Indonesian Journal of Electrical Engineering and Computer Science
Vol. 35, No. 3, September 2024, pp. 1865~1875
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v35.i3.pp1865-1875

Deep learning-based digitization of Kurdish text handwritten in the e-government system

Shareef Maulod Shareef, Abbas Mohamad Ali
Software and Informatics Engineering, College of Engineering, Salahaddin University-Erbil, Erbil, Iraq

Article Info

Article history:

Received Feb 2, 2024
Revised May 8, 2024
Accepted May 12, 2024

Keywords:

Deep learning
E-government
Feature extraction
Handwritten recognition
Machine learning

ABSTRACT

Many government institutions in developing countries such as the Kurdistan Region of Iraq (KRI) keep a variety of paper-based records that are available in printed or handwritten format. The need for technology that turns handwritten writing into digital text is therefore highly demanded. E-government in developed and developing countries is a crucial facilitator for the provision of such services. This paper aims to develop a deep learning model based on the mask region convolutional neural network (mask-RCNN) to effectively digitize kurdish handwritten text recognition (KHTR). In this research, typical datasets, which includes the isolated handwritten Central Kurdish character images, an extensive database of 40,410 images, and 390 native writers have been produced to determine the developed approach's performance in terms of identification rates. This approach achieves adequate outcomes in terms of training time and accuracy. The proposed model gives higher performance for detection, localization, and recognition when using a dataset containing many challenges, the results were 80%, 96%, and 87.6 for precision, recall, and F-score respectively. The findings revealed that the proposed model obtained better results compared to other similar works. The accuracy of optical character recognition (OCR) is more than 99%.

Corresponding Author:

Shareef Maulod Shareef
Department of software and Informatics Engineering, College of Engineering, Salahaddin University-Erbil
Erbil, Iraq
Email: shareef.shareef@su.edu.krd

1. INTRODUCTION

Many countries across the globe are embracing the digitization process, developing the structure and infrastructure of various areas in order to facilitate digitalization. Hence, transparency and efficiency are increased as a result of digitization. Numerous government agencies in Kurdistan, including government agencies, academic institutes, court systems, registration offices, and some businesses, keep various documents in Kurdish script, which are available in either handwritten or printed form. As a result, there is a high demand for technology that converts handwritten Kurdish writing into digital text. However, the revolution of technological advancements and the need to adapt to enormous changes in huge amounts of data in government institutions presents another problem for the government in dealing with massive amounts of data and efficiently implementing multi-channel platforms for digital transformation [1]. These huge amounts of data are produced on a regular basis and it is estimated that this number will rise significantly in the years to come. It can be seen that big data is a new phrase that indicates datasets that are difficult, if not impossible, to handle using conventional methods or even simple data mining tools. In particular, big data analytics [2], and big data mining [3], may be used to add value to data by extracting usable information from big datasets or streams of data. The need for new technology and applications to open these data and analyse big data it's a necessity.

In recent years, advanced technology has become a fact and created the explosion of big data for both public and private sectors gathering and processing them has increased dramatically, particularly in government institutions. In an attempt to turn the hidden information in this ocean of data into a useful one, the use of advanced technology, such as artificial intelligence (AI) technologies, machine learning, and deep learning have begun making inroads in an ever-broadening range of realms and professions. For example, in business, technology giants like Google, Facebook, and Amazon have been using smart technologies for years. In medicine, systems are expected to work effectively with physicians to provide better healthcare. In education, intended developments in enabling AI to provide challenges to academia in educating and training a new generation of skilled engineers and competent scientists, but the use of AI is rapidly spreading, with global corporate spending on software and platforms expected to reach billions in the coming years. In government institutions, for instance, the developed technology increases the effectiveness of the administration and brings it closer to the citizens. In this regard, the process of identifying handwriting provides a great challenge for machines, as there are many variances in many factors such as tilt, contrast, and stroke variations that vary from person to person.

The process of turning an image with handwritten text that the user inputs into a digital document that they may store on their computer for future use is known as "handwritten text digitization." Because digital papers are more portable, easier to store, and require less space than handwritten scripts or notes. Hence, it can be used to archive and digitize files and written documents particularly in government institutions. We are motivated to create such a model because, despite technical progress, the feeling of writing on paper cannot be replicated on a digital device. Eventhough, the digitization of individual handwritten words it's still challenging, and the word "handwritten text" is so broad, we sought to define it specifically for our purposes in order to focus the amis' scope. Deep learning is an emerging field of research that is focused on the digitization of handwritten Kurdish text. There have been several gaps in previous studies that have revealed important findings, below some of them:

a) In order to train deep learning models, they need large and diverse annotated datasets. Unfortunately, there is a lack of high-quality annotated text files for Kurdish handwritten documents, which can prevent the development of robust and accurate models.
b) The lack of standardized evaluation metrics and benchmarks for the recognition of Kurdish handwritten text files makes it challenging to evaluate deep learning techniques. This issue can affect the field's progress and the reproducibility of results
c) The recognition of Kurdish text files is also challenging due to the presence of various characters and diacritics, which are not typically represented in standard deep learning models. This issue can affect the accuracy of the models when dealing with different styles and cursive handwriting.
d) Due to the varying styles and characteristics of Kurdish handwriting, it is challenging for deep learning models to accurately generalize from a wide range of handwritten samples.
e) Less research is being conducted on the development and use of Kurdish language technology compared to other languages. This lack of attention has resulted in the availability of limited resources and expertise for the digitization of Kurdish handwritten text files.

To address these issues, collaborations among developers, researchers, and language experts are necessary. In addition, efforts toward establishing standardized datasets, enhancing character recognition systems, and promoting open research in the Kurdish language are crucial. This paper aims to develop a deep learning model to automatically digitize e-government services, particularly Kurdish handwritten text more accurately. This paper is comprised of six sections including these opening explanations. Section 2 tackles related works. The role of developed technology such as deep learning in e-government will be addressed in section 3. Section 4 illustrates the methodology that used to achive the aim of the research along with the proposed model for digitizing handwritten Kurdish text. Section 5 illustrate the results and discuss the findings of the research. Finally, the paper provides some conclusions and recommendations in section 6.

2. RELATED WORKS

Humans have been writing their thoughts in the form of letters, transcripts, and other documents for a long time in order to communicate them to others. However, with the introduction of computer technology, the format of handwritten text has swiftly evolved to computer produced digital language, necessitating the creation of a method that can convert handwritten text to digital text, since this makes data processing very quick and straightforward. Ray Kurzweil created the first well-known piece of optical character recognition (OCR) software in 1974, which allowed any typeface to be, recognized. The matrix approach was employed more extensively in pattern matching. Essentially, this would compare the template character's bitmaps to the read character's bitmaps in order to identify which character it most closely resembled. The disadvantage was that this program was sensitive to differences in size and writing styles. The OCR software began employing feature extraction rather than templating to enhance templating. For each letter, the program would look for qualities such as projection histograms, zoning, and geometric moments [4].

In addition, other studies have attempted to construct such a system in the past, and much more study in this area is still required. Many recognition studies have been conducted for offline and online handwritten characters in major languages used around the world, such as English, Chinese, and Indian scripts [5] but they all have some drawbacks, such as accuracy, low conversion speed, poor performance with noisy input, and low a higher false detection rate among others. Because of their vast application potential, recognition studies of handwritten character image samples are still significant. Aside from that, it's impossible to directly apply available English algorithms to other languages. Ghiasi and Safabakhsh [6] provide a set of online and offline Chinese handwriting databases, both of which contain data generated by the same group of writers using the Anoto pen at the same time.

On the other hand, [7] used the normalized and resampled contours of linked segments to produce feature vectors for each manuscript. They then used these feature vectors to create a codebook for writer recognition. They also solve cursive handwriting; the approach uses the occurrence histogram of the shapes in a codebook. Linked complements can be too lengthy and have a large range of forms, and connected complements can be too long and have a wide range of shapes. The authors exploited short chunks of related components and created two successful approaches for extracting code from contours to avoid complicated patterns. One solution uses the contoured pieces' actual pixel coordinates, while the other employs a linear piecewise approximation based on segment angles and lengths to eliminate part of the unneeded data. It aids with the recognition and grouping of related forms. Because this code is shorter, it can be applied more quickly and the codebook can be generated more quickly. In 2013, the authors evaluated their code on two English databases and one Persian database and found that it performed better than other current strategies. Although it is faster to build codebooks, the processing times of this method are lengthy.

Al-Maadeed *et al.* [8] proposed a method for segmenting the input handwritten Devanagari text query image, isolating individual characters, and assessing the resemblance of the individual character in its archive with the isolated characters of the query image within an expected range of resemblance using a convolutional neural network (CNN) on a Devanagari handwritten text recognition (DHTR). It has been discovered that the proposed system performs better. Some researchers [9] demonstrated the identification of diverse writers. They also advised that edge-based directional features be improved by utilizing a filled moving window instead of an edge moving window, and chain code-based features be enhanced by employing a fourth-order chain code list. This strategy was evaluated in the IAM and QUWI handwriting datasets. Shi-Ming and Yi-Song [9] introduced other approach for writer recognition termed digital library system and CNN (DLS CNN), which combined neural network (NN) and line segmentation. Many researches developed another technique named LSTP, but achieved Hamming distance at the classification stage based on NN [10]. Propose a projection profile-based algorithm for dealing with skewed text, overlapping and touched lines. The authors offer two strategies for DHTR in [11] one of which utilizes an artificial neural network (ANN) with pattern recognition (PR) tool and the other of which employs a CNN. Both methods were also compared by the authors. According to this research, CNN had the best outcomes with the PR tool when compared to ANN.

In the related case presented, Mohammed and Ahmed [12] proposed a clustering method for Kurdish handwritten texts, based on the most relevant aspects and utilising the same methodologies as the study of the human characteristics for handwriting. The impact of each feature was investigated using clustering methods and distance gauge behaviours, with each individual feature being examined and a combination of feature sets being tested to determine the performance of features on grouping systems and distance gauges in light of various group numbers. According to the findings, graph-based multiple runs for grouping techniques and distance measurement performed well, with f-measure values for the majority of features being similar across varying groups. However, Zhang *et al.* [13] proposed a method that entailed extracting short writing fragments by placing windows over the writing and grouping similar writing fragments into a codebook. Unlike traditional approaches that result in a codebook of graphemes, the codebook of short writing fragments is script independent and may be used to text in a variety of languages.

3. E-GOVERNMENT AND DEEP LEARNING APPROACH

E-government has traditionally focused on increasing communication and coordination of authorities in various sectors of government, inside the private sector, and even in the public sector. In addition, through simplifying procedures, cutting costs, boosting research capacity, and improving documentation and record-keeping. E-government has been shown to improve the speed and efficiency of operations. However, there is a growing focus that extends beyond the internal workings of e-government and focuses on how procedures and technology may directly engage citizens with more information and promote openness, accountability, and engagement. Mobile and advanced technology such as AI is on the increase all throughout the world, even in developing countries. Artificial Intelligence has existed in many theoretical forms and complex systems for eras. However, only recent improvements in computer power and big data have enabled AI to produce spectacular outcomes in an ever-increasing number of disciplines. For instance, AI has greatly advanced the areas of medical applications, computer vision, reinforcement learning [14], natural language processing [15], and numerous other sectors. AI is described as a computer's ability to emulate human intellect while simultaneously enhancing its own performance. Machine learning is a subset of the AI and is the ability of an algorithm to learn from previous data in order to develop intelligent behaviour and make the right decisions in many conditions it has never seen before. Deep learning, is a sub-set of machine learning, has evolved to overcome the limits of previous machine learning algorithms, unlike regular machine learning techniques. Deep learning is a mapping function that minimizes a loss function using some optimization methodologies to convert raw input data (e.g., a medical picture) to the intended output (e.g., diagnosis) [16].

In this regard, advanced deep learning algorithms can considerably increase the efficiency and cost-effectiveness of present e-government services and systems. Even if deep learning has improved state-of-the-art outcomes in various areas, it is clear that e-government applications confront a number of hurdles in terms of deep learning adaptation [17]. Deep learning applications in e-government services necessitate stringent data security and privacy rules. Despite the fact, there are numerous e-government resources and data that might be used in an increasing number of applications. Data is not being used in a way that facilitates and improves present e-government services utilizing data-driven techniques. However, obstacles such as citizen-government trust, transparency, and other technical difficulties associated with building and implementing safe systems continue to obstruct the development of actual data security and privacy standards. Many countries have recently implemented e-government services in various ministries as well as numerous independent applications [18]. While various studies have been undertaken to improve e-government services, only a few have addressed the use of current developments in deep learning in the automation of e-government services [19]. As a result, using deep learning algorithms to meet e-government difficulties and demands remains critical. While there is still ongoing study in other languages such as Kurdish, the rapid advancements in technology for automated systems oriented to English-like languages are developing and working well in real life.

4. METHOD

The three main components of the proposed model are pre-processing, classification of characterors, and word reconstruction. The research's technique is as follows: first, the handwritten texts are scanned or converted into digital images. Then, pre-processing methods including noise reduction, picture normalisation, and binarization are used to improve the image quality. Then, to identify and transcribe characters or words from the photos, deep learning models, typically mask regional CNN (mask-CNN) are trained using enormous datasets of labelled handwritten text samples. The mask-R-CNN is used in order to effectively digitize Kurdish handwritten text recognition (KHTR). To assess the efficiency of the established approach in terms of identification rates, this study built a large library of 40,410 photos, 390 native writers, and isolated handwritten Central Kurdish character images. With post-processing techniques like error correction algorithms and language model integration, transcription accuracy can be further increased. To ensure accurate and reliable digitization of handwritten text, the model's performance is ultimately assessed using metrics such as precision, recall, and F-score.

4.1. Evaluation parameter

Precision (P), recall (R), and F-measure are all used in the traditional evaluation methodologies for text detection, word spotting, and end-to-end recognition. Precision is the ratio of accurately identified text areas to all text regions that were detected. Recall is defined as the ratio of correctly identified text areas to all text regions in the dataset. Recall and precision were combined to generate the F-measure, a single quality metric [19]. These testing procedures are described as:

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{1}$$

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{2}$$

$$\mathrm{R} = 2 \text{ x } \frac{\mathrm{P \text{ x } R}}{\mathrm{P+R}} \tag{3}$$

4.2. Proposed model for digitizing handwritten Kurdish text

The process of translating handwritten text to digital text automatically will help automate e-government systems significantly. Hand-written text recognition, for example, can improve postal service filtration systems, which presently rely on human personnel to read the address on each envelope and deliver it to the right location. It may also be used to digitize and archive data and programs. In this section, we propose a deep learning model using ResNet 18 as a backbone of mask R-CNN that can detect Kurdish handwritten characters and convert them to digital text to help automate this service. The architecture of the proposed model comprises five major steps. The first step is scanning the document as an image, and the images are passed to the system as input for training purposes. The second step is the pre-processing of these images to remove noise. The third step is a segmentation of the characters using the ResNet 18 model, which will be entered into the model utilized for recognition of the Characters as the fourth step. The final step is the word reconstruction, and then the accuracy is calculated. Figure 1 illustrates the stepwise depiction of the proposed architecture.

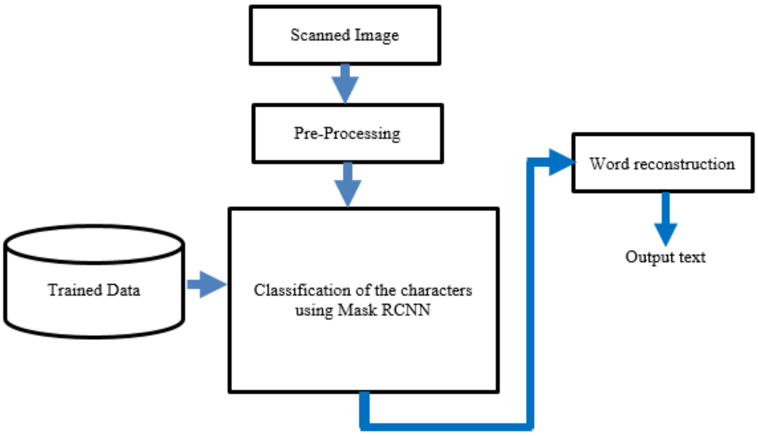

Figure 1. Architecture of the proposed model

4.2.1. Pre-processing

In this work the Kurdish handwritten characters (KHC) [20] will be used. Pre-processing is a process of converting the data row into a suitable form. The main goal is to reduce background noise, enhance the image's region of interest, and create a clear distinction between foreground and background. Noise filtering, binary conversion and smoothing procedures are done on the input image to achieve these aims [21]. A compressed representation of the input image is also part of the pre-processing. In order to determine the region of interest, edge detection is also used. The binary conversion ensures a clear distinction between foreground and backdrop. Edge detection is done during the pre-processing step as well. Figure 2 shows a sample of the characters for the Kurdish language.

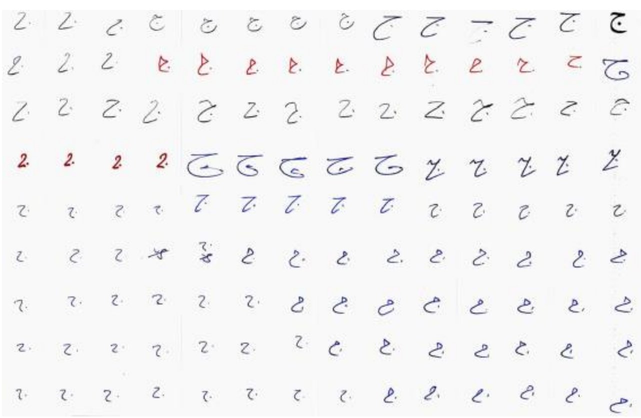

Figure 2. Sample of KHC dataset

4.2.2. Segmentation

The next important step is the technique of obtaining the characters from an image is known as segmentation. Only during the testing step is picture segmentation performed. A whole image is fragmented into a sequence of text/sub-images of a single text. In this method the edge detection and the space between the different characters are used to segment the image. In addition, pre-processing image used to remove noise such as salt and paper noise. Median filtering is one of the commonly used techniques used for this type of noise like small black or white dots on the image. Also, we can use morphological operation to remove and make the edges of the image's smoother. Following segmentation, the sub-divided portions are labelled and processed one at a time. The purpose of this labelling is to determine the total amount of characters in the image. After that, each sub-image is scaled (30×30) and normalized to itself. This aids in the extraction of image quality attributes. With the use of minima or arc sites in between the characters, the scanned picture is recognized for valid segmentation points, which are particularly easy to locate in handwritten text.

4.2.3. Mask R-CNN

In this part, we provide a summary of mask R-CNN, which is used in our method without any changes to its architecture but was retrained using the data from our methodology. A label is given to each pixel of the picture in the well-known instance segmentation network known as mask R-CNN [22]. The feature map is initially extracted from the input picture by the model using a Rest Net 18 as a backbone, and it is then sent to a region proposal network (RPN). The RPN creates several regions of interest (RoI) in the picture based on the feature map of the image. A fully connected network is then given the feature map of each created RoI to forecast its bounding box and semantic class. ResNet is also given the RoI feature map to produce as shown in Figure 3.

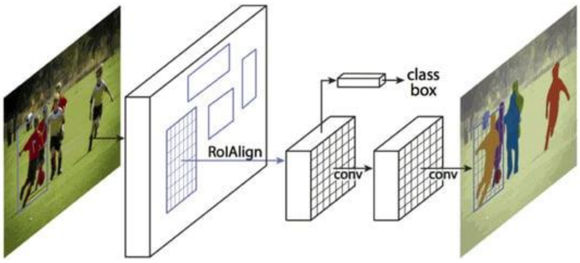

Figure 3. Mask R-CNN frameworks [22]

4.2.4. Classification

In most cases, an input document contains detailed information that is needed for direct categorization, but some of that information is not required for selecting the appropriate class. Therefore, the pattern recognition method relies primarily upon data extraction and measurement. These items are known as features, and feature development is an important aspect of every pattern recognition implementation domain. The characters will be classified into various Kurdish characters that are present in the training dataset, i.e., KHC, during the classification stage. Following the scanning of the words, each written or printed character is compared to a similar character kept within a predetermined class for classification. This stage uses ResNet to classify the Kurdish script into 34 (as the Kurdish language consists of 34 letters) classes. As a result, each character is classified based on the cropped image. A ResNet is a multi-layered neural network with a specialized architecture for detecting complicated data features. It can be used to categorize the contents of a variety of images. The images can be fed into the model as input. ResNet, like ANN, is influenced by the human brain's workings. ResNet can classify images by extracting features, much like the human brain does to recognize items by looking for features. Convolutional and max-pooling layers are included in the ResNet. The completely connected layer is connected to the *n*th pooling layer. To limit the loss, it performs a few backpropagation steps during the learning phase. Finally, it generates the output using an activation function such as Softmax or Tanh [23].

4.2.5. Feature extraction

The feature value is computed by transforming the pre-processed image to a (3×3)-pixel bitmap. Figure 4 depicts a few examples of the proposed model's bitmap versions of certain characters. In a smaller space/data length, the bitmap version keeps the major aspects of the input image. This cuts down on NN training time without compromising the precision of proper character identification. The bitmap pictures are then transformed into a single vector with a size of (15×1), which is used as an input vector to the ResNet.

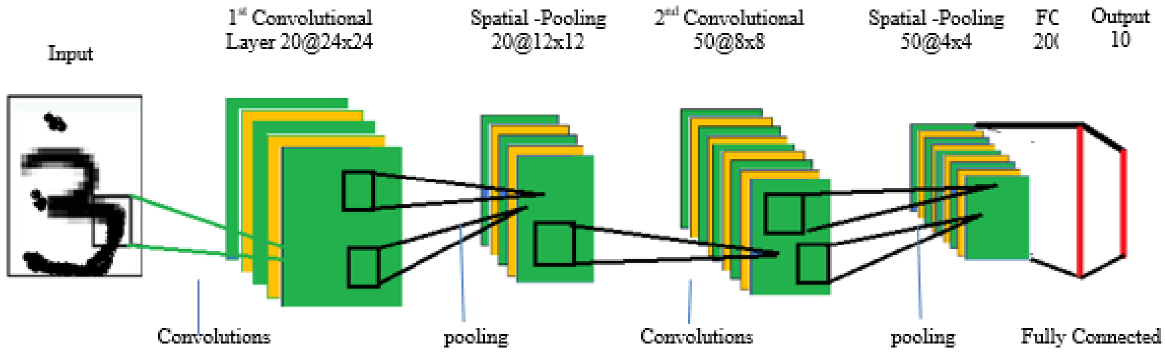

Figure 4. samples of the bit map version for various characters

4.2.6. Text detection and localization by mask RCNN

In the current proposal, we try to detect and localize the texts from images based on a deep-learning model called the mask RCNN model. To achieve the goal of this paper, mask RCNN must be trained on the part of the dataset, and then the training network is tested on the different data and measures the algorithm performance. In the training stage, we used dataset: Kurdish data set (KDS). Mask R-CNN was used to detect the text and localized it. Algorithm 1 shows the training steps. The trained model will be tested to determine the proposed method's performance. Testing the suggested model included many stages which are summarized in Algorithm 2.

Algorithm 1. Training phase
```
Input: KHC data-set.
Output: Text detection and localized
Method:
Step 1: Loading data-set
Step 2: Annotate (label) the images.
Step 3: preprocessing
Step 4: Create the Mask RCNN model as:
Step 5: Train Custom Mask RCNN Detector.
Step 6: detect text and determine its location.
```

Algorithm 2. Testing phase
```
Testing algorithm input: Scene text image.
Testing algorithm output: Real text as an attribute
Method:
Step 1: Enter the scene text image
Step 2: Preprocessing
Step 3: Apply the Mask RCNN model.
Step 4: Detect the text and bound the text region by drawing a box around the text region.
Step 5: Segment the text region.
Step 6: Cropped text from the text region
Step 7: Segment the Character(s) by the following process:
Step7.1: Cut only one character based on the dimensions of the bounding box [ x y w h]
Step7.2: Find the size of character [W H]
Step 8: Determine the font size.
Step 9: Recognize Word(s): achieved by the following process:
Step9.1: Apply Word(s) segmentation
Step9.2: Apply OCR for each word(s).
```

5. RESULTS AND DISCUSSION

The mask R-CNN framework takes into account the various attributes of an input image for the Kurdish character and predicts the likelihood of certain objects fitting into a given character. The initial step is to extract the feature maps from the input image of the Kurdish character using ResNet. After that, the RPN network then passes the generated feature maps into collections of object proposals, which are defined as candidate bounding boxes. A RoI pooling layer aligns the generated feature maps with the proposed objects, and it then filters and refines the proposals. A set of connected layers is then utilized to predict the candidate bounding boxes and the class probabilities, and it also refines the feature maps.

A fully connected layer is utilized to predict the binary masks for each proposal. To eliminate duplicated detections, a method known as non-maximum suppression is used. After this, the final set of candidates is output with the candidate bounding boxes, instance masks, and character probabilities. The proposed model is evaluated using two standard data sets and a new data set, and it is compared to several existing methods. Analysis and discussions regarding our model are also presented in detail. Figures 5 and 6 show the results of our proposed system.

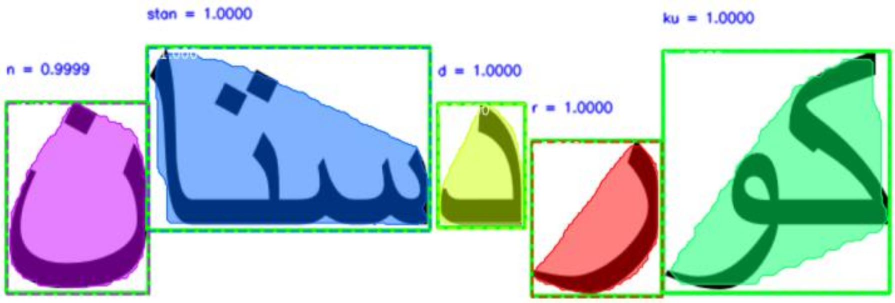

Figure 5. Detection and localization for Kurdish dataset using mask RCNN

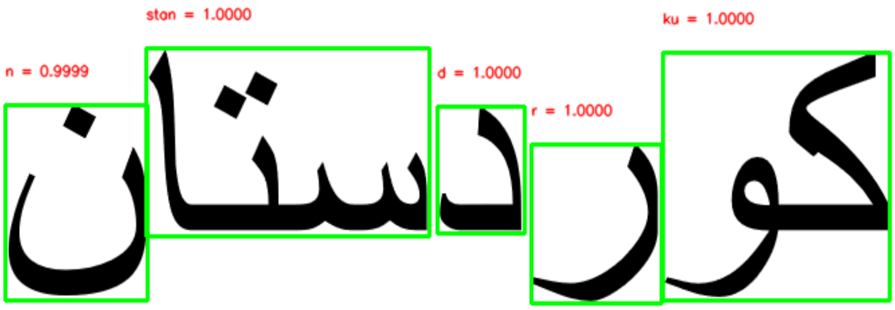

Figure 6. Text recognition using mask RCNN

The terms used in Figure 6 are: in text, recognizer border find the text border and remove the background from the text region (segment the characters and words exactly without background). In character detection, this includes detecting only character and then find font shape and its size. In detected words this mean detecting each word isolate from other words and the result of this step is to recognize each word.

The main problem of recognizing text in a scene is how to detect text inside an image. To do this, many challenges must be taken into consideration to deal with. Four challenges were solved during the creation of the new data-set model: first: colour which means characters with different colours are difficult to deal with; second character size, which means character size changing may lead to unpredictability to recognize text; third: font type: many font types can be used; Fourth: Direction, means not all characters is in the same direction. Figure 7 shows all words that were produced during the new data-set model and detected and localized using mask R-CNN.

The input image size must be 416×416 for all data-set types. For the KCD data set the separation of labelling images 75% training, 10% validation, and 15% testing. The collected images are labelled and separated into two of the mentioned sets, (training, value, and testing) using the labelling program VGG image annotator (VIA). The folder “datasets” is built first, followed by separating it into two subfolders processing, “Train” and “Val,” for keeping training and test samples, respectively. Each one contains photos that correlate to a JSON label information file of the same name. Train the KCD dataset with batch size=15, and the number of epochs=100. The result of mask R-CNN accuracy is 98%. Figure 8 shows the figure for the loss function, where the X-axis indicates the number of epochs. In addition, Figure 9 shows ROC cure, showing the proportion between the true positive and False positive rates with AUC =0.96.

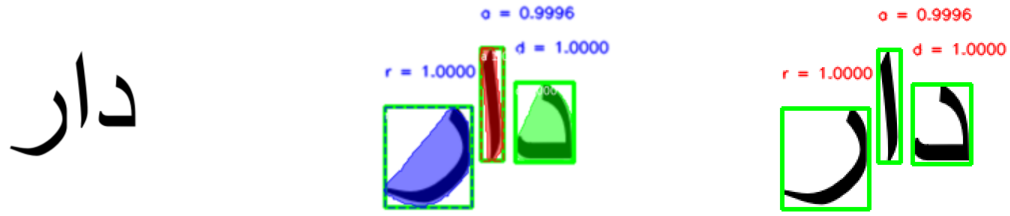

Figure 7. Kurdish character recognition using mask RCNN

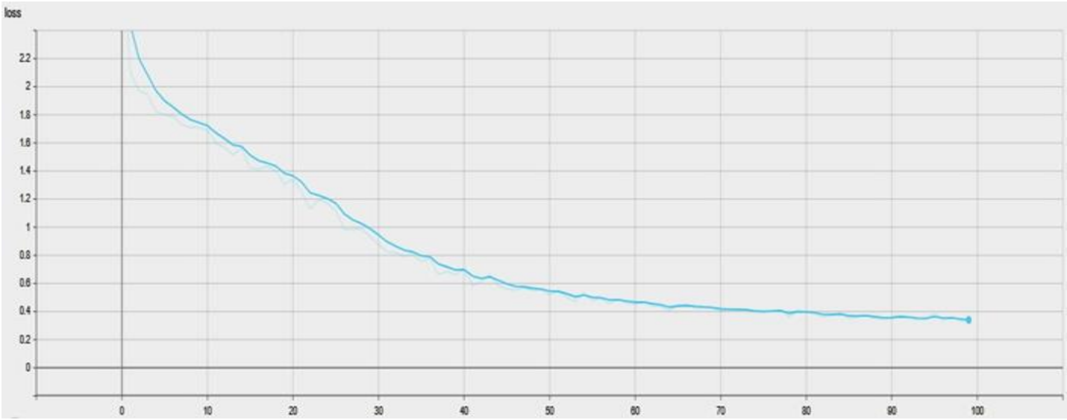

Figure 8. Loss function for the training of the proposed system

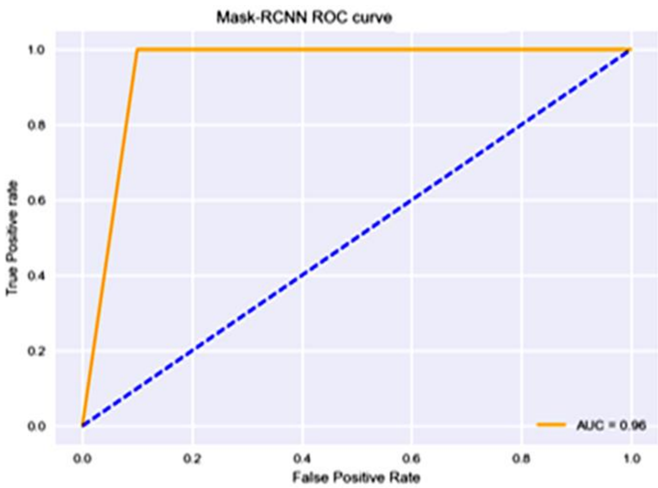

Figure 9. ROC curve for the proposed system

Unlike mask R-CNN, a region proposal network is used. It utilizes a feature extractor and then follows a set of key changes. The third change is the implementation of ROI-pooling, which is a procedure that involves generating output sizes that are suitable for use in a classifier. The ROI-pooling operation is no longer used in mask R-CNN. Instead, it is replaced by an operation called ROI align. Then, it adds a network head to perform the required instance segmentations this lets the Kurdish character be segmented well. The goal of mask R-CNN is to separate the Kurdish character predictions from the mask predictions. In this case, the network head makes its mask predictions. The framework of Mask R-CNN is built on the concept of CNNs. It aims to extract features from the data based on this technique. The use of a feature pyramid network can improve the accuracy of the system and let the characters be recognized. The framework is based on the ResNet-18, which are image analysis network. The main advantage of using a feature pyramid network is its multi-scale and hierarchical properties. A feature pyramid framework requires a network that serves as its backbone. This type of network is usually a convolutional one, and it can be pre-trained. Even though Kurdish character images may have different attributes, such as different shapes for characters, mask R-CNN can perform automated segmentation of them. Its success in character instance classification has made it a viable option for various applications. As many works have been proposed previously for classification, each has a different outcome. Table 1 shows several previous machine learning researches that classify Kurdish characters.

Table 1. Comparison with some other works

No	Works	Accuracy
1	[24]	97.23%
2	[25]	97.27%
3	[26]	96.38%
4	[27]	92%
5	[28]	96%
5	Our proposal	99%

6. CONCLUSION

We believe that this is the first use of mask R-CNN for detection and localization at the same time, unlike the previous methods that detect and localized separately. We detected and recognized text from images with many challenges such as (different font types, sizes, colors, orientations), and different image illumination. Build local challenges text (LCT) dataset is a new data-set model that is very useful because it solves many challenges that are still problems in other methods to solve the text recognition problems. A new way has been proposed to remove most of the background by drawing a box around the text and then segmenting only the text region. MULTI-Level feature extractions using maximally stable external regions (MSER) on the result from mask R-CNN. Character segmentation is a vital step to test whether the text is regular or not. Word isolation based on the word (s) segmentation leads to each word being recognized separately from others. Finally, recognizing text during OCR can be applied to the video, after converting the video into several frames. Thus, we concluded that using advanced technology such as deep learning algorithms will enhance the efficiency and effectiveness of the e-government services.

The results revealed that the proposed model has the potential to produce important advances for the researchers as well as our community. Novel methodology and tools for document analysis, character identification, and language understanding have been made available in research, which has aided in the creation of more precise and effective digitization procedures. This will allow academics to go deeper into historical archives and analyze massive amounts of handwritten data with never-before-seen speed and precision. It also has consequences for several other sectors, including data mining, information retrieval, and historical document preservation. Such technologies provide opportunities for digitizing and preserving priceless manuscripts, records, and literature for communities, especially those whose rich cultural heritage is preserved in handwritten documents. This ensures the materials' accessibility to a larger audience and protects cultural heritage for future generations. In future work, the collection of large-scale datasets will be developed for training mask R'CNN in analyzing the Kurdish handwritten text. This work may be extended to include some deep-learning models for Kurdish handwritten text detection and recognition to increase the accuracy of recognition.

REFERENCES

[1] Z. A. Al-Sai and L. M. Abualigah, "Big data and E-government: A review," in *2017 8th International Conference on Information Technology (ICIT)*, May 2017, pp. 580–587, doi: 10.1109/ICITECH.2017.8080062.

[2] Y. Charalabidis, C. Alexopoulos, E. Ferro, M. Janssen, T. Lampoltshammer, and A. Zuiderwijk, "The world of open data: concepts, methods, tools and experiences," *Public Administration and Information Technology*, pp. 2512–1839, 2018.

[3] N. G. B. Amma, "Big data mining," 2016, pp. 53–59.

[4] A. M. Obaid, H. M. El Bakry, M. A. Eldosuky, and A. I. Shehab, "Handwritten text recognition system based on neural network," *International Journal of Advanced Research in Computer Science and Technology*, vol. 4, no. 1, pp. 72–77, 2016.

[5] F. Zamora-Martínez, V. Frinken, S. España-Boquera, M. J. Castro-Bleda, A. Fischer, and H. Bunke, "Neural network language models for off-line handwriting recognition," *Pattern Recognition*, vol. 47, no. 4, pp. 1642–1652, Apr. 2014, doi: 10.1016/j.patcog.2013.10.020.

[6] G. Ghiasi and R. Safabakhsh, "Offline text-independent writer identification using codebook and efficient code extraction methods," *Image and Vision Computing*, vol. 31, no. 5, pp. 379–391, May 2013, doi: 10.1016/j.imavis.2013.03.002.

[7] S. D. Pande *et al.*, "Digitization of handwritten Devanagari text using CNN transfer learning – A better customer service support," *Neuroscience Informatics*, vol. 2, no. 3, p. 100016, Sep. 2022, doi: 10.1016/j.neuri.2021.100016.

[8] S. Al-Maadeed, A. Hassaine, A. Bouridane, and M. A. Tahir, "Novel geometric features for off-line writer identification," *Pattern Analysis and Applications*, vol. 19, no. 3, pp. 699–708, Aug. 2016, doi: 10.1007/s10044-014-0438-y.

[9] C. Shi-Ming and W. Yi-Song, "A robust off-line writer identification method," *Renhe Test*, vol. 46, no. 1, pp. 108–116, 2020.
[10] D. S. Prashanth, R. V. K. Mehta, and N. Sharma, "Classification of handwritten devanagari number – an analysis of pattern recognition tool using neural network and CNN," *Procedia Computer Science*, vol. 167, pp. 2445–2457, 2020, doi: 10.1016/j.procs.2020.03.297.
[11] A. A. Ahmed, "Kurdish handwritten classification based on human personality analysis," *Academic Journal of Nawroz University*, vol. 10, no. 1, p. 174, Mar. 2021, doi: 10.25007/ajnu.v10n1a862.
[12] T. L. Mohammed and A. A. Ahmed, "Offline writer recognition for kurdish handwritten text document based on proposed codebook," *UHD Journal of Science and Technology*, vol. 5, no. 1, pp. 21–27, 2021.
[13] Y. D. Zhang, Y. Zhang, X. X. Hou, H. Chen, and S. H. Wang, "Sevenlayer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed," *Multimedia Tools and Applications*, vol. 77, pp. 10521–10538, 2018.
[14] S. Venugopalan, H. Xu, J. Donahue, M. Rohrbach, R. Mooney, and K. Saenko, "Translating videos to natural language using deep recurrent neural networks," *arXiv preprint arXiv:1412.4729*, 2014. https://arxiv.org/abs/1412.4729.
[15] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*, Heidelberg: Physica-Verlag HD, 2010, pp. 177–186.
[16] A. Kankanhalli, Y. Charalabidis, and S. Mellouli, "IoT and AI for smart government: a research agenda," *Government Information Quarterly*, vol. 36, no. 2, pp. 304–309, 2019.
[17] A. T. Ho, "Reinventing local governments and the e-government initiative," *Public Administration Review*, vol. 62, no. 4, pp. 434–444, 2002.
[18] D. A. D. Putra *et al.*, "Tactical steps for e-government development," *International Journal of pure and applied mathematics*, vol. 119, no. 15, pp. 2251–2258, 2018.
[19] R. M. Ahmed, T. A. Rashid, P. Fatah, A. Alsadoon, and S. Mirjalili, "An extensive dataset of handwritten central Kurdish isolated characters," *Data in Brief*, vol. 39, p. 107479, Dec. 2021, doi: 10.1016/j.dib.2021.107479.
[20] R. Watkins, N. Pears, and S. Manandhar, "Vehicle classification using resnets, localisation and spatially-weighted pooling," *arXiv preprint arXiv:1810.10329*, 2018.
[21] Y. Jung, D. Kim, B. Son, and J. Kim, "An eye detection method robust to eyeglasses for mobile iris recognition," *Expert Systems with Applications*, vol. 67, pp. 178–188, Jan. 2017, doi: 10.1016/j.eswa.2016.09.036.
[22] M. A. Butt *et al.*, "Convolutional neural network based vehicle classification in adverse illuminous conditions for intelligent transportation systems," *Complexity*, vol. 2021, pp. 1–11, Feb. 2021, doi: 10.1155/2021/6644861.
[23] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 2980–2988, doi: 10.1109/ICCV.2017.322.
[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
[25] G. A. Abandah, M. Z. Khedher, M. R. Abdel-Majeed, H. M. Mansour, S. F. Hulliel, and L. M. Bisharat, "Classifying and diacritizing Arabic poems using deep recurrent neural networks," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3775–3788, Jun. 2022, doi: 10.1016/j.jksuci.2020.12.002.
[26] W. A. Yousef, O. M. Ibrahime, T. M. Madbouly, and M. A. Mahmoud, "Learning meters of Arabic and English poems with Recurrent Neural Networks: a step forward for language understanding and synthesis," *arXiv preprint arXiv:1905.05700*, 2019.
[27] S. Yousefi, "Rhythm recognition of persian poems via machine learning," *University of Tehran, Tehran*, 2019.
[28] R. M. Ahmed *et al.*, "Kurdish handwritten character recognition using deep learning techniques," *Gene Expression Patterns*, vol. 46, p. 119278, Dec. 2022, doi: 10.1016/j.gep.2022.119278.

BIOGRAPHIES OF AUTHORS

Dr. Shareef Maulod Shareef holds a Ph.D. in Computer Science in the School of Architecture, Computing, and Engineering at the University of East London-UK. He worked as an hourly paid lecturer at the University of East London-UK. In addition, he has published various papers along with participating in reviewing various conference papers, and international journals. He is also a member of many professional bodies in his area of specialization. He became a Fellow of the Higher Education Academy (FHEA) in the UK in 2011. He worked as the head of the Software and Informatics Engineering Department from 2012-2017 in the College of Engineering at Salahaddin University-Erbil-Iraq. Furthermore, Shareef worked as a Dean of the College of Education and Languages at Lebanese French University-Erbil-Kurdistan-Iraq, along with the Director of Postgraduate Studies from 2017-2021. He supervised many M.Sc. and Ph.D. students in his area of specialization. Currently, he is working as a Vice president for scientific affairs and Postgraduate studies at Slahaddin University-Erbil in Iraq. He can be contacted at email: shareef.shareef@su.edu.krd.

Dr. Abbas Mohamad Ali holds a Ph.D. in Computer vision from UKM - Malaysia, currently, He works as a part-time lecturer at many universities in KRG. In addition, he has published various papers along with participated in reviewing various conference papers, and international journals. He is also a member of many professional bodies in his area of specialization. He worked as the head of the Software and Informatics Engineering Department from 2022 to 2024 in the College of Engineering at Salahaddin University. Furthermore, he supervised many M.Sc. and Ph.D. students in his area of specialization. He is working as a staff member of Software and Informatics engineering at Salahaddin University-Erbil-Kurdistan-Iraq. He can be contacted at email: abbas.mohamad@su.edu.krd.