated characters. Similarly, the projections and statistical dimensional information was used for the Devanagari characters segmentation by Bansal and Sinha [30].

The problem of feature extraction and pattern recognition has also been addressed by the researchers to enhance the OCR performance. Various techniques like Syntactical Analysis, Neural Networks Template Matching, Hidden Markov Models, Bayesian Theory, etc. have been implemented for the efficient and robust recognition for different languages. E. B. Lacerda and C. A. Mello [31] presented a Self-Organizing Maps (SOM) based approach for the separating the touching characters. They have used the skeletonization process to cluster the feature points. Elnagar and Alhaji [32] proposed a technique to isolate the handwritten digit strings by normalizing and thinning which helped in identifying the feature points. These points are derived through the decision line from the deep points in the image. Gattal et al. [33] extended the research by combining different segmentation approaches based on configuration links between overlapped digits. They have used the sliding window Radon transform of these segmentation techniques to take the decision about selecting or discarding a digit image. Histograms of the vertical projection have also been used here for the contour analysis.

The computation cost associated with these segmentation techniques has posed serious concern during the real time implementation of OCR techniques. To deal with the issue of over-segmentation, several researchers have transformed the problem of segmentation and classification into a sole problem of recognition which does not involve these heavy segmentation algorithms. D. Ciresan [34] and A. G. Hochuli et al. [35] derived a convolutional neural network (CNN) framework for the recognition of these digits. These CNNs are trained over the datasets including isolated and touching text. Both these research works have avoided heavy segmentation issues, but their performance was subject to the availability of large datasets. H. Zhan et al. [36] have improved the OCR performance by combining Connectionist Temporal Classification (CTC) with Recurrent Neural Network (RNN). The features of the input image were extracted through the residual network and RNN was employed to derive the contextual information through these features. CTC model was derived to tune the parameters of the network for accurate classification. Zhang et al. [37] replaced RNN with DenseNet to further enhance the efficiency of the previous designs. These OCR frameworks have improved the recognition accuracy and efficiency, but the high computation complexity has remained the point of concern for the researchers.

Various configurations of deep learning have been applied for the problem of character recognition, but the performance of these techniques is subject to the availability of large training dataset. Also, their performance in dynamic and random environments cannot be guaranteed due to its

dependency on initial conditions and uncertain convergence characteristics. The learning characteristics and weight adaptation is also found to be affected by the choice of a suitable activation function. Zhang et al. [38] has recently proposed a novel framework of neural network by replacing the sigmoid function by fast decaying wavelet function which presented better convergence rate and learning capabilities due to the space and frequency localization property of wavelets. Owing to the superior learning characteristics, this modified configuration of neural network known as Wavelet Neural network (WNN) has been used by the researchers in the fields of engineering where the feature space is highly dynamic and uncertain like object tracking, automatic control, advance communication, etc. [39]. WNN is used in this paper for the character recognition in overlapped and touching characters as the textual features distribution is uncertain and random in these types of images.

III. DEEP WAVELET NEURAL NETWORK FRAMEWORK

The real potential of any neural network configuration is dependent on the activation function used for the learning of the network. Various activation functions like sigmoid, ReLU, fuzzy, etc. have been used by the researchers traditionally for variety of applications [40]. However, the performance of these functions for the data distribution with high diversity and randomness could not be guaranteed due to the globally defined nature. DWNN are the modified framework of conventional neural network architecture with translated and dilated versions of wavelet functions as activation function. The spectral characteristics of wavelets are explored to select the best suited wavelet for a typical data sample space. The dyadic translation and binary translation of the wavelets in a subspace $L^2(\mathfrak{R})$ are used as basic functions for the data processing at the nodes of the wavelet network. Universal approximation property of neural network architecture deduces that the linearized combination of these basis functions is further used to evaluate the estimation function $\zeta \in L^2(\mathfrak{R})$. A typical configuration of the DWNN is shown in Fig. 1. It is a four layered architecture namely input layer, wavelon layer, product layer and output layer. The preprocessing of the data is done at the input layer while the wavelon layer passes the data through the wavelet activation function. The deep features of the data are extracted at this layer through the translated and dilated versions on the wavelet function. The product of the outcome of these processed outcomes is then evaluated at the product layer and the decision is provided through the output layer.

The n dimensional biased network with m nodes generates the output as

$$\zeta = \omega^T \varphi(x, \tau, \sigma) + \mu^T \phi(x, \tau, \sigma) \quad (1)$$

where $x = [x_1, x_2, \dots, x_n]^T \in \mathcal{R}^n$ is the input vector, $\tau = [\tau_1, \tau_2, \dots, \tau_m]^T \in \mathcal{R}^{m \times n}$ and $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_m]^T \in \mathcal{R}^{m \times n}$ are translation and dilation parameters respectively, $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_m]^T \in \mathfrak{R}^m$ is wavelet function and

$\phi = [\phi_1, \phi_2, \dots, \phi_m]^T \in \mathfrak{R}^m$ represents the respective bias function. The weights and bias function of the network are represented as $\omega = [\omega_1, \dots, \omega_m]^T \in \mathfrak{R}^m$ and $\mu = [\mu_1, \dots, \mu_m]^T \in \mathfrak{R}^m$ respectively.

Optimization of the network parameters is the most important aspect of any deep learning network as it governs the nature of the learning curve of the network. The best approximation of the desired function can be attained through the optimal parameter vectors. The optimization of the network is achieved through an estimate function defined as:

$$\hat{\zeta}(x(n)) = \hat{\omega}^T \phi(x(n)) + \hat{\mu}^T \varphi(x(n)) \quad (2)$$

where $\hat{\omega}, \hat{\mu}$ are the estimates of the optimal values of the network parameters ω^*, μ^* respectively. The problem of optimization is reframed in terms of estimation error defined as

$$\tilde{\zeta}(x(n)) = \zeta(x(n)) - \hat{\zeta}(x(n)) = \{ \tilde{\omega}^T(n)\phi(x(n)) + \tilde{\mu}^T(n)\varphi(x(n)) + \varepsilon(x(n)) \} \quad (3)$$

The objective is now to reduce the value of estimation error $\tilde{\zeta}$ to an arbitrarily small value by carefully selecting the number of resolutions. The adaptive algorithm used to tune the weights of the NN framework is derived through gradient descent algorithm. The respective tuning laws for the weights are derived as:

$$\omega(n+1) = \omega(n) + \Delta\omega(n) = \omega(n) + \mu \left(-\frac{\partial e(n)}{\partial \omega(n)} \right) \quad (4)$$

where the learning rate and tuning weights are represented as μ and ω respectively. The weights are modified till the estimation error is not minimized to exceedingly small value.

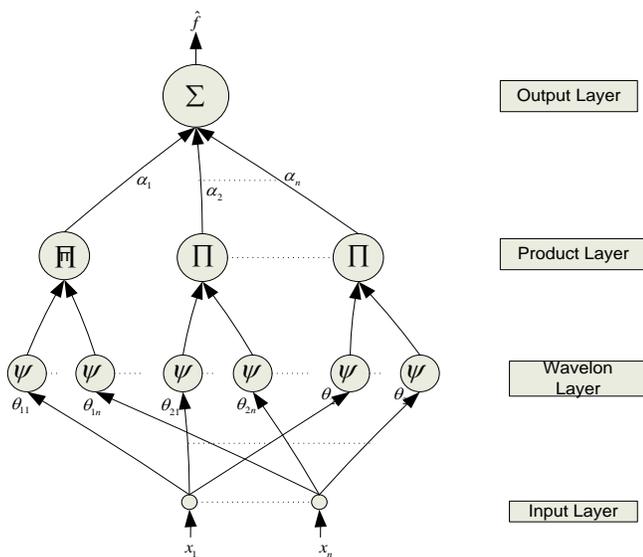


Fig. 1. Architecture of Wavelet Neural Network (WNN).

IV. PROPOSED METHODOLOGY

The text recognition framework for the overlapped characters using DWNN is shown in Fig. 2. The overall process is divided into following phases: Preprocessing, segmentation, Classification, and recognition. The proposed framework for the DWNN based OCR for overlapped characters is shown in Fig. 2. The process is discussed in detail below:

Preprocessing: The preliminary process of the images is performed at preprocessing stage which includes binarization, noise filtering, skew correction and thinning. Binarization is performed over the image to transform it from RGB to grayscale. A threshold is chosen in this process to separate the foreground and background information. However, it is overly complex problem in cases where the contrast between text pixels and background is low. The thin text strokes or non-uniform illumination during image capturing may result in background bleeding into text pixels during digitization. Multi-thresholding is performed here to identify the relevant gray level information and eliminate the noise and isolated points. The pixel values of the input image are first provided to the multi-threshold module which compares their values with the upper and lower threshold values. If the data is within the range defined through the threshold levels, the pixel data is assigned as 1, otherwise it will be assigned the value of 0. It results into the binary distribution of the text image which could further be used for the respective analysis. The resultant digital image may be subject to some disturbances or noise due to the undesired aspects in optical scanning devices and camera. This noise is removed during preprocessing through suitable filters. The typical noise present in the scanning process is salt and pepper noise because of the quality of paper scanned and parasitic components. The filtering of these noises may result into some edge losses; therefore, median filter is used in this work so as to preserve the edges of the text images.

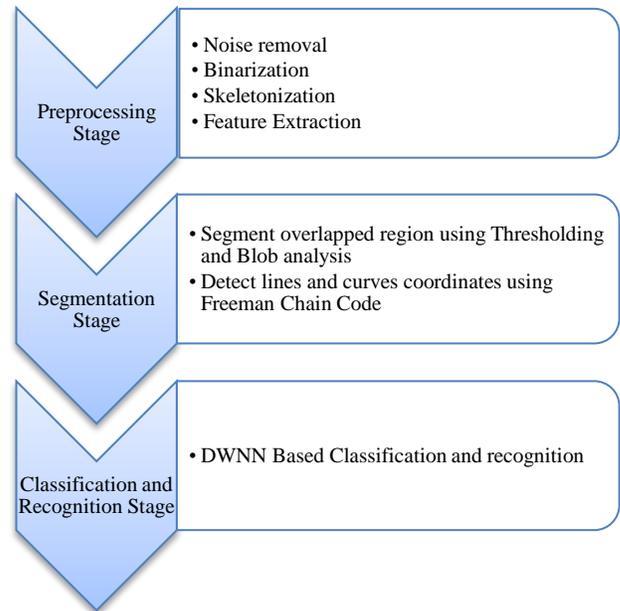


Fig. 2. Proposed Methodology.

Skew correction is then performed over the filtered images to remove any kind of disorientation of text in the images. The skew of the scanned document image specifies the deviation of its text lines from the horizontal or vertical axis. The projection profiles are used as a suitable feature for skew detection after extracting the black text pixels for analysis. The shape of the characters is then extracted using the thinning process by exploring the pixel distribution of the characters smoothly. The process of preprocessing used in this work is shown in Fig. 3.

Segmentation: The most important process of the proposed work is segmentation which is responsible for efficient split of the overlapping and touching characters in a text image. The overall performance of the OCR is dependent on the correctness of segmentation as the corresponding features of the segmented images are used for the recognition and classification. Keeping the seriousness of this process in view, Blob analysis-based method is used for the segmentation in this paper. The filtered pixels are classified or clustered in Blob analysis based on the respective pixel values. If the pixel value is nearly equal to the neighboring pixel value, then they are kept in a same blob. This process results into number of clusters or blobs having same kind of pixel distribution and spectral characteristics. The segmentation in this paper is performed using the blob adjacency analysis where each pixel is checked with its eight neighbor pixels on vertical, horizontal, and diagonal axes.

The projection profiles of these blobs and the connected component analysis are used to derive the blob classification. The performance for the overlapped and touching characters is further improved by applying the dilation and erosion over the merged characters.

The shape and size of the blobs are identified in this work using the Freeman chain coding with eight connectivity approach which derives the boundary of the blob contour. Chain codes use the connected sequence of straight-line segments of specified length and orientation to identify the boundary. Freeman chain coding is a linear structure derived through the quantization of the trajectory traced by the centers of adjacent boundary elements in a pixel array. The respective code is generated by the clockwise or anticlockwise scanning of the boundary and assigning the orientation value to the segment connecting each pair of pixels. This process for the shape and size identification results into a compact and translation independent representation of a binary contour. It also provides a lossless compression and preserving capacity for all morphological or topological information.

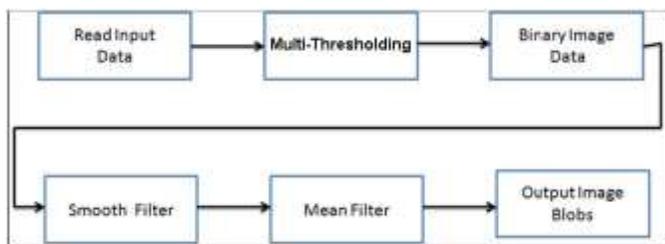


Fig. 3. Preprocessing Framework.

Classification and Recognition: Various features of blobs are extracted in this work to derive the dataset which is used for the training of the deep learning network. The selection of the features plays a particularly important role in defining the performance of the character recognition. In this work, the parameters selected and extracted for the training are based upon the blob analysis. These features resemble the shape, texture, blob area, perimeter, corners, etc. which are generated through the skeletonization process. The respective feature vector is derived by dividing the blob zones into various windows of equal size and the respective chain codes are used to derive the parameters like number of horizontal lines, vertical lines, right diagonal lines and left diagonal lines. These features are then used for the training of the proposed DWNN framework. It is responsible for the final decision making about the characters which are to be extracted from the text. The feature dataset is divided into two parts, training dataset, and testing dataset. The derived feature vector is denoted as X as.

$$X = f_1, f_2, \dots, f_n \tag{5}$$

Where n denotes the number of blob zones and f represents the respective feature sets. DWNN is trained with these feature sets to derive a recognition model for the overlapped and touching characters in the text.

The recognition model is derived from (1) as

$$\zeta = \omega^T \varphi(X, \tau, \sigma) + \mu^T \phi(X, \tau, \sigma) \tag{6}$$

The wavelet (φ) used in the proposed DWNN model is Mexican hat wavelet which has shown best resolution performance in analyzing the sharp features in the data distribution. It has two side-lobes and central peak as shown in Fig. 4 and therefore known as second Gaussian wavelet (g_2 wavelet) as well. It is the negative normalized second derivative of a Gaussian function. The mathematical description of Mexican hat wavelet can be derived by evaluating the second derivative of the Gaussian probability density function (pdf) as

$$\varphi(t) = \frac{2}{\pi^{\frac{1}{4}} \sqrt{3} \rho} \left(1 - \frac{t^2}{\rho^2} \right) e^{-\frac{t^2}{2\rho^2}} \tag{7}$$

where ρ represents the standard deviation of the Gaussian pdf.

The proposed DWNN model used for the text extraction comprises of three nodes and the weights and bias of the network are tuned and optimized using the gradient descent algorithm over the error value derived in (4) as:

$$\omega(n+1) = \omega(n) + \Delta\omega(n) = \omega(n) + \mu \left(-\frac{\partial e(n)}{\partial \omega(n)} \right) \tag{8}$$

The recognition accuracy is considered to evaluate the performance of the proposed DWNN based text extraction strategy.

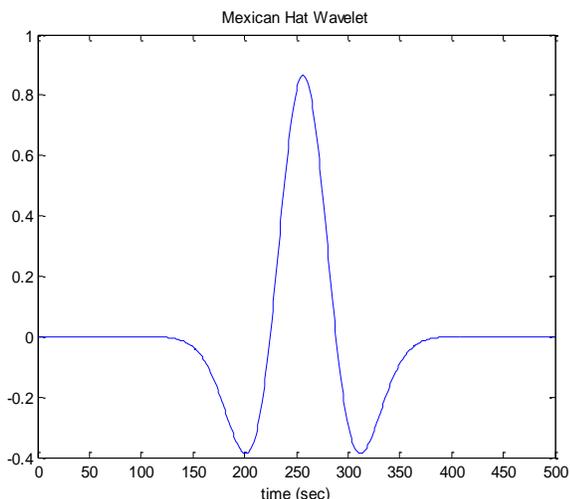


Fig. 4. Mexican Hat Wavelet.

V. EXPERIMENT RESULTS AND DISCUSSION

The performance of the proposed DWNN based text recognition with overlapped characters is assessed through the experimental analysis in MATLAB. The recognition capability of the proposed deep learning network is evaluated based on some metrics namely estimation error, cost function and Accuracy. More than 100 images with overlapped and non-overlapped characters are taken for the experiment with 80 % used for training and 20% for testing purpose. Fig. 5 shows the variation of cost function with respect to time and reflects the efficacy of the optimization algorithm for the tuning of the proposed DWNN. The variation of estimation error is shown in Fig. 6 which clearly reflects the learning characteristics of the network. The values of the estimation error are converging and ranging around zero as the learning goes on.

The performance of the proposed DWNN based character recognition is also shown by the accurate segmentation of the characters in the input images. The input images along with their respective segmentations are shown in Fig. 7 to 10. Although the accuracy of OCR in case of overlapped or touching character solely depends upon the quality of segmentation which relies on the quality of the image, the recognition capability has been evaluated in terms of average accuracy and found to be around 94 percent. The fast convergence rate of the proposed deep learning network can easily be deduced from the sharp decrease in the value of estimation error. It also represents the impact of using wavelet function as the activation function in the conventional neural network framework. The capability of wavelet as kernel in the neural network framework is also reflected from the efficient classification.

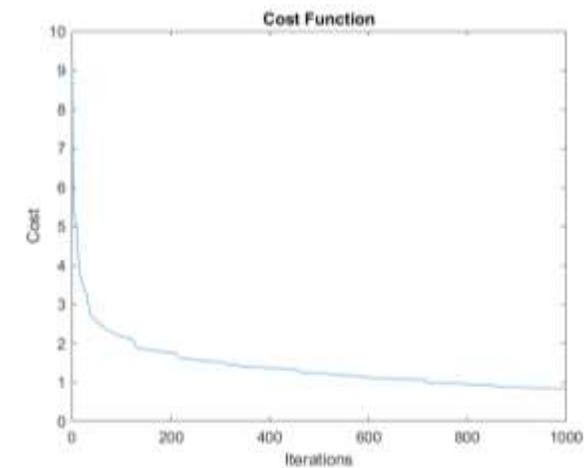


Fig. 5. Cost Function v/s Time.

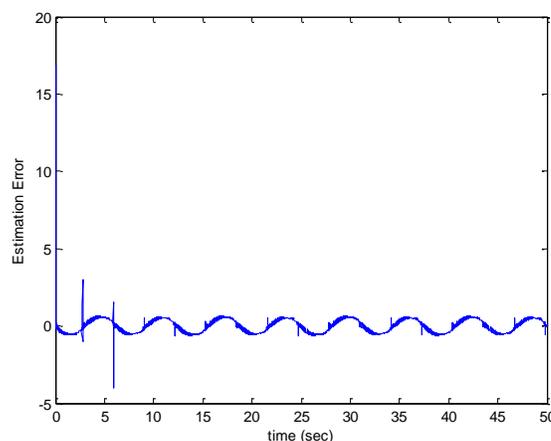


Fig. 6. Estimation Error v/s Time.



Fig. 7. Input Image 1.

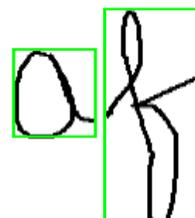


Fig. 8. Segmented Image 1.

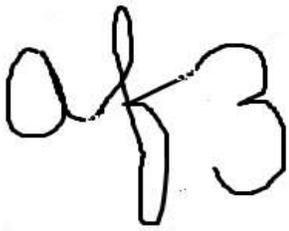


Fig. 9. Input Image 2.

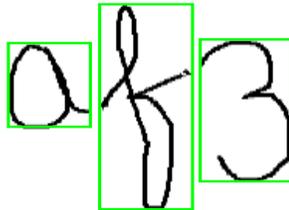


Fig. 10. Segmented Image 2.

VI. CONCLUSION

A deep learning based intelligent text recognition is presented in the paper which has used DWNN to enhance the learning capabilities of the conventional deep learning models. Mexican hat wavelet has been used as the activation function to mitigate the impact of uncertainties and randomness in the data distribution over an image. The challenging problem of segmentation and recognition of overlapped or touching characters has been addressed in this paper. The segmentation is achieved through the combination of blob adjacency analysis and Freeman chain coding technique. The extracted features are used for the training of the proposed DWNN model. Cascaded layered architecture of the translated and dilated versions of wavelet function is used to derive the complete framework of the DWNN. Tuning laws have been derived through gradient descent algorithm to achieve the optimal learning characteristics. The performance of the proposed deep learning framework is evaluated through the experimental analysis in terms of estimation error and cost function which has reached to a desirable range within an exceedingly small amount of time. The segmentation part of the process is found to be the most crucial aspect of OCR and it is subject to the nature of connectivity between the characters. The future aspect of this research is to add more robustness during the segmentation process to make the proposed model more effective even for the characters with high degree of superposition and overlapping.

REFERENCES

- [1] S. Nomura, K. Yamanaka, O. Katai, H. Kawakami, and T. Shiose, "A novel adaptive morphological approach for degraded character image segmentation," *Pattern Recognit.*, vol. 38, no. 11, pp. 1961-1975, Nov. 2005.
- [2] A. Rehman and T. Saba, "Performance analysis of character segmentation approach for cursive script recognition on benchmark database," *Digit. Signal Process.*, vol. 21, no. 3, pp. 486-490, May 2011.
- [3] P. P. Roy, U. Pal, J. Lladós, and M. Delalandre, "Multi-oriented touching text character segmentation in graphical documents using dynamic programming," *Pattern Recognit.*, vol. 45, no. 5, pp. 1972-1983, May 2012.

- [4] N. Arica and F. T.Yarman-Vural, "Optical character recognition for cursive handwriting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 6, pp. 801-813, Jun. 2002.
- [5] K. S. Dash, N. B. Puhan, and G. Panda, "BESAC: Binary external symmetry axis constellation for unconstrained handwritten character recognition," *Pattern Recognit. Lett.*, vol. 83, no. 3, pp. 413-422, Nov. 2016.
- [6] K. S. Dash, N. B. Puhan, and G. Panda, "Handwritten numeral recognition using non-redundant Stockwell transform and bio-inspired optimal zoning," *IET Image Process.*, vol. 9, no. 10, pp. 874-882, Sep. 2015.
- [7] G. Raju, B. S. Moni, and M. S. Nair, "A novel handwritten character recognition system using gradient based features and run length count," *Sadhana*, vol. 39, no. 6, pp. 1333-1355, Dec. 2014.
- [8] P. Singh, A. Verma, and N. S. Chaudhari, "Feature selection based classifier combination approach for handwritten Devanagari numeral recognition," *Sadhana*, vol. 40, no. 6, pp. 1701-1714, Sep. 2015.
- [9] U. Garain and B. B. Chaudhuri, "Segmentation of touching characters in printed Devnagari and Bangla scripts using fuzzy multi-factorial analysis," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 32, no. 4, pp. 449-459, Nov. 2002.
- [10] Y.-K. Chen and J.-F. Wang, "Segmentation of single- or multiple-touching handwritten numeral string using background and foreground analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1304-1317, Nov. 2000.
- [11] V. Bansal and R. M. K. Sinha, "Segmentation of touching and fused Devanagari characters," *Pattern Recognit.*, vol. 35, no. 4, pp. 875-893, Apr. 2002.
- [12] P. Sahare and S. B. Dhok, "Review of text extraction algorithms for scene text and document images," *IETE Tech. Rev.*, vol. 34, no. 2, pp. 144-164, Apr. 2016.
- [13] B. Zhu and M. Nakagawa, "A robust method for coarse classifier construction from a large number of basic recognizers for on-line handwritten Chinese/Japanese character recognition," *Pattern Recognit.*, vol. 47, no. 2, pp. 685-693, Feb. 2014.
- [14] G. A. Montazer, H. Q. Saremi, and V. Khatibi, "A neuro-fuzzy inference engine for Farsi numeral characters recognition," *Expert Syst. Appl.*, vol. 37, no. 9, pp. 6327-6337, Sep. 2010.
- [15] K. Verma and R. K. Sharma, "Comparison of HMM and SVM-based stroke classifiers for Gurmukhi script," *Neural Comput. Appl.*, vol. 28, pp. 51-63, Dec. 2016.
- [16] A. A. Desai, "Gujarati handwritten numeral optical character reorganization through neural network," *Pattern Recognit.*, vol. 43, no. 7, pp. 2582-2589, Jul. 2010.
- [17] Sampath AK, Gomathi N, "Decision tree and deep learning based probabilistic model for character recognition", *J Cent S Univ* 24:2862-2876, 2017.
- [18] Semwal VB, Gaud N, Nandi GC, "Human gait state prediction using cellular automata and classification using ELM", In: *Machine Intelligence and Signal Analysis*, p 135-145, 2019.
- [19] M. Sharma, A. Kulkarni and A. Verma, "Wavelet Adaptive Output Tracking Control for a Class of Delayed Uncertain MIMO Nonlinear Systems Subjected to Actuator Saturation", *Proceedings of IEEE international conference on advances in computing, control and telecommunication technologies, ACT-09, India*, pp. 705-710, 2009.
- [20] M. Sharma, A. Kulkarni and A. Verma, "Wavelet adaptive observer based control for a class of uncertain time delay nonlinear systems with input constraints", *Proceedings of IEEE international conference on advances in recent technologies in communication and computing, ARTCOM-09, India*, pp. 863-867, 2009.
- [21] Guruprasad P, Majumdar J, "Multimodal recognition framework: an accurate and powerful Nandinagari handwritten character recognition", *Modelb. Procedia Comput Sci* 89:836-844, 2016.
- [22] Kanimozhi VM, Muthumani I, "Optical character recognition for English and Tamil script", *International Journal of Computer Science and Information Technologies* 5(2):1008-1010, 2014.
- [23] Priyadarshni, Sohal JS, "Improvement of artificial neural network based character recognition system", *using SciLab*. 10:1-9, 2016.

- [24] A. Antonacopoulos and D. Karatzas, "Semantics-Based Content Extraction in Typewritten Historical Documents," in Eighth International Conference on Document Analysis and Recognition, pp. 48-53, 2005.
- [25] G. A. Farulla, N. Murru, and R. Rossini, "A fuzzy approach to segment touching characters," *Expert Syst. Appl.*, vol. 88, pp. 1-13, Dec. 2017.
- [26] U. Garain and B. B. Chaudhuri, "Segmentation of touching characters in printed Devnagari and Bangla scripts using fuzzy multi-factorial analysis," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 32, no. 4, pp. 449-459, Nov. 2002.
- [27] S. Nomura, K. Yamanaka, O. Katai, H. Kawakami, and T. Shiose, "A novel adaptive morphological approach for degraded character image segmentation," *Pattern Recognit.*, vol. 38, no. 11, pp. 1961-1975, Nov. 2005.
- [28] J. Tian, R. Wang, G. Wang, J. Liu, and Y. Xia, "A two-stage character segmentation method for Chinese license plate," *Comput. Elect. Eng.*, vol. 46, pp. 539-553, Aug. 2015.
- [29] L. Zheng, A. H. Hassin, and X. Tang, "A new algorithm for machineprinted Arabic character segmentation," *Pattern. Recogn. Lett.*, vol. 25, no. 15, pp. 1723-1729, Nov. 2004.
- [30] V. Bansal and R. M. K. Sinha, "Segmentation of touching and fused Devanagari characters," *Pattern Recognit.*, vol. 35, no. 4, pp. 875-893, Apr. 2002.
- [31] E. B. Lacerda and C. A. Mello, "Segmentation of connected handwritten digits using self-organizing maps," *Expert Syst. Appl.*, vol. 40, no. 15, pp. 5867-5877, Nov. 2013.
- [32] A. Elnagar and R. Alhadj, "Segmentation of connected handwritten numeral strings," *Pattern Recognit.*, vol. 36, no. 3, pp. 625-634, Mar. 2003.
- [33] A. Gattal and Y. Chibani, "SVM-based segmentation-verification of handwritten connected digits using the oriented sliding window," *Int. J. Comput. Intell. Appl.*, vol. 14, no. 1, Mar. 2015.
- [34] D. Ciresan, "Avoiding segmentation in multi-digit numeral string recognition by combining single and two-digit classifiers trained without negative examples," in Proc. 10th Int. Symp. Symbolic Numeric Algorithms Sci. Comput., Sep. 2008, pp. 225-230.
- [35] A. G. Hochuli, L. S. Oliveira, A. Britto, Jr., and R. Sabourin, "Handwritten digit segmentation: Is it still necessary?" *Pattern Recognit.*, vol. 78, pp. 1-11, Jun. 2018.
- [36] H. Zhan, S. Lyu, and Y. Lu, "Handwritten digit string recognition using convolutional neural network," in Proc. 24th Int. Conf. Pattern Recognit. (ICPR), Sep. 2018, pp. 3729-3734.
- [37] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499-1503, Oct. 2016.
- [38] Q. Zhang and A. Benveniste, "Wavelet networks," *IEEE Transactions on Neural Networks*, Vol. 3, no. 6, pp. 889-898, November 1992.
- [39] J. Zhang, G. G. Walter, Y. Miao, and W. Lee, "Wavelet neural networks for function learning," *IEEE Transactions on Signal Processing*, Vol. 43, no. 6, pp. 1485-1497, June 1995.
- [40] B. Delyon, A. Juditsky, and A. Benveniste, "Accuracy analysis for wavelet approximations," *IEEE Transactions on Neural Networks*, Vol. 6, no. 2, pp. 332-348 March 1995.