Journal of Computing and Communication Vol. 4, No. 2, PP. 89-99 , 2025

DEEP Learning Based Handwritten Recognition using Checkerboard Pattern

Sameh Zarif*[a,b], Mahmoud Hammam[c], Marian Wagdy[d]

[a]Department of Information Technology, Faculty of Computers and Information, Menoufia University, Menoufia, Egypt

[b]Department of Artificial Intelligence, Faculty of Artificial Intelligence, Egyptian Russian University, Egypt

[c]Department of Artificial Intelligence, Faculty of Computers and Information, Banha University, Banha, Egypt

[d]Department of Information Technology, Faculty of Computers and Information, Tanta University, Tanta, Egypt

*Corresponding Author: Sameh Zarif [***sameh.shenoda@ci.menofia.edu.eg, sameh-zarief@eru.edu.eg***]

ARTICLE DATA

Article history:
Received : 30 June 2025
Revised :26 July 2025
Accepted : 30 July 2025
Available online

Keywords:
LSTM
Checkerboard pattern
Handwritten recognition
OCR
document processing

ABSTRACT

This research addresses Egypt's Vision 2030 by developing an AI system for automating optical character recognition for documents. This paper uses a deep learning to eliminate manual data entry and improve data accuracy, saving time and resources for organizations, which aligns with the Vision's goals of using AI for document processing and better decision-making. The paper uses mainly the checkerboard pattern to fix the de-warpping data on the bills, then the proposed method will be able to recognize the content of each bill separately and categorize the data inside those bills. The proposed method uses the mix of image correction for de-warpping the image and the usage of LSTMs to recognize the oriented text in the handwritten document. Also the checkerboard is aiding in the usage of the LSTM algorithm which advances its performance to be better than the other state of the art techniques for both the RIMES and IAM datasets.

1. Introduction

The utilization of Optical Character Recognition technology has witnessed significant growth across various sectors, revolutionizing the way textual information is processed, extracted, and utilized. Particularly in the context of handwritten text, OCR plays a pivotal role in enabling the digital transformation of documents, facilitating efficient storage, retrieval, and analysis of textual data [1].

Moreover, we highlight the significance of OCR technology in commercial applications, emphasizing its role in streamlining document processing workflows, enhancing data accuracy, and enabling seamless integration with digital systems. The widespread adoption of OCR applications in industries such as finance, retail, logistics, and healthcare underscores the pressing need for advancements in handwritten OCR techniques, particularly in addressing challenges associated with document deformations.

Despite notable advancements, the accurate recognition of handwritten text remains a challenging task, especially when dealing with documents characterized by deformations, such as bills, invoices, and receipts. The inherent variability in handwriting styles, coupled with distortions caused by scanning or imaging processes, poses substantial hurdles for conventional OCR systems [2].

In recent years, deep learning approaches, mainly Recurrent Neural Networks such as Long Short Term Memory networks, have exposed remarkable promise in improving the accuracy of OCR systems for handwritten text recognition. By leveraging the temporal dependencies inherent in sequential data, LSTMs have demonstrated superior performance in capturing the contextual information crucial for deciphering handwritten characters [3].

Furthermor, addressing the de-warping problem, wherein the geometric distortions present in scanned documents hinder accurate character recognition, is paramount for enhancing the effectiveness of OCR systems, especially in the context of bill processing [4]. Traditional approaches to de-warping often fall short in handling complex distortions present in real-world documents.

The novelty of the proposed methodology is the integration of the power of LSTM networks with the checkerboard calibration pattern to tackle the de-warping problem in handwritten OCR, particularly focusing on bill processing scenarios. The checkerboard calibration pattern, widely recognized for its effectiveness in correcting geometric distortions, works in tandem with the sequence modeling power of LSTMs to achieve robust and precise recognition of handwritten text in warped or distorted documents. The proposed approach aims to contribute to the advancement of handwritten OCR technology, paving the way for more accurate, efficient, and versatile solutions capable of meeting the demands of modern document processing tasks, especially in commercial settings.

The rest of the paper is structured as follows: Section II provides a review of relevant works. Section III introduces the proposed idea of handwritten recognition. Section IV presents the experimental results obtained from the proposed method. Finally, Section V concludes the paper and discusses avenues for future research.

2. Related Work

Sign languages use the visual-manual modality to convey meaning, which makes them distinct from spoken languages in terms of articulation. While spoken languages rely mainly on the throat, nose, and mouth as articulators, sign languages primarily use the fingers, hands, and arms. In addition to these manual elements, sign languages also rely heavily on non-manual features—such as eyebrow movements, facial expressions, mouth shapes, eye and cheek movements, upper body shifts, head tilts, and shoulder movements. These non-manual features form an essential part of a meaningful sign. Without them, even if the signs are produced in the correct syntactic order, the intended message would lack clarity and be incomplete.

In the realm of Handwritten Text Recognition (HTR), recent research heavily leans on methods built with neural networks, similar to many computer vision tasks. Bouillon et al. [5] proposed an approach for children's handwriting; he focused on symbols and letters in different geometric shapes. They employed fuzzy inference systems with discriminative capabilities for evaluating handwritten symbols. Simonnet et al. [6] designed a system employing styluses for gathering handwritten words as signals on digital touchscreens. The signals go on correction using the EVOLVE classifier, it is emphasizing the French language dataset. Notably, a group of examples in this dataset featured missing letters.

Li et al. [7] introduced a method for addressing document image quality assessment that showed a no-reference method called MSER. They integrated OCR accuracy as a parameter to measure the quality of the document image. This approach is working by assuming that essential features were located in the form of patches, which might not perfectly correspond to human perception. It showed the impact of the overall quality of the document on aesthetic perception and used a dataset that featured characters in uniform shapes.

Schneider et al. [8] introduced a method that derives a vector field from a generated warping mesh using local orientation information. They corrected image distortions by approximating the nonlinear warp with multiple linear projections. To obtain the necessary local orientation features, they applied a detector. Meanwhile, B. Gatos et al. [9] proposed segmenting the document into individual words, then determining the orientation of each word. Finally, each word is rotated to correct the warping distortion.

Previous research has investigated different approaches for correcting document image curvature. Bukhari et al. [10] introduced a method that employs curled baseline pairs, mapping characters between corresponding top and bottom baselines to enhance image quality. Zhang et al. [11] used segmentation and thin plate splines for the restoration process. Lu et al. [12] identified text baselines and vertical stroke boundaries for each line. Bolelli et al. [13] proposed a technique that encloses each letter within a quadrilateral cell, adjusting the orientation by computing the center of each letter, resulting in a flattened word.

Leifert, G.[14], Strau, proposed a recurrent neural network model fitting the sequential nature of handwriting text. Krishnan, P., Dutta, K., Jawahar, C.[15] proposed a complex augmentation schemes and novel architectures s (e.g. Seq2Seq/Transformers, Spatial Transformer Networks [16], deformable convolutions [17]) and multi-task losses with auxiliary training feeds (e.g. n-gram training [18]).

The common approaches rely on text information, performing well when segmentation is accurate. However, these techniques have limitations, particularly when dealing with significant curvature or variable line spacing. In addition to their sensitivity to image resolution and the words used. Text based approaches are time-consuming and struggle with complex layouts, including documents containing graphics and tables. Additionally, these approaches are prone to numerous distortions, leading to a high rate of segmentation errors.

3. Proposed Work

Improving the accuracy of recognition can be achieved by eliminating issues resulting from document capturing, such as poor lighting and distortion. Eliminating these issues before recognizing the characters makes a significant difference in the quality of the result. The proposed idea investigates how the checkerboard calibration pattern can be integrated with LSTMs to achieve improved performance in handwritten recognition, specifically focusing on handling the challenges of dewarping handwritten characters. The idea of this research is divided into two parts: the first is addressing the aforementioned issues in the captured documents, and the second is document recognition based on deep learning. Image dewarping and illumination adaptation are crucial pre-processing steps for handwritten recognition as it corrects distortions in handwritten characters. In the first part, the checkerboard calibration pattern algorithm is applied for dewarping due to its efficiency and robustness. Retinex theory is demonstrated for the illumination enhancement. In the second part, a deep learning architecture specifically designed for recognizing handwritten bills is employed. Figure 1 presents the architected diagram of the proposed technique.

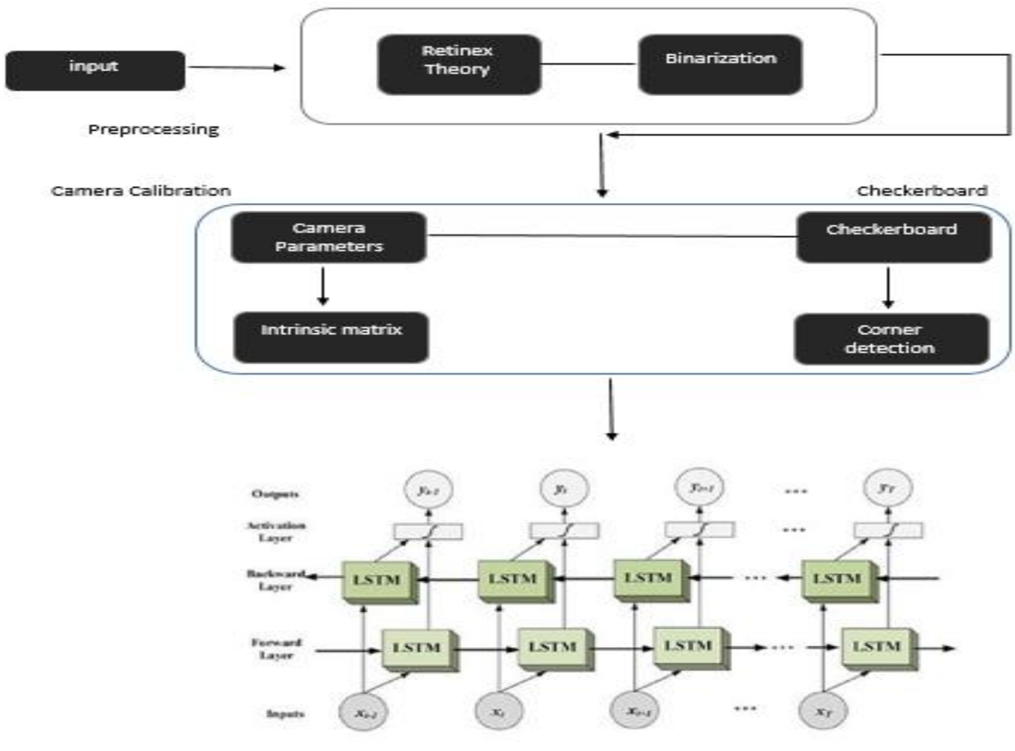

FIGURE 1. The general block diagram for the proposed English OCR system

3.1. Addressing Document Image Quality

The quality of captured bills can significantly impact OCR performance. Here, this paper proposes incorporating document image approaches to enhance the accuracy. The initial step analyzes the captured bill and identifies potential issues that may hinder accurate OCR, such as illumination variations and geometric distortions.

3.1.1. Illumination variations

Uneven lighting conditions can lead to shadows and uneven character recognition. Here, this paper leverages a technique based on Retinex theory to normalize illumination and enhance image quality as illustrated in figure 2. The Retinex concept is derived from the biological processes of the visual system. The Retinex theory is expressed in Equation 1, where I(x, y) represents the image intensity, with R(x, y) indicating reflectance and L(x, y) representing illumination, which can be approximated using the low frequency component of the image. If I(x, y) is a damaged document image, the Lightness image R(x, y) is acquired by dividing it by its smoothed one L(x, y), as shown in Equation 2.

$$I(x,y) = R(x,y).L(x,y) \quad (1)$$

$$L(x,y) = I(x,y)/M(x,y) \quad (2)$$

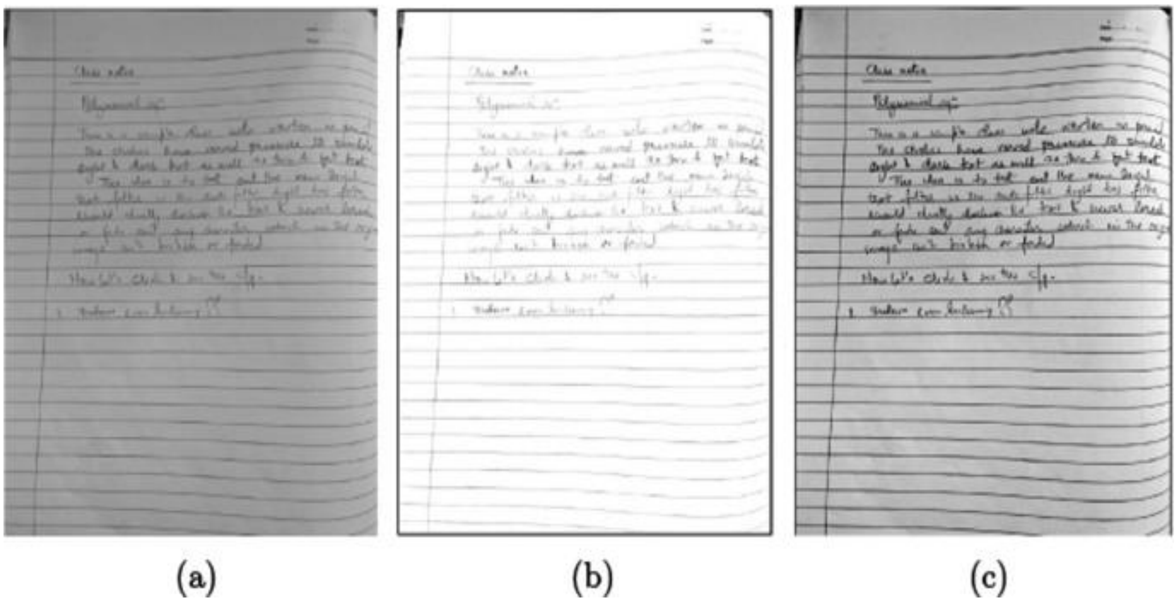

FIGURE 2. Enhancing image quality for the inputs entered to the OCR system

3.1.2. Geometric distortions

This section addresses warping and skewing issues, shown in figure 3, commonly encountered in images captured from bounding books or using mobile phone cameras, which utilizes checkerboard calibration patterns and camera calibration to correct distortions.

FIGURE 3. Geometric distortions of the images entered to the proposed OCR model

Checkerboard Calibration pattern is used to find the 3D shape of document image. Calibration patterns are common tools in computer vision for various purposes, like setting reference points, figuring out camera properties, and correcting image distortions as shown in figure 4. Here, the corners of the checkerboard in the image act as markers that help to calculate the 3D shape of the document. It's important to choose a clear checkerboard with at least 6 squares by 6 squares. The more squares there are, the better the results will be. The size of the checkerboard should also match the size of the document in the image.

FIGURE 4. Applying the checkerboard calibration pattern for the images after the enhancement

Checkerboard Corner detection: The image 3D shape is simulated by setting the 3D points $P(x, y, z, 1)$ of the checkerboard pattern. The markers are subsequently used to reconstruct the shape by calculating the corners of the checkerboard pattern. The corners are extracted from every white and black square that touches each other. Four squares are surrounding the corners which are (upper right, upper left, lower right, lower left,) as shown in figure 5.

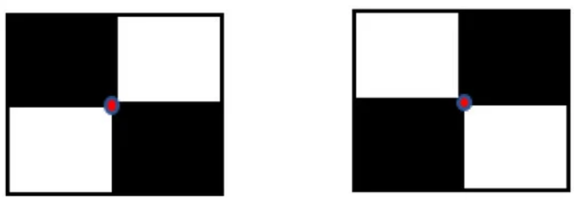

FIGURE 5. Checkerboard corners in the preprocessing step in the model

Camera Calibration: This section incorporates camera calibration techniques to rectify document images captured from various viewpoints. Camera calibration is crucial for accurately transforming 3D world points on the document into corresponding 2D pixel coordinates in the image.

The camera calibration process relies on two fundamental sets of parameters: intrinsic and extrinsic.

- **Intrinsic parameters:** These inherent characteristics of the camera define its internal properties, including:
    - Focal length (f): The distance, measured in millimeters, between the camera’s optical center and the image plane.
    - Principal point (x_c, y_c): Corresponds to the center of the image plane, often referred to as the camera center.
- **Extrinsic parameters:** Describe the pose of camera (position and orientation) relative to the world. They include:
    - Rotation vector (R): Represents the camera's rotation around the x, y, and z axes.
    - Translation vector (t): Denotes the camera's displacement in 3D space relative to the origin of the world coordinate system.

The relationship between these parameters and the image formation process can be expressed mathematically using the camera matrix (P).

Camera Matrix and Image Formation

The camera matrix (P) is a 3x4 matrix that combines both intrinsic and extrinsic parameters as shown in equarion 3:

$$P = K[R|t] \qquad (3)$$

where:

- K is the 3x3 intrinsic camera matrix:

$$K = \begin{bmatrix} f & 0 & x_c \\ 0 & f & y_c \\ 0 & 0 & 1 \end{bmatrix}$$

- [R | t] represents the 3x4 extrinsic parameter matrix, consisting of the rotation matrix (R) and the translation vector (t).

In the context of document image correction, the intrinsic camera parameters (f, x_c, and y_c) are typically unknown. However, by leveraging a well-defined calibration pattern like a checkerboard, we can estimate these parameters using the equations 4 and 5:

The center of the image plane that corresponding to camera center (x_c, y_c) can be calculated based on the checkerboard's dimensions (width: W, height: H) as represented by equation 4:

x_c, y_c = (W / 2, H / 2) (4)

The focal length (f) can be estimated using the checkerboard dimensions, field of view (assumed to be 180°), and the arctangent function as shown in equation 5:

$$f = W/2 * [tan(\Theta/2)]^{-1} \quad (5)$$

where Θ represents the field of view (180° in this case).

Once the intrinsic camera matrix (K) is estimated, it can be used to map a 3D world point (Dw) on the document to its corresponding 2D pixel coordinate (xp) in the image as shown in equation 6:

$$x_p\,(x_s\,, y_s) = K * D_w \quad (6)$$

This equation allows us to transform points from the 3D space of the document to their corresponding pixel locations in the captured image, facilitating the rectification process for accurate document analysis.

By addressing these quality issues, the pre-processed document images become more suitable for accurate character recognition in the subsequent stages.

3.2. Deep Learning-Based OCR for Handwritten Bills

This paper employs a deep learning architecture specifically designed for recognizing handwritten bills, incorporating the following key components:

Convolutional Neural Network (CNN) for Feature Extraction: The CNN extracts relevant features from the pre-processed image, capturing crucial information about the shape and structure of characters, employing convolutional layers with pooling operations to achieve effective feature extraction.

- Then, the Recurrent Neural Network (RNN) for Sequence Modeling: the proposed work used Long Short-Term Memory (LSTM) network. It is employed to capture the sequential nature of the script. This allows the model to effectively analyze the relationships between characters within a word or sentence, improving recognition accuracy, particularly for words with similar-looking characters. The proposed architecture will utilize a Bi-directional LSTM.

-Connectionist Temporal Classification (CTC) loss layer is a specialized loss function widely used in sequence-to-sequence tasks where the alignment between input and output sequences is unknown or variable in length. The CTC loss enables the model to learn to map an input sequence to a target sequence of labels without requiring explicit alignment between them. It does this by introducing a special blank token and allowing repetitions, then summing over all possible valid alignments that can produce the target sequence. In the context of handwriting recognition, the CTC loss layer allows the network to output variable-length text sequences directly from feature maps extracted by convolutional or recurrent layers, making it particularly suitable for recognizing text in scanned documents or natural scene images where character positions are not explicitly labeled.

- Named Entity Recognition (NER): After character recognition, Named Entity Recognition (NER) technique will be employed to identify and extract key information from the recognized text. This information may include names, dates, amounts, and other relevant entities specific to the bill format. By automatically extracting these entities, NER enables systems to structure unstructured text, making it more searchable and analyzable. For example, after converting handwritten or printed text into machine-readable form using OCR, NER can help detect and categorize important terms, which facilitates tasks like indexing, summarization, and knowledge graph construction.

The combined capabilities of these components enable the system to accurately recognize handwritten characters, understand the context of words and sentences, and extract critical information from the bill. Table 1 shows the architecture summary.

TABLE 1: The proposed architecture summary

Layer (type)	**Output shape**	**Param #**	**Connected to**
Image (InputLayer)	[(None, 64, 32, 1)]	0	[]
Conv1 (Conv2D)	(None, 64, 32, 32)	320	[‘image[0][0]’]
Pool1 (MaxPooling2D)	(None, 32, 16, 32)	0	[‘Conv1[0][0]’]
batch_normalizaion_2 (BatchNormalization)	(None, 64, 32, 32)	128	[‘pool1[0][0]’]
Reshape(Reshape)	(None, 32, 512)	0	[‘batch_normalization_2[0][0]’]
dense2 (Dense)	(None, 32, 16)	8208	[‘reshape[0][0]’]
Batch_normalization_3 (BatchNormalization)	(None, 32, 16)	0	[‘dense2[0][0]’]
bidirectional_1 (Bidirectional)	(None, 32, 256)	64	[‘batch_normalization_3[0][0]’]
label1 (InputLayer)	[(None, None)]	148480	[]
dense3 (Dense)	(None, 32, 42)	0	[‘bidirectional_1[0][0]’]
ctc_loss (CTCLayer)	(None, 32, 42)	10794	[‘label[0][0]’, ‘dense3[0][0]]
Total params **Trainable params:** **Non-trainable params:**		**167,994** **167,898** **96**	

4. Experimental Results

The proposed system will be evaluated using a comprehensive dataset of handwritten bills, encompassing diverse formats and handwriting styles. Metrics such as character recognition accuracy, word error rate (WER), and F1 score will be estimated to assess the system's performance. Feedback from the evaluation will be used to further refine and improve the model's robustness and accuracy.

The results are made as a comparison between normal pre-processing and the processing through applying the checkerboard technique for the wrapped image. The dataset, composed of 5,160 images, goes through augmentation to increase its size to 15,480. The dataset was split into training and testing sets, where 80% for the training and 20% for testing and validation. A total of 500 epochs was applied followed by choosing Adam optimizer with an exponential decay scheduler for the learning rate at a maximum value of 0.0001. A small learning rate is applied to the beginning to trigger the overfitting before achieving accuracy. The training loss reached 1.5730 at the final epoch, while the validation loss was 1.1795.

The proposed work used Character Error Rate (CER) and Word Error Rate (WER) metrics to assess the model’s performance across the dataset to compare across the testing set by comparing training and validation losses, which states that the total number of characters (n) including the spaces and determines the minimum number of insertions (i), substitutions (s) and deletions (d) of characters required to match the ground truth, the CER and WER can be calculated as the shown in equations 7 and 8.

$$CER\ = [numer\ of\ incorrect\ characters/n] * 100 \tag{7}$$
$$WER\ = [numer\ of\ incorrect\ words/total\ numer\ of\ words] * 100 \tag{8}$$

The comparison is made for both the flat image and the wrapped image to check the efficiency of the proposed method with and without checkerboard de-warping algorithm. Samples of the results taking into consideration CER and WER evaluation metrics are illustrated in table 2 and figure 6.

TABLE 2: Proposed work accuracy with and without de-warping

	Validation		Test	
	CER	WER	CER	WER
Using checkerboard De-warping technique	3.89	10.89	4.95	13.55
Without using checkerboard De-warping	4.09	11.65	5.23	14.40

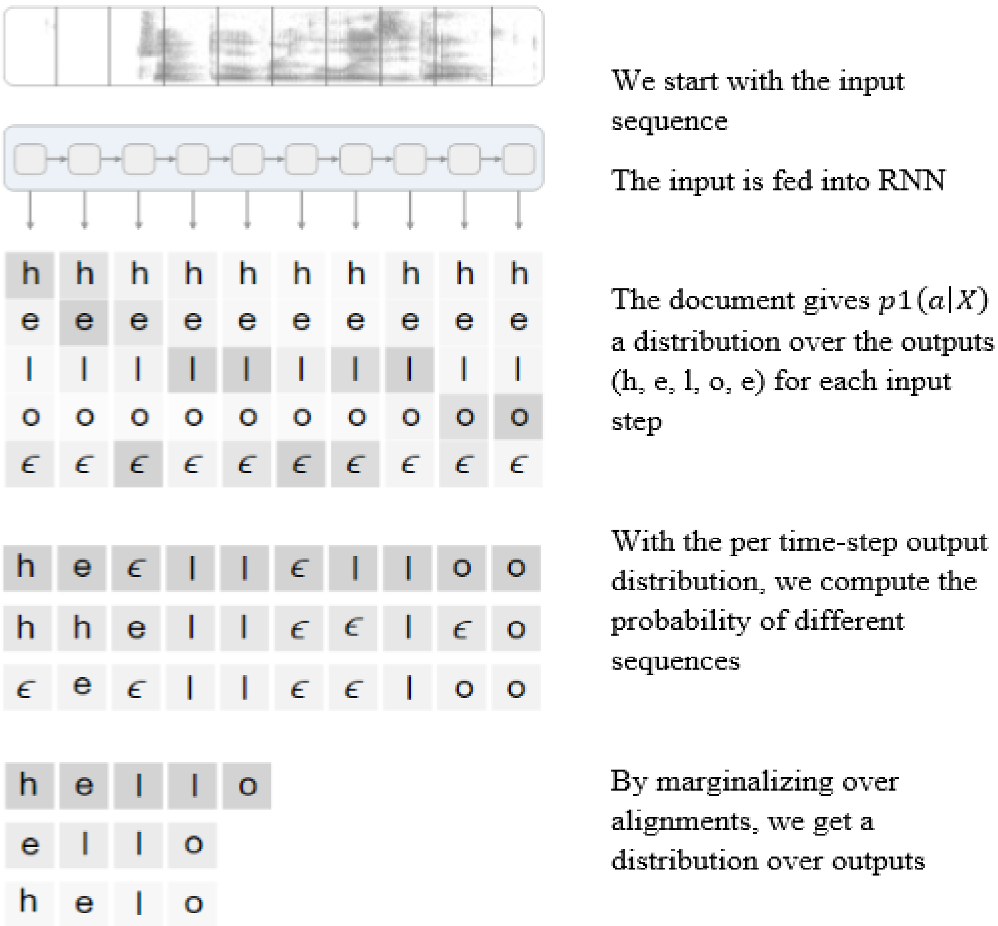

FIGURE 6. The output predictions for the OCR model after entering a random text

Comprehensive comparisons between the proposed method and the most state of the art methods are presented in table 3. The proposed model is compared on IAM and RIMES datasets, moreover our proposed system shows precise results compared to the other methodologies in handwriting recognition based on the same datasets. However, the accuracy of the proposed method is limited by its reliance on correctly detecting the document boundaries. Moreover, it is unable to reconstruct the document's shape when dealing with complex layouts, figures and tables.

TABLE 3: Proposed work compared to the other state of the art methods

	IAM		RIMES	
Method	CER(%)	WER(%)	CER(%)	WER(%)
In [18]	11.15	34.55	8.29	30.5
In [19]	10.8	35.1	6.8	28.5
In [14]	9.78	32.89		
In [20]	8.10	16.70	3.59	9.60
In [21]	6.2	20.2	2.60	10.7
In[22]	6.14	20.04	3.34	11.23
In [15]	5.8	17.8	5.07	14.7
In[24]	5.24	-		
In [17]	5.18	17.68		
In [25]	4.9	-		
In [16]	4.55	16.08	3.04	10.56
In [26]	4.62	15.89	2.75	9.93
Proposed	**4.53**	**15.99**	**2.67**	**9.83**

5. Conclusion

In conclusion, this research proposes a novel methodology for recognizing and interpreting handwritten bills. The proposed method effectively combines image correction for de-warping with LSTMs for recognizing oriented text in handwritten documents. By integrating document image quality valuation with a deep learning based OCR specifically tailored for script, the proposed method achieved superior accuracy and efficiency in automating data extraction from these critical business documents. By applying the proposed straightforward modifications, the proposed model achieved results that are outperforming the compared techniques for both the RIMES and IAM datasets.

References

[1] Ibrahim, M., Wagdy, M., AlHarithi, F. S., Qahtani, A. M., Elkilani, W. S., & Zarif, S. An Efficient Method for Document Correction Based on Checkerboard Calibration Pattern. Applied Sciences, 12(18), 9014, 2022.

[2] Baird, H.S., 2007. The state of the art of document image degradation modelling. Digital Document Processing: Major Directions and Recent Advances, pp.261-279.

[3] Wagdy, M., Amin, K., & Ibrahim, M. Dewarping document image techniques: survey and comparative study. International Journal of Image and Graphics, 21(03), 2150031, 2021.

[4] Gatos, B. and Ntirogiannis, K., 2007, February. Restoration of arbitrarily warped document images based on text line and word detection. In 4th IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (pp. 203-208).
[5] Bouillon M, Anquetil E. Handwriting analysis with online fuzzy models. In17th Biennal Conference of the International Graphonomics Society (IGS) 2015 Jun 21.
[6] Simonnet D, Girard N, Anquetil E, Renault M, Thomas S. Evaluation of children cursive handwritten words for e-education. Pattern Recognition Letters. 2019 Apr 15;121:133-9.
[7] Li H, Zhu F, Qiu J. CG-DIQA: No-reference document image quality assessment based on character gradient. In2018 24th International Conference on Pattern Recognition (ICPR) 2018 Aug 20 (pp. 3622-3626).
[8] Schneider, D.; Block, M.; Rojas, R. Robust document warping with interpolated vector fields. In Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Curitiba, Brazil, 23–26 September 2007; Volume 1, pp. 113–117.
[9] Bukhari, S.S.; Shafait, F.; Breuel, T.M. Dewarping of document images using coupled-snakes. In Proceedings of the Third International Workshop on Camera-Based Document Analysis and Recognition, Barcelona, Spain, 23–24 July 2009; pp. 34–41.
[10] Zhang, Y.; Liu, C.; Ding, X.; Zou, Y. Arbitrary warped document image restoration based on segmentation and Thin-Plate Splines. In Proceedings of the 2008 19th International Conference on Pattern Recognition, Tampa, FL, USA, 8–11 December 2008; pp. 1–4.
[11] Lu, S.; Chen, B.M.; Ko, C.C. A partition approach for the restoration of camera images of planar and curled document. Image Vis. Comput. 2006, 24, 837–848.
[12] Bolelli, F. Indexing of historical document images: Ad hoc dewarping technique for handwritten text. In Italian Research Conference on Digital Libraries; Springer: Cham, Switzerland, 2017; pp. 45–55.
[13] Leifert, G., Strau, T., Gr, T., Wustlich, W., Labahn, R., et al.: Cells in multidimensional recurrent neural networks. Journal of Machine Learning Research 17(97), 1–37 (2016).
[14] Krishnan, P., Dutta, K., Jawahar, C.: Word spotting and recognition using deep embedding. In: 2018 13th IAPR International Workshop on Document Analysis Systems (DAS). pp. 1–6. IEEE (2018).
[15] Dutta, K., Krishnan, P., Mathew, M., Jawahar, C.: Improving CNN-RNN hybrid networks for handwriting recognition. In: 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). pp. 80–85, 2018.
[16] Retsinas, G., Sfikas, G., Nikou, C., Maragos, P.: Deformation-invariant networks for handwritten text recognition. In: 2021 IEEE International Conference on Image Processing (ICIP). pp. 949–953, 2021.
[17] Tassopoulou, V., Retsinas, G., Maragos, P.: Enhancing handwritten text recognition with n-gram sequence decomposition and multitask learning. In: 2020 25th International Conference on Pattern Recognition (ICPR). pp. 10555–10560, 2021.
[18] Chen, Z., Wu, Y., Yin, F., Liu, C.L.: Simultaneous script identification and handwriting recognition via multi-task learning of recurrent neural networks. In: 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). vol. 1, pp. 525–530, 2017.
[19] Pham, V., Bluche, T., Kermorvant, C., Louradour, J.: Dropout improves recurrent neural networks for handwriting recognition. In: 2014 14th international conference on frontiers in handwriting recognition. pp. 285–290. 2014.
[20] Chowdhury, A., Vig, L.: An efficient end-to-end neural model for handwritten text recognition, 2018.
[21] Puigcerver, J.: Are multidimensional recurrent layers really necessary for handwritten text recognition? In: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). vol. 1, pp. 67–72. 2017.
[22] Markou, K., Tsochatzidis, L., Zagoris, K., Papazoglou, A., Karagiannis, X., Symeonidis, S., Pratikakis, I.: A convolutional recurrent neural network for the handwritten text recognition of historical greek manuscripts. In: International Workshop on Pattern Recognition for Cultural Heritage (PATRECH), 2020.
[23] Wick, C., Z¨ollner, J., Gr¨uning, T.: Transformer for handwritten text recognition using bidirectional post-decoding. In: International Conference on Document Analysis and Recognition. pp. 112–126, 2021.
[24] Michael, J., Labahn, R., Gr¨uning, T., Z¨ollner, J.: Evaluating sequence-to-sequence models for handwritten text recognition. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). pp. 1286–1293, 2019.
[25] Yousef, M., Hussain, K.F., Mohammed, U.S.: Accurate, data-efficient, unconstrained text recognition with convolutional neural networks. Pattern Recognition 108, 107482, 2020.
[26] Retsinas, G., Sfikas, G., Gatos, B., & Nikou, C. " Best practices for a handwritten text recognition system. arXiv:2404.11339v1, 17 Apr 2024.