Journal of Computing and Communication Vol.3 , No.1 , PP. 44-54 , 2024

Handwritten Arabic Bills Reader and Recognizer

Sara Sweidan *[a,b], Hesham Mahmoud Yasser Hammam [a,c]

[a] *Faculty of Computers and Artificial Intelligence, Artificial Intelligence Department, Benha University, Benha 13512, Egypt.*

[b] *New Mansoura University, Faculty of Computer Science and Engineering, Gamasa, 35712, Egypt.*

[c] *Faculty of Artificial Intelligence, Egyptian Russian University, Cairo 11829, Egypt.*

*Corresponding Author: Sara Sweidan [***Sweera20@yahoo.com]***

ARTICLE DATA

Article history:
Received 10 Oct 2023
Revised 17 Dec 2023
Accepted 27 Jan 2024
Available online

Keywords:
Handwritten Bills
Deep learning
Named Entity Recognition.
LSTM

ABSTRACT

In pursuit of Egypt's Vision 2030, which emphasizes the pivotal role of governance within state institutions and society, artificial intelligence (AI) stands as a transformative force. A central tenet of this vision involves harnessing AI technologies to accelerate the digitization of documents and their seamless integration into a unified system. This fosters more informed decision-making processes and revolutionizes processing and utilizing information. Our current research project aligns with this broader goal by deep learning capabilities to support organizations involved in enterprise applications. By incorporating AI-driven solutions, we aim to empower these organizations to manage their operations and optimize resource allocation efficiently. The proposed model eliminates manual input of handwritten invoices into ERP applications, resulting in substantive cost savings.

Furthermore, the proposed model integrates an entity classification system enhanced by LSTM, significantly improving invoice data's clarity and accuracy. This streamlined approach saves valuable time and enhances the overall effectiveness of resource allocation and decision-making processes. In essence, by integrating AI into document management and enterprise operations, we are not only contributing to the realization of Egypt's vision but also spearheading a technological transformation that has far-reaching implications for governance, efficiency, and progress in the digital age.

1. Introduction

Nowadays, the utilization of computers and digital media has surged across nearly all industries. Creating a paperless setting requires converting documents like receipts, invoices, assignments, and answer sheets into digital formats. Many tools like scanners, digital cameras, mobile phone cameras, and other devices can be employed. A document represents many types of written, printed, or electronic content that serves as an official record providing evidence or information. Types of documents include printed materials, handwritten notes, and mixed media. Documents can undergo various processes such as scanning, transmission, printing, decompression, and compression, leading to degradation. This degradation may occur singularly or in combinations, resulting in unexpected outcomes [1]. Additionally, lighting conditions during document capture can affect the quality of the image due to environmental or physical influences. Handwriting is an important skill for conveying information and is associated with legibility and kinematics [2], which influence document quality [3].

Current methods for assessing document quality heavily rely on manual labor, having several drawbacks, including inconsistent evaluation standards among individuals, low expertise, and great labor costs. While recognition, document image classification, and character detection often involve image processing, handwritten document quality assessment is often overlooked [3]. Evaluation can occur at different granularities, including page, phrase, word, and character levels. Present assessment models predominantly focus on character-level evaluations, which involve comparing symbols to a predefined reference model. These models typically encompass five criteria: direction, kinematics, order, form, and position relative to reference lines [4]. Page-level degradation includes challenges such as background noise (e.g., salt and pepper noise), ink spreading, cluttering, and structural distortions caused by skewing, warping, folding, and curling. [1]. Document Image Quality Assessment (DIQA) typically follows one of two approaches: traditional or modern approaches, such as Machine Learning.

Conventional techniques in image processing have traditionally been employed to analyze images, which include metrics or a reference for comparison. Image acquisition, edge detection, pre-processing, and feature extraction are commonly used in character or word evaluation. To begin, A physical document in hardcopy form is digitized through a scanner or by capturing an image of the document using imaging tools such as a digital camera or the camera on a mobile phone. The resulting image is initially considered raw data, necessitating further processing to remove any noise or distortions that might adversely affect the quality of the results. This processing phase also includes resizing the images to create a standardized dataset, eliminating the requirement for costly recurrent symbol alignment and establishing a consistent baseline for analysis, simplifying the data organization process. [5-6].

Edge Detection is a method used to segment an image into regions of discontinuity, facilitating the identification of features in an image that exhibit substantial changes in gray levels. This process trims the data in an image while retaining its essential structural attributes. A diverse array of features greatly enhances the recognition of handwritten characters. The extraction of robust features is key to representing handwritten text during the feature extraction stage [6]. Training and testing times can be reduced by reducing the number of features and focusing on relevant ones [7].

Conversely, traditional approaches that depend on a single feature extraction method or deep learning methods inspired by the human nervous system have enhanced performance in diverse domains. They also facilitate the exploration of certain features that may not be easily observable through visual examination. [8]. Deep learning offers the benefit of encompassing various-level features, whereas conventional state-of-the-art methods concentrate on low-level or high-level features. Despite the remarkable performance delivered by deep networks and the advantages of decreased training expenses and time, there is still a potential for overfitting and an escalation in training errors [9]. Integrating an automated system for reading handwritten Arabic bills into Enterprise Resource Planning (ERP) systems is a critical motivation in digital transformation, aiming to enhance efficiency and accuracy in business operations. The research goals outlined below will guide the development of this system:

- Develop a robust system for recognizing and extracting information from handwritten Arabic bills.
- Implement state-of-the-art deep learning techniques, including Optical Character Recognition (OCR), Named Entity Recognition (NER), and Long Short-Term Memory (LSTM) networks.
- Address challenges specific to Arabic script, considering variations in handwriting styles and contextual nuances.
- Design the system to eliminate the need for manual data entry, reducing processing time and minimizing the risk of errors.
- Ensure the system can handle various formats and styles of handwritten Arabic bills commonly encountered in business transactions.
- Develop an interface seamlessly integrating the recognized handwritten bill data into existing Enterprise Resource Planning (ERP) systems.
- Ensure real-time or near-real-time synchronization to provide up-to-date information for decision-making.
- Conduct thorough testing to validate the system's performance across different handwriting styles and bill formats.
- Iterate and refine the system based on feedback and performance metrics.

The proposed model empowers businesses to make informed, data-driven decisions while fostering growth and competitiveness through enhanced operational efficiency.

In the paper, section 2 discusses existing literature, section 3 describes the proposed method, section 4 provides a conclusion for the problem statement, section 5 discusses the results obtained, followed by a conclusion in section 6.

2. Literature Review

We can categorize DIQA methods into learning-based and metric-based assessment methods and reference and no-reference approaches based on feature extraction techniques. The following sections will delve into DIQA within two categories: the Traditional approach and the Deep Learning-based methods [10].

1-Traditional approaches

Bouillon et al. [11] presented an approach to assess children's handwriting, encompassing symbols and letters in different geometric shapes. They employed fuzzy inference systems with generative and discriminative capabilities for evaluating handwritten symbols, focusing on geometric attributes such as shape and direction. The investigation considered three sets of features to appraise symbol morphology, sequence, and direction, underscoring a significant reliance on the input data.

Akouaydi et al. [4] focused on assessing the quality of Arabic letter handwriting. They adopted a methodology incorporating the Beta elliptical character segmentation model to meticulously scrutinize individual characters and the Cartesian Fourier Descriptor model for character shape and boundary analysis. Data acquisition involved the use of a tablet, which required preprocessing. It's worth noting that the system was trained on a relatively small dataset consisting of just 20 correctly written samples for each letter. This limited dataset size could potentially raise concerns regarding the results' reliability.

Simonnet et al. [12] designed a system that employed styluses for gathering handwritten words as online signals on digital touchscreens. The collected signals underwent correction using the EVOLVE classifier, primarily emphasizing the French language dataset. Notably, a subset of examples in this dataset featured missing letters, which spurred an investigation into different segmentation techniques and hypotheses.

Simonnet et al. [13] pursued an alternative approach that focused on creating a cursive writing analysis system, providing holistic feedback derived from global, directional, geometric, and sequential characteristics. This system compared scores between classes and within the same class against a reference model. The multi-criteria classifier combined results from individual classifiers assessing the above features to yield a comprehensive outcome. The evaluation used IntuiScript, a digital notebook, and integrated stable attributes of block letters to augment the approach.

Li et al. [10] introduced a method for Document Image Quality Assessment (DIQA) that harnessed a no-reference technique known as MSER. They additionally integrated OCR accuracy as a parameter for measuring quality. This approach operated assuming that essential features were located within patches, which might not perfectly correspond to human perception. It recognized the impact of the overall quality of the document on aesthetic perception and used a dataset that featured characters with uniform shapes.

Kulesh et al. [14] presented a model that relied on low-level characteristics, including aspect ratio, zero-crossing distributions, and width distributions along the character's height. These attributes were then transformed into high-level feature vectors. The assessment and scoring of each handwritten letter were achieved through a fusion of artificial neural network and expert system approaches. However, it should be noted that these methods require a degree of complexity in comprehension and application.

Nayef et al. [15] utilized no-reference techniques and metric-based approaches to evaluate the quality of documents. Their assessment criteria included sharpness quality and character quality metrics. Sharpness was quantified using the LPC Sharpness Index and Log-Gabor filters of different scales. Character quality metrics considered factors such as black and white noise and the presence of overlapping or touching characters. The primary emphasis of the study was on the specific evaluation of distortions in scanned and mobile-captured images of documents, including bills and receipts. Additionally, the study encompassed the analysis of 175 images from the Tobacco dataset, with results being compared to human visual perception.

Alaei's study [16] introduced a novel full-reference Document Image Quality Assessment (DIQA) method, relying on Hast derivations. This approach involved generating similarity maps for both the reference and distorted images and evaluating the quality of the distorted document image by applying average pooling. The study's findings demonstrated that the second-order Hast derivation exhibited superior performance in the context of document images. In contrast, the first-order Hast derivation proved more effective for images of natural scenes.

2- Deep-learning approaches

Learning-based DIQA approaches employ discriminative features to address different document degradation forms using techniques like deep learning [17,18]. In their work, Gao et al. [3] introduced and evaluated a significant algorithm that relies on Convolutional Neural Networks (CNN) to evaluate English handwriting. They conducted a comparative study against well-established classification methods. The dataset was divided into two categories: award-winning and non-award-winning handwriting. To address the dataset's class imbalance, they trained the ResNet-18 model using weighted samples, utilized the soft-max function for loss computation, and employed Stochastic Gradient Descent (SGD) as the optimization technique. Kang et al. [17] adopted a machine-learning approach to evaluate the quality of two datasets: the SOC dataset, which comprised 175 color images, and the Newspaper dataset, which consisted of 571 grayscale images. Their method employed a Convolutional Neural Network (CNN) with two convolution layers, location-blind max-min pooling, and Rectified Linear Units (ReLU) in the fully connected layers. They computed patch scores and averaged them across the entire image to derive the document score. However, it's important to note that the dataset primarily comprises printed documents, and as such, it lacks significant diversity. While segmenting the image into patches can potentially provide more samples for the CNN, it may not necessarily align with or reflect human visual perception. Similarly, Lu et al. [19] proposed an approach by training and fine-tuning a Deep CNN with three fully connected layers while incorporating OCR as a quality descriptor.

Peng et al. [18] introduced a method for Document Image Quality Assessment (DIQA) that assesses the quality score of the target image utilizing sparse code learning. This methodology entails encoding the quality of an image into a codebook using a pooling technique, effectively converting the entire document image into a vector. Consequently, each patch within a training image can be represented as a combination of code words, thereby establishing a model for linking images with OCR confidences. The training process incorporates the use of linear regression. Similarly, Alaei et al. [20] utilized an unsupervised approach, trained on the ITESOFT dataset, in which the image was partitioned into patches to generate a Bag of Words (BoW) representation. In the testing phase, features extracted from these patches are allocated to clusters and subjected to average pooling. This process yields an Image Quality Assessment (IQA) score on a scale from 0 to 1, where 0 represents lower quality, while a score closer to 1 signifies higher quality.

Li et al. [21] introduced a Document Image Quality Assessment (DIQA) model based on Recurrent Neural Networks (RNN). This model incorporates a crucial convolutional layer for feature extraction from specific image regions identified through a spatial glimpse mechanism. The training process is carried out incrementally, utilizing reinforcement learning and Stochastic Gradient Descent (SGD) on the Smartdoc-QA and SOC datasets. Nevertheless, it encountered challenges when dealing with issues such as shadows in the Smartdoc-QA dataset, ultimately leading to performance that fell short of optimality.

From the previous studies, image quality assessment can be categorized from traditional methodologies to more advanced as deep learning methods. Current state-of-the-art research predominantly emphasizes character-level and word-level degradation analysis. However, this paper's primary concern is tackling degradation at the Arabic page level. It's worth noting that there is currently no easily accessible benchmark dataset encompassing a wide array of real-time degradations. The main Contribution of this research is to eliminate the need for manual input of handwritten invoices into ERP systems, resulting in major cost and time savings. It incorporates an entity classification system enhanced by LSTM, significantly improving the accuracy and clarity of extracted invoice data. By integrating AI-driven handwritten text recognition into document management and business operations, this research enables data-driven decision-making, enhanced resource allocation, and progress toward Egypt's digital transformation vision.

3. Proposed Methodology

1- Proposed system

This research paper proposes a methodology that leverages OCR (Optical Character Recognition), Named Entity Recognition (NER), and Deep Learning techniques to define and recognize Arabic handwritten bills. The initial phase involves the collection of a suitable dataset comprising Arabic handwritten bills, followed by comprehensive data preprocessing steps to enhance image quality and prepare the data for analysis. We will employ a deep learning-based OCR model for the OCR component, specifying the architecture, framework, and training process, including loss functions, optimization algorithms, and data splits for model assessment. Simultaneously, NER techniques will be applied to identify and extract key named entities from the recognized text, aiding in bill categorization and organization. This integrated approach aims to efficiently process Arabic handwritten bills, offering potential applications in document digitization, organization, and data extraction tasks.

2- Workflow of the model

In pursuit of our goal to develop an intelligent system for recognizing and interpreting handwritten Arabic bills, we have designed a multi-layered architecture leveraging state-of-the-art deep learning techniques. This section outlines the key components of the proposed handwritten Arabic bills (HAB) framework as presented in Figure 1, including the Convolutional Neural Network (CNN) feature extraction phase, batch normalization, and the subsequent Recurrent Neural Network (RNN) layers.

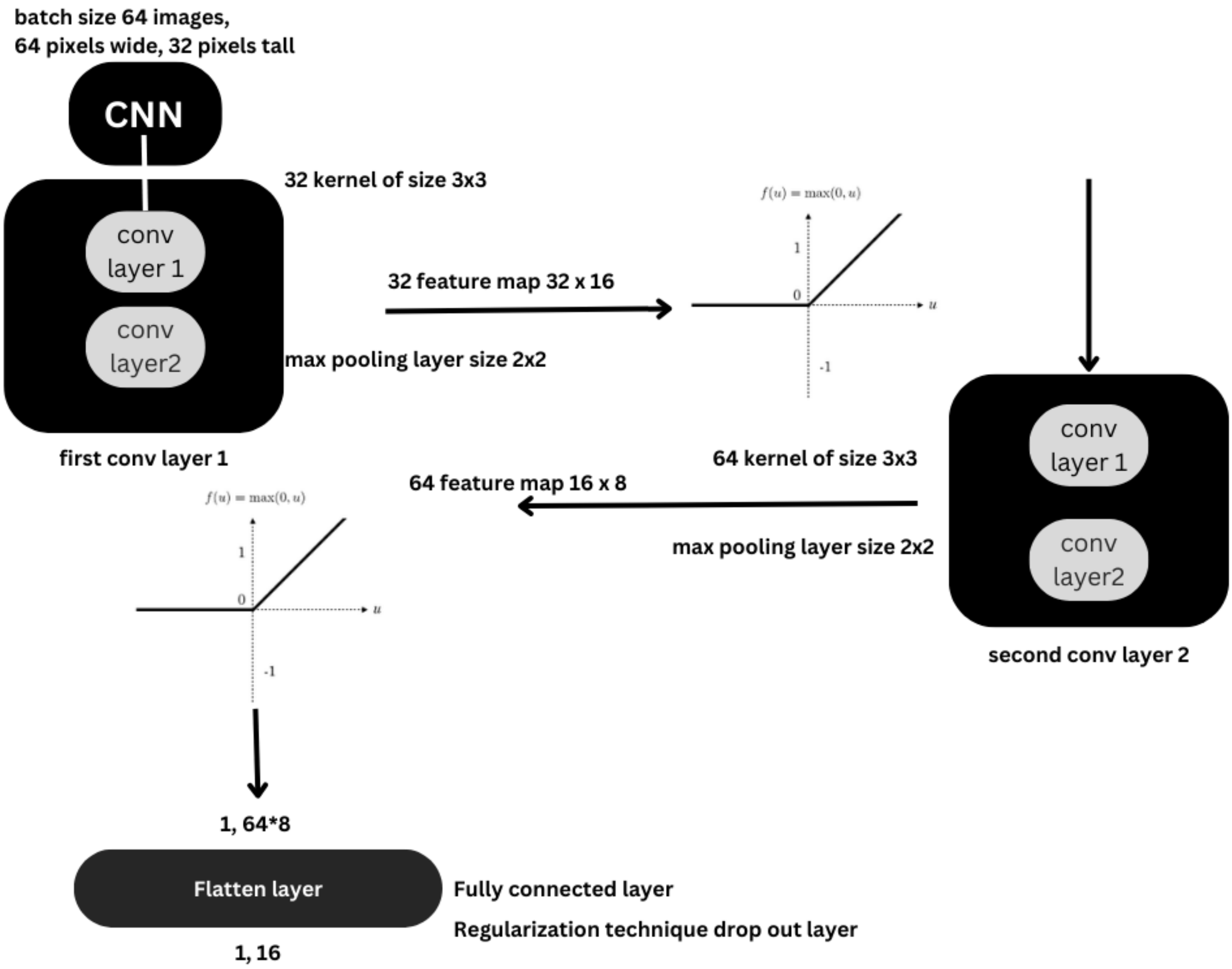

Figure 1: proposed Handwritten Arabic Bills (HAB) framework.

2.1 - Feature extraction

Our architecture begins with the Convolutional Neural Network (CNN) for feature extraction. This phase consists of two convolutional layers: a batch normalization layer, a Rectified Linear Unit (ReLU) activation function, and a fully connected dense layer. The first CNN layer employs 32 kernels of size (3,3) and includes a max pooling layer of size (2,2). This configuration yields 32 feature maps, each with dimensions of 32 pixels in width and 16 pixels in height. The second CNN layer further refines these features with 64 kernels of size (3,3) and another max pooling layer of size (2,2). This results in 64 feature maps, each measuring 16 pixels in width and 8 pixels in height. These feature maps are then flattened into a sequence of numerical values to be processed by the subsequent Recurrent Neural Network (RNN).

2.2 - learning algorithm

After the CNN feature extraction phase, the processed feature sequence is passed to the Recurrent Neural Network (RNN) component. Specifically, we employ a bidirectional Long Short-Term Memory (Bi-LSTM) layer with 256 LSTM units in both forward and backward directions, as shown in Figure 2 (a), totaling 128 units in each direction using equations (1-6). This bidirectional architecture allows the model to effectively capture dependencies between data entries in the sequence. We have chosen a dropout ratio of 0.35 to mitigate overfitting for regularization purposes.

$$i_t = \sigma(x_t\, U^i + \; h_{t-1} W^i) \tag{1}$$

$$f_t = \sigma\,(x_t U^f + h_{t-1} W^f \qquad (2)$$

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \qquad (3)$$

$$C_t = tanh(x_t U^g + h_{t-1} W^g) \qquad (4)$$

$$c_t = \sigma(f_t * C_{t-1} + i_t * C_t) \qquad (5)$$

$$h_t = \tanh(c_t) * o_t \qquad (6)$$

The final layer in our architecture is the soft-max output layer, which provides the recognized and interpreted results. In summary, our architecture's combination of CNN for feature extraction, as shown in Figure 2 (b), batch normalization, a bidirectional Bi-LSTM layer, and soft-max output presents a robust and effective solution for recognizing and interpreting handwritten Arabic bills. These components work in tandem to capture intricate patterns in the data and yield accurate results for integration into business.

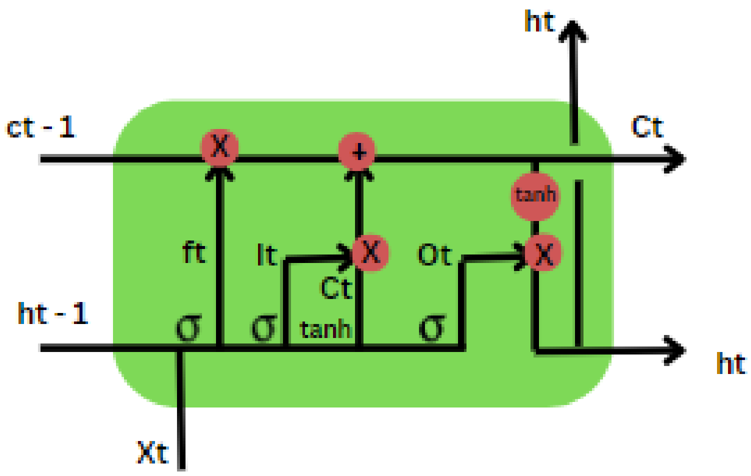

Figure 2 (a)**:** learning algorithm

Within our architecture, the Soft-max output layer is pivotal in deciphering the recognized and interpreted results. It is configured to have one additional unit beyond the number of characters.

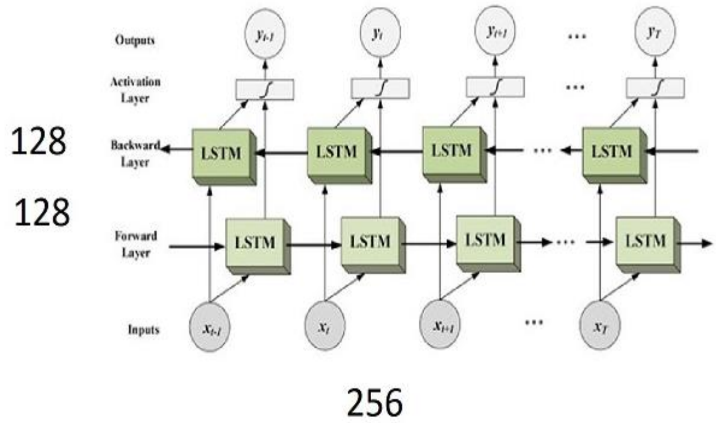

Figure 2 (b): model architecture

present in our defined vocabulary. The activations of these units are interpreted as the probabilities associated with observing specific characters or a "blank" character at a particular time step in the sequence. We employ the Connectionist Temporal Classification (CTC) loss function to facilitate extracting meaningful information from the sequence. The CTC loss function serves as a crucial tool in determining the most probable labeling for the input sequence. It accomplishes this by utilizing a many-characters-to-one-character mapping mechanism, eliminating redundant or blank characters from the predictions. By doing so, the CTC loss function calculates the conditional probability of encountering a given character in the sequence.

To decode the most probable labeling from the predictions, we apply the prefix/beam search decoding algorithm. This decoding method enhances the accuracy of the results by efficiently exploring and selecting the most likely character sequence, aligning with our objective of providing precise and meaningful interpretations of the handwritten Arabic bills. To prevent overfitting, we have incorporated a dropout layer into our architecture. Furthermore, a batch size of 64 images per batch was selected to optimize the training process.

In summary, our architecture's Soft-max output layer, combined with the CTC loss function and the prefix/beam search decoding algorithm, contributes significantly to the accuracy and reliability of our system's character recognition and interpretation capabilities.

2.3 - Dataset Description

In this research section, a crucial step involved the creation of a custom dataset by meticulously transcribing handwritten Arabic bills. This dataset was meticulously curated to encompass various handwriting styles, formats, and content, ensuring the model is exposed to a representative sample of real-world handwritten bills. Each handwritten bill was painstakingly replicated to provide a labeled training dataset, which served as the foundation for training the OCR and NER components of the proposed system. This step was essential in tailoring the model's capabilities to recognize and process the unique intricacies and variations found in Arabic handwritten bills, ultimately enhancing the system's accuracy and robustness in real-world applications.

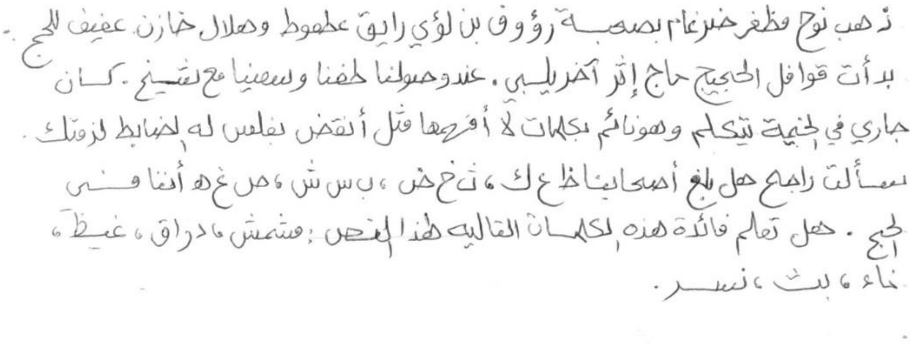

Figure 3: Raw dataset

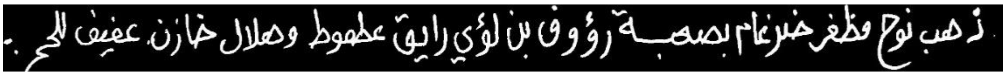

Figure 4: Sample dataset after preprocessing

4. Results and Evaluation

4.1 - Evaluation Overview

The initial dataset of 5,160 images underwent augmentation to increase its size to 15,480. The dataset was then split into a training-testing set, with 80% allocated for training and 20% for testing and validation. A total of 450 epochs were employed, influenced by the choice of optimizer and learning rate. The ADAM optimizer was utilized with an exponential decay scheduler for the learning rate, commencing at a maximum value of 0.0001. A smaller initial learning rate was favored as larger values tended to trigger overfitting prematurely before achieving acceptable accuracy. The training loss reached 1.5730 at the final epoch, while the validation loss was 1.1795.

In Figure 5, we used the Character Error Rate (CER) metric to evaluate the model's performance across the test dataset by comparing training and validation losses. The Character Error Rate (CER) assesses, for a given transcription, the total number of characters (n), including spaces, and determines the minimum number of insertions (i), substitutions (s), and deletions (d) of characters required to match the Ground Truth result. The CER calculation is as follows:

$$CER = [(i + s + d) / n] * 100 \tag{7}$$

Insertions refer to extra characters that the model predicts but are not in the ground truth label (e.g., label = 'cat', insertion = 'caat'). Deletions represent missing characters the model fails to predict, which are present in the ground truth label (e.g., label = 'cat', deletion = 'at'). Substitutions denote characters incorrectly predicted by the model (e.g., label = 'cat', substitution = 'bat'). Evaluation of the model on the testing dataset revealed a CER of approximately 13%, indicating that the model successfully predicted 87% of the characters in all the ground truth labels within the testing set, including punctuation marks and spaces.

Figure 5 is a visualization of the model's predictions versus the image samples for reference:

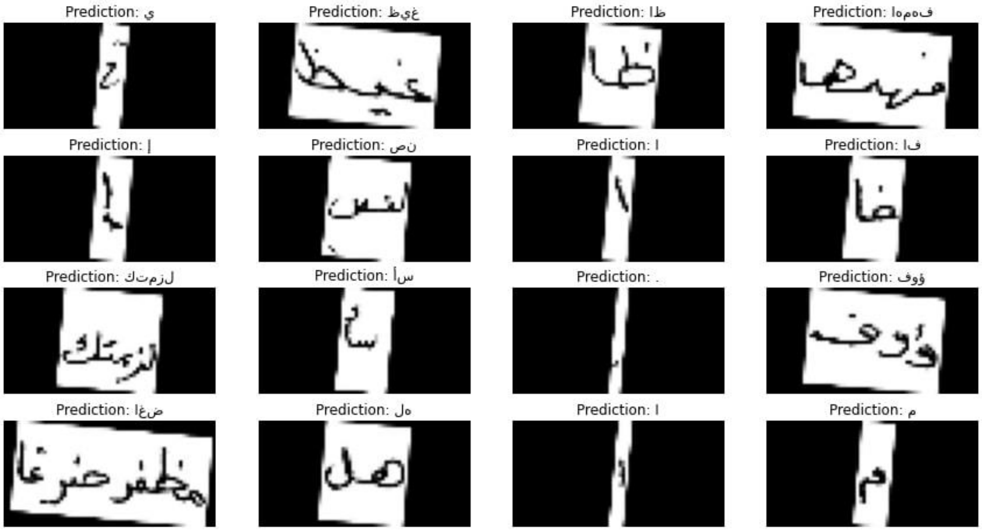

Figure 5: Sample of model results

4.2 Model Summary

According to the model training, we have reached the proposed summary for the whole work inside the model. Table 1 illustrates the summary of the handwriting recognizer model.

Table 1: Model parameters

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\0	(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)	64	[‘dense2[0][0]’]
bidirectional_1 (Bidirectional)	(None, 32, 256)	148480	[‘batch_normalization_3[0][0]’]
label1 (InputLayer)	[(None, None)]	0	[]
dense3 (Dense)	(None, 32, 42)	10794	[‘bidirectional_1[0][0]’]
ctc_loss (CTCLayer)	(None, 32, 42)	0	[‘label[0][0]’, ‘dense3[0][0]]
Total params: 167,994 Trainable params: 167,898 Non-trainable params: 96			

5. Conclusion

Our research addresses the challenges of adapting education to the digital age and streamlining business processes. We introduced innovative teaching methods like Active Note-taking and explored handwritten content recognition. Our work aligns with Egypt's digital transformation goals, providing an automated system for handwritten bill processing. This contributes to business efficiency, cost reduction, and data accuracy. Our deep learning model demonstrates high accuracy in recognizing Arabic characters. In a digital world, our research emphasizes adaptability and technology's role in solving real-world challenges, making handwritten content relevant today.

References

[1] Ye P, Doermann D. Document image quality assessment: A brief survey. In2013 12th International Conference on Document Analysis and Recognition 2013 Aug 25 (pp. 723-727). IEEE.

[2] Dinehart LH. Handwriting in early childhood education: Current research and future implications. Journal of Early Childhood Literacy. 2015 Mar;15(1):97-118.

[3] Gao Y, Yu R, Duan X. An English Handwriting Evaluation Algorithm Based on CNNs. In2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM) 2019 Oct 16 (pp. 18-21). IEEE.

[4] Akouaydi H, Hamdi Y, Boubaker H, Zaied M, Alaya Cheikh F, Alimi A. Children's Online Handwriting Quality Analysis.

[5] Ptucha R, Such FP, Pillai S, Brockler F, Singh V, Hutkowski P. Intelligent character recognition using fully convolutional neural networks. Pattern recognition. 2019 Apr 1;88:604-13.

[6] Katiyar G, Katiyar A, Mehfuz S. Off-line handwritten character recognition system using support vector machine. American Journal of Neural Networks and Applications. 2017 Dec;3(2):22-8.

[7] Ghadekar P, Ingole S, Sonone D. Handwritten digit and letter recognition using hybrid dwt-dct with knn and svm classifier. In2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) 2018 Aug 16 (pp. 1-6). IEEE

[8] Reddy AS, Juliet DS. Transfer learning with ResNet-50 for malaria cell-image classification. In2019 International Conference on Communication and Signal Processing (ICCSP) 2019 Apr 4 (pp. 0945-0949). IEEE.

[9] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770-778).

[10] 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). IEEE.

[11] Bouillon M, Anquetil E. Handwriting analysis with online fuzzy models. In17th Biennal Conference of the International Graphonomics Society (IGS) 2015 Jun 21.

[12] 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.

[13] Simonnet D, Anquetil E. Handwriting Quality Analysis of Block Letters and Cursive Words. Handwriting Today, Journal of the National Handwriting Association. 2016 Dec(15):15-21.

[14] Kulesh V, Schaffer K, Sethi I, Schwartz M. Handwriting quality evaluation. InInternational Conference on Advances in Pattern Recognition 2001 Mar 11 (pp. 157-165). Springer, Berlin, Heidelberg.

[15] Nayef N, Ogier JM. Metric-based no-reference quality assessment of heterogeneous document images. InDocument Recognition and Retrieval XXII 2015 Feb 8 (Vol. 9402, p. 94020L). International Society for Optics and Photonics.

[16] Alaei A. A new document image quality assessment method based on hast derivations. In2019 International Conference on Document Analysis and Recognition (ICDAR) 2019 Sep 20 (pp. 1244-1249). IEEE.

[17] Kang L, Ye P, Li Y, Doermann D. A deep learning approach to document image quality assessment. In2014 IEEE International Conference on Image Processing (ICIP) 2014 Oct 27 (pp. 2570-2574). IEEE.

[18] Peng X, Cao H, Natarajan P. Document image quality assessment using discriminative sparse representation. In2016 12th IAPR Workshop on Document Analysis Systems (DAS) 2016 Apr 11 (pp. 227-232). IEEE.

[19] Lu T, Dooms A. A deep transfer learning approach to document image quality assessment. In2019 International Conference on Document Analysis and Recognition (ICDAR) 2019 Sep 20 (pp. 1372-1377). IEEE.

[20] Alaei A, Conte D, Martineau M, Raveaux R. Blind document image quality prediction based on modification of quality aware clustering method integrating a patch selection strategy. Expert Systems with Applications. 2018 Oct 15;108:183-92.

[21] Li P, Peng L, Cai J, Ding X, Ge S. Attention based RNN model for document image quality assessment. In2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) 2017 Nov 9 (Vol. 1, pp. 819-825). IEEE.