ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/332682920

# Deep Learning Algorithms for Arabic Handwriting Recognition: A Review

**Article** *in* International Journal of Engineering & Technology · September 2018

DOI: 10.14419/ijet.v7i3.20.19271

CITATIONS
7

READS
1,843

**5 authors**, including:

Ahmed Alsaffar
Universiti Kebangsaan Malaysia
22 PUBLICATIONS 230 CITATIONS
SEE PROFILE

Suryanti Awang
Universiti Malaysia Pahang
36 PUBLICATIONS 171 CITATIONS
SEE PROFILE

Sabrina Tiun
Universiti Kebangsaan Malaysia
80 PUBLICATIONS 473 CITATIONS
SEE PROFILE

Ahmed Salih Al-Khaleefa
Universiti Kebangsaan Malaysia
23 PUBLICATIONS 133 CITATIONS
SEE PROFILE

**Some of the authors of this publication are also working on these related projects:**

Project Image classification using texture descriptors View project

Project Natural Language Processing View project

All content following this page was uploaded by Ahmed Alsaffar on 25 January 2020.

The user has requested enhancement of the downloaded file.

*International Journal of Engineering & Technology,* *7 (3.20) (2018) 344-353*

**International Journal of Engineering & Technology**

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***

# Deep Learning Algorithms for Arabic Handwriting Recognition: A Review

**[1] Ahmed AL-Saffar , [1]Suryanti Awang , [2]Wafaa AL-Saiagh , [2]Sabrina Tiun and [3]A. S. Al-khaleefa**

*[1] Faculty of Computer System and Software Engineering,*
*University Malaysia Pahang UMP, Pahang, Malaysia*
*[2]Faculty of Information Science and Technology,*
*Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia*
*[3]Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia*
*Coresponding Arthur Email:* ahmed_saffar5@siswa.ukm.edu.my

**Abstract**

Computer vision (CV) refers to the study of the computer simulation of human visual science. Major task of CV is to collect images (or video) so that they could be used for analysis, gathering information, and making decisions or judgements. CV has greatly progressed and developed in the past few decades. In recent years, deep learning (DL) approaches have won several contests in pattern recognition and machine learning. (DL) dramatically improved the state-of-the-art in visual object recognition, object detection, handwritten recognition and many other domains. Handwritten recognition technique is one of this tasks that targeted to extract the text from documents or another images type. In contrast to the English domain, there are a limited works on the Arabic language that achieved satisfactory results, Due to the Arabic language cursive nature that induces many technical difficulties. This paper highlighted the pre-processing and binarization methods that have been used in the literature along with proposed numerous directions for developing. We review the various current deep learning approaches and tools used for Arabic handwritten recognition (AHWR), identified challenges along this line of this research, and gives several recommendations including a framework based (DL) that is particularly applicable for dealing with cursive nature languages.

**Keywords—***Arabic OCR; Deep Convolutional Neural Networks; pattern recognition; image processing, Text recognition.*

## 1. Introduction

This Image recognition is considered a major research direction in the computer vision field. It has a very vital role and impact for the acquisition and processing of image-based intelligent data. Using image recognition technology can help effectively handle the identification and detection of specific target objects (e.g. face, objects, or handwritten characters), classification of images, and subjective assessment of image quality, among other issues. Currently, the technology of image recognition has great commercial market and a good prospect for applicability in Internet applications such as commodity recommendation, image search, face recognition, and user behavior analysis. Simultaneously, advanced technology such as intelligent robots, unmanned aerial vehicle industry and biology, unmanned driving, geology, medicine and several other disciplines have broad prospects for applicability. In the early systems for image recognition, feature extraction methods like histogram of oriented gradients [5] and scale-invariant feature transform [7] were used. Furthermore, the extracted feature input to a classifier was used for recognition and classification. Essentially, these features are considered a manual design feature. For various identification problems, the extracted features directly affect the performance of the system. Therefore, the researchers need to determine the problem areas that need to be examined in order to design adaptability and give it better features, which would enhance system performance as well. Generally, this part of the image recognition system is meant

**Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**

for a specific identification task. Furthermore, data size is not large and the ability for generalization is poor. Thus, in the practical application of the problem, it is difficult to reach an accurate identification effect.

Over four hundred million people worldwide speak the Arabic language. This language is also used as a method for the transcription other languages such as Turkish, Persian, Kurdish, Urdu, and Malay. Unlike the Chinese and Latin domain, off-line techniques for Arabic Handwritten Recognition (AHWR) are not well-developed yet, because the cursive nature of the language gives rise to numerous technical difficulties (Khémiri, Kacem, & Belaïd, 2014; [2]).

HMM-based printed Arabic text recognition is presented by [10] for various scenarios. This is a procedure for partitioning the sliding window adaptively into cells. The procedure makes use of the writing line attribute of ink-pixel distributions and Arabic text. On the other hand, unseen-font and mixed-font text recognition was recommended with a two-step methodology in which the input text line image is linked with the nearest known font in the first step. In the second step, the HMM-based text recognition is applied using the recognizer that was trained on the linked font's text. [11] suggested a Hybrid Classification Scheme to enhance the performance of traditional HMM classifiers. The procedure is composed of hybrid k-NN classifiers and PAW-based HMMs. To derive the accuracy results, the IFN/ENIT Arabic Word Database is applied. [12] recommended a technique that makes use of three types of HMMs: two for the slanted windows and one for the vertical sliding window. In order to combine the three HMMs, a fusion scheme is used. Every classifier examines the image from a certain orientation. When these classifiers were combined, the recognition rate significantly increased. An online handwritten mathematical expression recognition system was suggested by [13]. Accordingly, a 2D-SCFG based standard statistical framework and its consequent CYK-based parsing algorithm were defined. As proposed by [14], there can be a two-recognizer system, based on multi-stream HMM, for the offline recognition of an Arabic handwritten Tunisian city name. The two key recognizers are deemed as a dynamic Bayesian network that primarily emphasizes on combining the extracted attributes at each slice time t simultaneously. Although the second methodology models the two observations separately, with the help of an HMM for each stream, the interaction takes place through another hidden layer. However, enhancing the hidden Markov model can improve the recognizer performance considerably and make it more efficient. Only a consistent estimate of the attributes or parameters can be applied to analyze the sequence and decode it efficiently for even the most complex of problems.

This paper aims to accommodate a review of (AHWR) task by exploring the methods that have been used in the previous works. In addition, we proposed a general methodology based on deep learning technique for (AHWR). We will highlight the pre-processing and binarization methods that have been used in the literature, and we will propose some directions for its development. Next section, explain the deep learning algorithms and applications in a deferent domain, followed by a review on its application on (AHWR) recent works.

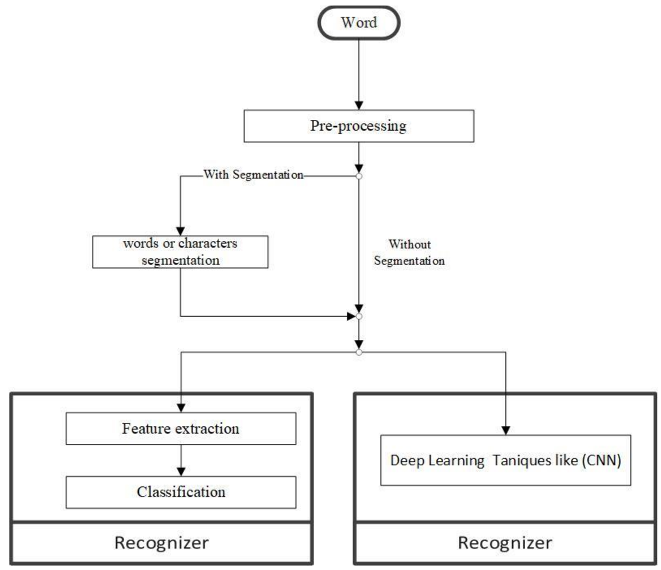



Figure 1 General Steps for Arabic Handwritten Word Recognition

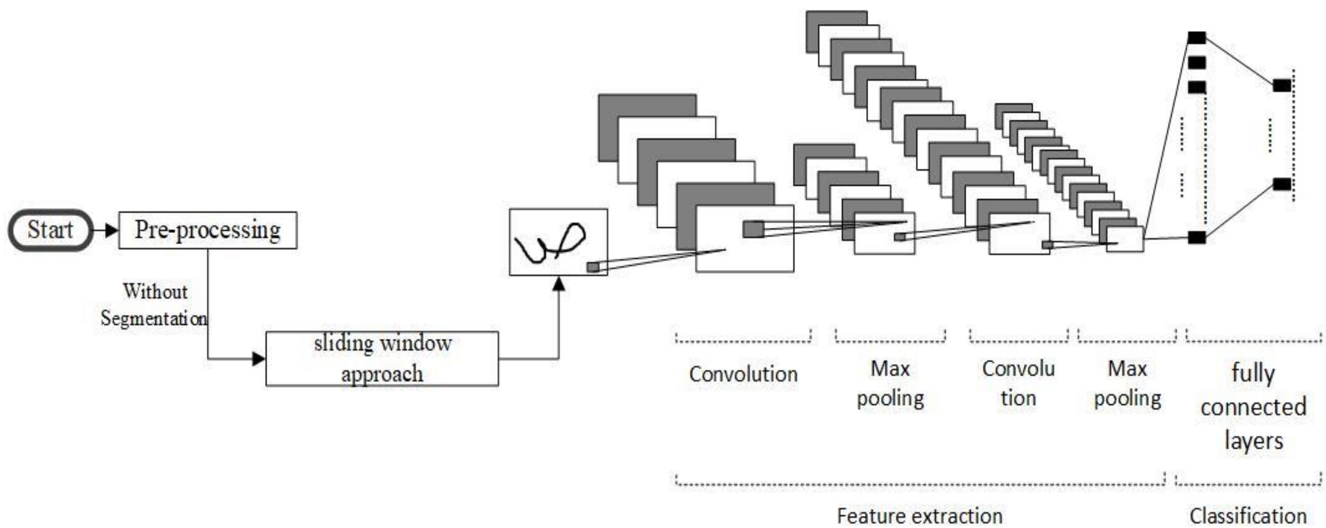
## 2. Handwritten Recognition Architectural

In Arabic handwritten recognition models the input will be a text image and output are the text that recognized. Pre-processing is the first important step that start with any Arabic handwritten model. Due to the high frequency of noise in the input text image that usually come from a scanner, a camera, or simply a print-screen of a page. The noise can be generated from the degradation of old manuscripts or additional non-text contents such as pictures, equations, tables. Many of pre-processing related works will be presented in next section of, Arabic OCR Pre-Processing and Layout Analysis.

According to [15] in 2015, the process of recognizing Arabic words can be perform without segmentation as whole word, or recognized words based on its segmented characters as mentioned in fig 1. Whilst the Arabic sim-cursive nature makes the segmentation step like an errors bank for Arabic handwritten recognition models. A recent study by [14] investigated the effects of two words models are implemented based on the holistic and analytical approaches without any explicit segmentation. Jayech, [14] used a sliding window approach over the columns and rows to

extract a combination of statistical and structural features that extracted.

Figure 2 flowchart of Arabic Handwritten recognition based on CNN model




In the recent years can merely note, the state of art researches focused on implemented the deep learning techniques for Arabic Handwritten recognition [6, 8, 9, 16]. they can fallow a new hypothesis that consists of only two phases to solve the Arabic OCR task. Whilst, traditional approaches applied the three phases, firstly; pre-processing, word segmentation phase, followed by machine learning algorithms as classification phase. Nevertheless, the high capability of deep learning technique specially (CNN) to capture word features promote to recognize the whole word directly without performing word segmentation task. Arabic Handwritten recognition-based DL approach flowchart, can be includes only pre-processing phase followed by the (CNN) technique as mentioned on fig.2

## 3. Pre-Processing Arabic (Hwr)

Optical Character Recognition (OCR) refers to the recognition of written or printed text characters using a computer. However, despite years of studies and the existence of numerous off the shelf OCR systems, OCR processes often produce outputs that still contain errors. In the past several years, the focus of researchers has been on the recognition of degraded historical handwritten document images. Bleed-through of ink, uneven intensities, dirt, marginal noise, uniform illustrations, and other kinds of damage are often present in old documents as well. Standard OCR packages do not have the capacity to accommodate this range of document recognition needs and diversity of content Rani and Lehal [17].

The pre-processing stage is a vital OCR stage Naz, Hayat [18]. Choosing the best image improvement method is an important part of the OCR task. For instance, an individual method has a greater tendency to harm OCR results instead of improving them Borovikov and Zavorin [19]. Several kinds of pre-processing techniques are needed before the goal of text recognition can be targeted. Handling degradation documents for transcription using optical character recognition (OCR) is seen as more difficult compared to processing a normal document. One of the major and most important steps includes automating the noise removal process while preserving the content of the text.

Shafait and Breuel [20] proposed an algorithm which is a combination of connected component removal and projection profile analysis to determine the noise regions' borders on the printed binary document image. . Stamatopoulos, Gatos [21] proposed a technique to help in the detection of the optimal page frames. This method is based on the horizontal and vertical white run projections of double page binary documents such as newspapers and books. Chen, Wei [22] devised a page segmentation method that can be used for the colour historical handwritten document images. This proposed technique solved the page segmentation task by combining colour, coordinates, and texture features, followed by the classifier algorithm and the feature selection process. Chakraborty and Blumenstein [23] suggested a text preserving method that also eliminates the non-textual noise from its environment. This is directly applicable to greyscale images. The basis of this method is the analysis of the area between the text area and the marginal noise. However, using only a single method to handle various degradations documents images is difficult Farahmand, Sarrafzadeh [24]. The high inter / intra variation between the document background and the text stroke is still considered as one of the major obstacles as it makes the task of handwriting recognition more difficult.

Arabic Handwriting recognition (AHW) had two main steps to start recognition phase, pre-processing and segmentation. Some works in literature considered the segmentation phase as a part of pre-processing as well. Many different schemes are suggested for extracting text blocks from machine-printed Arabic documented images. Segmentation of text to line blocks in AHW usually used with the help of horizontal projections. Meanwhile, methods of braking this words to characters, still need more developing [25]. In the case of multiple-column documents, in one study, the researchers, ,[26] had proposed the top-down approach which used a combination of the structural segmentation along with the functional labelling. The segmentation approach is developed using the X–Y cut algorithm and is directed depending on the knowledge regarding the layout features of the document. This helped in constructing a structural documented tree after recursively studying the horizontal and the vertical projection profiles of the partitioned areas. Thereafter, these vertical and horizontal projection profiles were used for dividing the page into small blocks like columns and also paragraphs [27, 28].

In contrast to the approach used by Nagy et .al 1992 , another author [29] used the method of the structural page layout which was further based on the bottom-top, nearest neighbor clustering of the various page components, and the author grouped the smaller components (starting from the pixels which were used as connected components) into larger components till all the blocks were seen on the document page. Thereafter, they subdivided the text blocks into singular text lines using horizontal projection. Using these techniques, the authors addressed the major problem of tackling the local nonlinearity seen for the text lines. For the past several years, various researchers have proposed numerous layout analytical algorithms for handling different layouts and many of these were robust against the noise in the text. A large proportion of these proposed algorithms has been used for the analyzing the text images in many scripts. In the study conducted by [30], the authors developed the smearing approach based on the nonlinear, run-length smoothing algorithm for task segmentation and the classification of the various digital printed texts documents into smaller areas containing images and text. Using the general text line features, a linear adaptive classification approach is able to distinguish the text areas from other regions. Furthermore, [31] proposed a novel method for analyzing the whitespace in the document and also addressed the problem related to the isolation of the blocks containing the machine-printed text. Their method was generally based on the computational-geometrical algorithms for the off-line enumeration of the maximum number of white rectangles along with the on-line rectangle unification. They used the complicated layouts from many technical journals in English as their main data. In 2002, [32] applied the geometric algorithms in order to solve some key problems noted in the layout analysis; which involved finding a background whitespace cover in the document like maximum empty rectangles and also determining a constrained maximal likelihood match of the geometric text line models observed when geometric obstacles were present. The author also proposed the evaluation function which could accurately identify a maximum number of empty rectangles which corresponded to the column boundaries. Using the evaluation function in combination with the geometric functions helped in an easier implementation of the layout analytical system. In another study by [33], the author proposed a novel scheme involving page segmentation which was based on an estimated region of the Voronoi diagram. Their method had the following characteristics:

(1) This Voronoi diagram helped in obtaining the various boundary candidates of the document components using the page images with the help of the skew and a non-Manhattan layout.

(2) The authors used the candidate for estimating the interline gaps and the inter-characters without using the domain-specific parameters for boundary selection.

In contrast to the Arabic OCR, there have been very few reports which have analyzed the Urdu or Arabic and the Jawi documents. In their study, [34] applied the morphological operators for extracting the text lines from the printed Jawi or Arabic text documents. Furthermore, they also applied these on the font with 2 sizes for verifying the approach used by the morphological operators.

In 2012, [35] proposed the layout analytical system for the extraction of text lines in a reading order from various scanned images in the Arabic scripted documents written in many other languages (Urdu, Arabic, Jawi and Arabic etc.) and many styles (i.e., Naskh or Nastaliq, etc.). They also showed that this improved multiresolution morphology-based text or non-text segmentation method was suited for analyzing Arabic texts.

Also, it is seen that a greyscale image is first converted to a binary form with the help of the suitable binarization approach. Many reports have been published in the literature which have applied the binarization approaches [36-39]. In the study by Otsu [36], the author proposed a nonparametric and an unsupervised technique for selecting automatic thresholds for image segmentation. In their technique, they used the $0^{th}$ and the $1^{st}$ order cumulative moments present in the grey-level histograms. This was a simple technique which could be extended for multi-threshold problems. Also, [37] proposed a binarization method for solving the problems related to adaptive document imaging, wherein a page was compiled of various components like text, background and images. They applied 2 algorithms for determining a local threshold for every pixel. These algorithms were studied using various images which had various text components and degradations.

Furthermore, in one study, the authors [40] highlighted the Urdu-like script peculiarities, by presenting the text and image databases. They compiled their attempts into 3 parts, i.e., printed, handwritten, along with the online character recognition. For every part, they analyzed the different studies using a traditional OCR pipeline and laid more emphasis on pre-processing, line segmentation, feature extraction, data classification, and recognition.

Also, Saberi [41] analyzed the legibility of various Arabic scripts for the Malaysian mosque users and proposed a novel quantitative tool called the ANFIS process. They identified the most significant parameters which affected the understanding and the readability of the different decorative and cursive Arabic scripts. Furthermore, they also studied the manner in which the demographic aspects and the cognitive skills affected the capability to accurately interpret the Arabic scripts.

Today, several studies applying the Optical Character Recognition (OCR) make use of the off-line images as the dataset in the various OCR models. All these studied images from a type of a historical manuscript which are transferred to their digitized counterparts after photographing them. However, there are some issues related to the performance of an OCR model we can mention as First, Photograph Resolution the resolutions can be very low (some cases showed 200 dpi). Additional, Binarization artefacts: these could occur due to the bad quality of the various built-in binarization algorithms which were installed in many inexpensive scanners. Finally, Noise: this occurred in different forms like lines or rule lines, holes and blobs, dark background, logos etc.

Sauvola and Pietikäinen [37] introduced a local threshold binarization method that can be used for the adaptive document image. In this approach, the page is taken as a collection of subcomponents like background, text, and picture. It utilized two algorithms so that a local threshold can be identified for each pixel. Testing of the algorithms was done with images, which include various kinds of degradations and document components. Su, Lu [42] introduced a historical document image binarization approach that can tolerate various kinds of document degradation, such as document smear and uneven illumination. This suggested technique utilizes the image contrast that is assessed based on the local maximum and minimum (LMM). Su, Lu [43] designed an adaptive contrast technique that can be especially used for degraded document images. Furthermore, the adaptive image contrast refers to a mixture of the local image contrast and the local image gradient that can tolerate variations in the text and background as a result of various kinds of document degradations. Chen and Wang [44] introduced a new framework that is useful for binarization of the degraded document image and the restoration of the document images' quality.

However, the method being proposed is considered a variation of Wellner's adaptive thresholding that has its basis on the histogram. Furthermore, the method being proposed passes through the image from the left and from the right alternatively. It is able to locate the text regions if the pre-processed image's value goes beyond the threshold d. The Rosenfeld's method is used to determine the threshold d. However, this method has difficulties achieving a high accuracy rate. Furthermore, the background noise resulting from the process causes severe degradation issues in the binary algorithm.

The targeted image enhancing algorithms for some noise-related categories have been described earlier. The problem of surrounding noise was addressed by Agrawal Agrawal and Doermann [45] where they treated the binary images with different noise types which were sized like the diacritics (a secondary component) and nearer the text constituent named, "Stroke-like Pattern Noise (SPN)". Depending on the text characteristics, the image components were categorized as the Prominent Text Components (PTC) when they could be identified using no neighborhood context, and as non-Prominent Text Components (non-PTC) when they needed the neighborhood context. Once the image constituents were classified as PTC or others (like non-PTC components and noise etc.), the set containing textual features was extracted and used by a supervised algorithm which used an SVM and a Radial Basis Function (RBF) kernel. Thereafter, in the next stage, the results were used for filtering and removing the various non-PTC constituents. For this purpose, the stroke width and the cohesion property (based on the distance transform for backgrounds) were computed for every constituent using a K-means clustering approach. Finally, the verification was carried out for removing the misclassified text components. In another study by Agrawal and Doermann [46], the authors also proposed a novel algorithm for detecting and reducing the clutter noise seen in the texts associated with the binary document images. They assumed that this clutter noise was higher than 2-times the text maximal stroke width, and hence, the distance transform was applied for detecting and reducing the noise using an iterative manner. In Stage 1, the pixels in the foreground were thinned to 50% of the maximum distance, which yielded an image containing only the cluttering constituents. Thereafter, the authors extracted a set of features from the image and used a Support Vector Machine (SVM) classifier for determining if the extracted image contained clutter components. In Stage 2, the authors removed the associated clutter noise and maintained the text components. The major idea of using this technique was to regenerate the clutter constituents from the half-residual core noted in the earlier stage. Therefore, the authors computed the distance transform for every component associated with the clutter noise, from the clutter boundary. The steps needed (pixels having similar distances) for regeneration were seen to rapidly change for the component edges. The distance (i.e., number of steps required) was used for removing all pixels in the image, thereby removing all the related clutter pixels.

Some studies have also reported the use of other enhancement algorithm filters like the modified median filters. Toh [47] proposed an NAFSM filter for effectively removing the salt-pepper noises. This NAFSM filter was a type of recursive double-staged filter, which detected the intensities of the salt-pepper noises and then located the positions of these probable noise pixels. After detecting the "noise pixel", it was subjected to a subsequent filtering stage. However, if the pixel was categorised as "noise-free," then, it would be maintained, and no filtering action would be carried out, thus preventing any kind of alterations to the finer image textures or details. Also, in one study Chinnasarn, Rangsanseri [48]; the researcher presented a novel technique which was based on a kFill algorithm, which was carried out by scanning the image using a single-pass. This algorithm could remove the salt-pepper noises in the image, which were sized smaller that the document objects. Many studies have been presented in the literature, which investigate the deferent filters. Some popular filters used for handling the salt-pepper noises were the Optimal Boolean and the Modified Directional Morphological filters (MDMFs). Lee Lee and Fan [49] used an Optimal Boolean filter for improving the binary document images which contained many uniform noises or image backgrounds with uniformly-distributed specific graphical patterns. Ping Ping and Lihui [50] proposed a noise eradication model for removing the noises from binary document images. Their model was based on 2 filtering algorithms, i.e., MDMFs and the Image Geometric Structure Filters (IGSF). However, if the image showed several types of noises, these techniques were unable to resolve the issues.

## 4. Binarization

The process of binarization involves the conversion of the greyscale or colour images to the binary images. Many techniques were proposed earlier for binarization, which are divided into two types of approaches: i.e., global and local thresholding. Carrying our binarization is difficult for the images if they have a lower contrast, significant noise levels, variable background reflectance, and complex textual patterns. Hence, no single binarization process can provide completely satisfactory results for different types of images. Therefore, numerous local adaptive methods have been proposed for preserving the textual information as accurately as possible. Many experiments were carried out for investigating and examining the available local adaptive binarization methods like [51-56].

Another problem arises when many historical documents are converted into images, they need to develop uneven backgrounds method. The text in the foreground must be separated from the backgrounds for better readability. Hence, the thresholding that occurs along a scanline must be in the form of a curve instead of a straight line/ lines which were determined using conventional global or local thresholding algorithms [37].

Earlier studies have described kinds of threshold algorithms which were used for solving the OCR problems. Global thresholding techniques was used for images include Otsu's thresholding technique [36]. Which is an effective and a common method, followed by Entropy technique [57] and the minimum error method by Kittler and Illingworth [58].

However, for images that possess a higher contrast background, it was observed that the Integral Ratio technique performed better than the Entropy algorithm and the Otsu's technique. [42] formulated an adaptive contrast technique that can be especially implemented for degraded document images. Furthermore, the adaptive image contrast refers to a mixture of the contrast of the local image as well as the local image gradient that can tolerate the background and text variations that come from various kinds of document degradations. First, an adaptive contrast map is built so that it could be used for an input-degraded document image. Afterwards, the contrast map is binarised and mixed with Canny's edge map so that the text stroke edge pixels can be identified. A local threshold is then used to further segment the document text. Estimation of this local threshold is based on the intensities of the detected text stroke edge pixels that are found within a local window. (Chauhan, Sharma, & Doegar, 2016) have stated that adaptive contrast technique performs better than other popular binarisation methods, such as Niblack's and Sauvola's adaptive thresholding methods (Niblack, 1985; [37]) and Otsu's global thresholding method (Level Otsu, 1979).

Ntirogiannis [59] conducted a study that determined that the process of image binarisation was an important part of image analysis and pipeline recognition since it affects the latter stages in the process of recognition. Examination of the binarisation techniques helps in the assessment of the process' algorithmic behaviour and the determination of its efficacy by studying the qualitative and quantitative performance of the techniques. The authors proposed the utilisation of a pixel-based binarisation evaluation method to study the historical document images which could either be machine-printed or handwritten. [60] utilised an optimisation approach for image enhancement and the reduction of the effect of non-uniform illumination. In a real-time industry environment, a genetic algorithm is used to determine the threshold value of binarisation. The binarisation technique is based on genetic algorithm outperformance global threshold binarisation methods such as the Otsu and local threshold binarisation methods similar to Bernsen.

[61] conducted a research that performed an optimisation algorithm that can be used for the intelligent binarisation of images. The binarisation method was designed based on the whale optimisation algorithm and makes use of a fuzzy *c*-means fitness function as a means of separation between the foreground (text) and the background on the Arabic handwritten manuscript image. Literature has presented many various binarisation techniques. However, determining optimal threshold values is still known as an open issue, especially for historic handwritten manuscripts in Arabic. When these manuscripts are digitised and captured, the resulting images normally have different artefacts such as smear, faint characters, bleed-through, stains, and non-uniform illumination.

However, the local threshold typically needs to be determined dynamically, which can be time-consuming and will often lead to large amounts of calculation. Thus, the local threshold method will not be able to meet the real-time requirement within the product line. Furthermore, the local threshold technique will also result into pseudo-shadow, which negatively affects image recognition.

Singh, Sharma [62] presented an adaptive binarization method for the degraded document images. Variable background, non-uniform illumination, and blur caused by humidity are the addressed degradations. Local contrast analysis, contrast stretching, thresholding, and noise removal steps are used to recover text from the document images containing non-uniform and degraded background. Wen, Li [63] Proposed binarization method for non-uniformly illumination document images. They combined Curvelet transform and Otsu's method to solve binarization task. However, most of the works reported in the field of OCR are on good-quality documents. But still it remains a highly challenging task to implement an OCR that works under all possible degradation and non-uniform illumination conditions and gives highly accurate results. Elaborate study and research on poor-quality documents need further exploration to make it a viable solution in the development of robust OCR system. In addition, segmentation error existing mainly in the peripheral image and the regions with interference is still a problem.

## 5. Deep learning and its applications

Recently, researchers have given deep learning significant attention. It has also achieved great success in terms of the various machine learning tasks. Popular public sources like Wikipedia describe deep learning as the machine learning algorithms that utilise architecture that is made up of multiple non-linear transformations so that high-level data abstractions can be modelled. In general, deep learning methods are based on artificial neural networks (ANNs). ANNs have been used less compared to other shallow models like vector machines (SVM). This is due to the fact that finding optimal solutions with ANNs is difficult. Also, a vast amount of computation is required by their training processes. The value of ANNs has been recently discovered by the research community. Furthermore, recent research can achieve excellent performance when several modern techniques are applied to address the traditional ANN flaws and attain optimal efficacy.

Deep learning is rooted in human information processing mechanisms (e.g. speech and vision), which is suggestive of the need for deep architecture to extract complex structure and construct internal representation using input signals. Typical systems for deep learning are constructed with deep architectures that are made up of various layers of nonlinear processing stages. In this system, the outputs of each lower layer are used as input and fed to its immediate higher layer. Unlike shallow architectures, deep learning methods typically possess greater representational and modelling power, which gives it the ability to handle more complex real-world applications like understanding unconstrained visual scenes and natural images. Guo, Liu [64] conducted a comprehensive study regarding deep learning and developed a categorisation system for the analysis of the current deep learning literature. The deep learning algorithms are divided into four categories based on the basic model they developed from: Convolutional Neural Networks CNN, Autoencoder, Restricted Boltzmann Machines RBM, and Sparse Coding. There is a discussion and an analysis of the approaches for these four classes. For computer vision domain applications, the advancements of CNN based schemes are the ones that are mostly reported since they are the most commonly used and the most appropriate for images.

Among the various streams of deep learning techniques, CNNs are the most popular and commonly used techniques. They are considered as top performing methods for the task of handwriting recognition [64-70]. CNNs are a representative of a specific type of multi-layer neural networks that have been developed to take advantage of how objects typically go through shifts or translational image variations. Feature detectors that function well on one portion of the image will typically find success across the whole image. Given this knowledge, every neuron found in a convolutional layer is obliged to receive input coming from a small set of nearby neurons within the previous layer. This area is referred to as the receptive field. Furthermore, the neurons found at various places (with varying receptive fields) are obliged to have identical weights. Typically, each convolutional layer is succeeded by a spatial pooling operation, which is responsible for improving its tolerance to feature variation. For every convolutional layer, the output is partitioned into small sets. Furthermore, from each set, one value is sampled (mean or the max) and this serves as the input for the next layer. For the local receptive fields, neurons found in the first convolutional layer obtain low level visual features like dots, edges, and corners. The following layers will then combine these features so that higher-order features can be detected.

CNNs have the ability to integrate classification/regression model learning and feature learning into one process. Furthermore, it is capable of directly recognising visual patterns from pixel images while involving minimal pre-processing. With the broad application of CNNs recently, hand-crafted features have begun to lose their focus in the community of object recognition. Domain knowledge that comes from experts is losing its importance when it comes to the design of visual recognition systems. CNNs have performed better on many standard handwriting recognition benchmarks [71-78], and it was observed to have a large margin compared to other types of techniques.

With the help of deep learning techniques, several problems have been addressed in terms of image processing and understanding, including document image categorisation, handwritten recognition, and blind image quality assessment. Formerly, these issues were addressed using different shallow-structure learning methods and handcrafted features. Although some methods in the past have shown reasonable performance, these problems have remained unsolved for the most part and still need some improvement. Instead of depending largely on heuristics or domain knowledge, deep learning methods can work directly on the image domain, allowing them to learn complex features and classifiers/regressors within the network.

## 6. Deep learning for Arabic Handwriting Recognition

this section covers the related work on Arabic Handwriting recognition. Hence, a number of Deep learning techniques [1, 3, 6, 8, 9] compared with the stat-of-art techniques[2, 4]. Nevertheless, all details of works were presented such as, a dataset that used, methods, numerical/characters/word level recognition, and results in accuracy for each research.

Early research into Arabic handwriting based on Deep Learning algorithms proposed by Porwal, Zhou [1]. The system utilised the deep belief network, which incrementally learns the data’s complex structure by representing it in a manner that is more compact and abstract. To assess the efficacy of the deep belief network, the system was tested on the AMA Arabic PAW dataset. This data set contains 848 testing images and 6464 training images, and also has 34 paw classes. It achieved significantly lower results compared to the results obtained from the bootstrapped dataset because the model did not have enough data to capture the distributions.

Tamen, Drias [2] developed a model that made use of the multiple classifier system to improve the recognition of Arabic handwritten words. The system was assessed on the IFN/ENIT database so that the efficiency of the model can be demonstrated. The focus of the study is the two steps involved in the recognition system, the classifiers combination and the feature extraction. The Chebyshev moments (CM) that was improved using some statistical and contour-based features (SCF) to describe word images was used during the feature extraction phase. Then, several classifiers were combined and integrated at the decision level during the classification phase. The study also took the multilayer perceptron (MLP), the extreme learning machine (ELM), and the support vector machine (SVM) classifiers into account. The system considers word recognition as the basic units and does not use any segmentation step.

Ashiquzzaman and Tushar [3] Stated that the Distributed Arabic handwriting digit recognition systems efficiency was enhanced by employing a deep learning neural networks algorithm to improve the recognition rate. The model is trained and tested on the CMATERDB 3.3.1 Arabic handwritten digit dataset. The entire dataset consists of 3000 images, making it a source of 3000 unique samples. dataset divide to 2000 training samples and 1000 test samples. Dropout used to reduce overfitting as regularization layer, that configured to randomly exclude 25% of neurons in the layer. The system applied CNN model uses several convolutional layers along with ReLU activation, with dropout used as a regularization layer. Then the output was fed into a fully connected layer with softmax activation to obtain prediction for each class. However, the added convolutional layers in the proposed CNN method enhance the accuracy in digit recognition, making the accuracy 97.4 percent. This is the recorded highest accuracy for the (CMATERDB) Arabic handwritten digit dataset.

In the study by Rani and Lehal [17], the author proposed a system which carried out offline recognition of the Arabic cursive handwritten texts using the Hidden Markov Models (HMMs). This was an analytical system with no explicit segmentation and made use of the embedded training for carrying out and improving the character recognition. They used the IFN/ENIT benchmark database, which contained around 26,459 handwritten words containing the names of 946 Tunisian towns or villages, written by many writers. In their method, the baseline estimation was carried out followed by the extraction features, which were geometric and statistical and helped in integrating the text peculiarities and the pixel distribution characteristics seen within the document image. The extraction features were modelled using the HMMs and they were trained using the embedded training. All experiments carried out using the images from the IFN/ENIT benchmark database showed a recognition rate of 87.93%.

Elleuch, Mokni [6] presented a deep learning model using SVM called as the Deep SVM (DSVM) for carrying out Offline Arabic Handwritten Recognition. This DSVM model was developed by stacking the SVMs which permitted the extraction and the learning of the features automatically based on raw images and carried out classification. The authors selected and tested a Multi-class SVM along with the RBF kernel, as a non-linear discriminative feature for carrying out the classification of the Handwritten Arabic Characters Database (HACDB). This HACDB consisted of 6.600 shapes of various handwritten characters which were written by 50 people. Every writer generated 2 forms for 66 shapes; wherein 58 shapes represented characters and 8 shapes represented the overlapping characters (consisting of 24 basic characters or overlapping characters with no dots). The DSVM model used the raw data as the input data and generated automatic strong features without using an extra-feature engineering stage. The system showed a high accuracy due to a discriminatory strength of the automatically extracted features from the input raw data which used the deep architecture (i.e., DSVM) along with the regression capabilities of the SVM classifiers.

In another study by Elleuch, Maalej [8], the authors also presented a model which used the Convolutional Neural Network (CNN) along with the SVM for the offline Arabic handwriting recognition. This method used the HACDB and the IFN/ENIT databases for proving the model efficiency. The authors studied the CNN-based SVM model performance with and without the dropout for carrying out training and recognition of the Arabic characters. The authors parameterised a convolutional layer for a setting architecture, based on the size and the map number, skipping factors, kernel sizes, along with a connection table. In the case of the IFN/ENIT database, the pre-processing step makes use of processes like the noise reduction, binarization, and filtering the input text image for improving the image quality. However, in the case of the HACDB database, the authors did not apply any kind of pre-processing step for their experiments. For the feature extraction step, the authors used CNN as the compact end-to-end model,

| Literature | Brief Description | Dataset | dataset Type | Results |
|---|---|---|---|---|
| Porwal, Zhou [1] | Deep Belief Networks which incrementally learns complex structure of the data | AMA Arabic PAW dataset | Arabic Handwriting Character | 75,05% |
| Tamen, Drias [2] | Combine several classifiers integrated at the decision level in classification phase | IFN/ENIT | Arabic Handwritten Words | 87,00 |
| Ashiquzzaman and Tushar [3] | CNN model uses several convolutional layers + Dropout used to reduce overfitting | CMATERDB 3.3.1 | Arabic Handwritten Digit | 97.4 |
| Rabi, Amrouch [4] | Hidden Markov Models (HMMs) | IFN/ENIT | Arabic Handwritten Words | 87.93%. |
| Elleuch, Mokni [6] | Deep learning model based on Support Vector Machine(DSVM) | HACDB | Arabic Handwriting Character | Error rate: 8.86% |
| Elleuch, Maalej [8] | CNN based-SVM model without and with dropout for training and recognizing Arabic characters | HACDB and IFN/ENIT | Arabic Handwriting Character | Error rate: HACDB : 5.83 IFN/ENIT :7.05 |
| Elleuch, Tagougui [9] | Deep Belief Network (DBN) + with dropout/dropconnect | Arabic handwritten script recognition | HACDB | Error classification rate of 2.73% using dropout. 2.27% using dropconnect |

which resulted in the input data in the network to be raw images. Then, they defined the optimal kernel and the penalty parameters for the SVM. Then, lastly then stated that the CNN-based SVM classifiers could offer significant state-of-the-art results without laying a lot of emphasis on the feature extraction and the pre-processing steps. In another study by Elleuch, Tagougui [9], the authors used a Deep Belief Network (DBN) for carrying out Arabic handwritten script recognition. The DBN was seen to be a probabilistic generative model which contained several stochastic hidden variable layers along with a visible neuron layer. The DBN could be trained using an efficient algorithm after greedily learning every layer, similar to an RBM, for initialising a deep network. This network contained 2 major steps, i.e., Step 1 was the unsupervised feature learning whereas Step 2 comprised of a supervised learning of the discriminating functions. The authors also applied the dropconnect for preventing the system from being over-fitted and improving the performance. The dropconnect was similar to the dropout but was used for the weights, W. Furthermore, the authors tested their model performance using the HACDB database and studied the efficiency of the unsupervised feature learning approach with the help of the DBN model in combination with the dropout/dropconnect. They obtained favourable results, where the error classification rates were 2.73% when using dropout and 2.27% when dropconnect was used.

However, Porwal, Zhou [1], Ashiquzzaman and Tushar [3], Elleuch, Mokni [6], Elleuch, Maalej [8], Elleuch, Tagougui [9] applied several different deep learning techniques, but they had not yet to discuss the developing of special (DL) for (AHW).

## Conclusion

This paper produced an overview of the research which has conducted models based deep learning for the Arabic language of handwritten Recognition task. This research has highlighted the general architectural methodology that can use for Arabic handwritten Recognition based on deep learning thoroughly and criticized. Moreover, these architectural parts had explained in detail with their corresponding related works. In addition, special attention had given to the stat of art work of handwritten Recognition with reference to the limitation on the domain of Arabic language. On the other hand, we presented a table of the literature works on Arabic domain contains a comparison between Deep learning techniques and that of other recent techniques.

## Acknowledgements

We would like to show our gratitude to Universiti Malaysia Pahang (Grant No. PGRS-180308), for supporting this study.

## References

1. Porwal, U., Y. Zhou, and V. Govindaraju. Handwritten arabic text recognition using deep belief networks. in Pattern Recognition (ICPR), 2012 21st International Conference on. 2012. IEEE.
2. Tamen, Z., H. Drias, and D. Boughaci, An efficient multiple classifier system for Arabic handwritten words recognition. Pattern Recognition Letters, 2017.
3. Ashiquzzaman, A. and A.K. Tushar. Handwritten Arabic numeral recognition using deep learning neural networks. in Imaging, Vision & Pattern Recognition (icIVPR), 2017 IEEE International Conference on. 2017. IEEE.
4. Rabi, M., et al. Recognition of cursive Arabic handwritten text using embedded training based on HMMs. in Engineering & MIS (ICEMIS), International Conference on. 2016. IEEE.
5. Dalal, N., et al., Histograms of oriented gradients for human detection. Computer Vision and Pattern Recognition CVPR Society Conference on, 2005. **1 SRC - GoogleScholar**: p. 886-893.
6. Elleuch, M., R. Mokni, and M. Kherallah. Offline Arabic Handwritten recognition system with dropout applied in Deep networks based-SVMs. in Neural Networks (IJCNN), 2016 International Joint Conference on. 2016. IEEE.
7. Lowe, D.G., Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004. **60**(2 SRC - GoogleScholar): p. 91-110.
8. Elleuch, M., R. Maalej, and M. Kherallah, A New Design Based-SVM of the CNN Classifier Architecture with Dropout for Offline Arabic Handwritten Recognition. Procedia Computer Science, 2016. **80**: p. 1712-1723.
9. Elleuch, M., N. Tagougui, and M. Kherallah, Optimization of DBN using Regularization Methods Applied for Recognizing Arabic Handwritten Script. Procedia Computer Science, 2017. **108**: p. 2292-2297.
10. Ahmad, I., S.A. Mahmoud, and G.A. Fink, Open-vocabulary recognition of machine-printed Arabic text using hidden Markov models. Pattern Recognition, 2016. **51**: p. 97-111.
11. Hafiz, A.M. and G.M. Bhat, Arabic OCR Using a Novel Hybrid Classification Scheme. Journal of Pattern Recognition Research, 2016. **1**: p. 55-60.
12. Azeem, S.A. and H. Ahmed, Effective technique for the recognition of offline Arabic handwritten words using hidden Markov models. International Journal on Document Analysis and Recognition (IJDAR), 2013. **16**(4): p. 399-412.
13. Álvaro, F., J.-A. Sánchez, and J.-M. Benedí, Recognition of on-line handwritten mathematical expressions using 2D stochastic context-free

Recognition Letters, 2014. **35**: p. 58-67.
14. Jayech, K., M.A. Mahjoub, and N.E.B. Amara, Synchronous Multi-Stream Hidden Markov Model for offline Arabic handwriting recognition without explicit segmentation. Neurocomputing, 2016. **214** : p. 958-971.
15. Lawgali, A., A survey on arabic character recognition. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015. **8**(2): p. 401-426.
16. Amrouch, M. and M. Rabi. Deep Neural Networks Features for Arabic Handwriting Recognition. in International Conference on Advanced Information Technology, Services and Systems. 2017. Springer.
17. Rani, S. and G.S. Lehal. Recognition based classification of Gurmukhi manuscripts. in Colossal Data Analysis and Networking (CDAN), Symposium on. 2016. IEEE.
18. Naz, S., et al. Challenges in baseline detection of cursive script languages. in Science and Information Conference (SAI), 2013. 2013. IEEE.
19. Borovikov, E. and I. Zavorin, A multi-stage approach to Arabic document analysis, in Guide to OCR for Arabic scripts. 2012, Springer. p. 55-78.
20. Shafait, F. and T.M. Breuel. A simple and effective approach for border noise removal from document images. in Multitopic Conference, 2009. INMIC 2009. IEEE 13th International. 2009. IEEE.
21. Stamatopoulos, N., B. Gatos, and T. Georgiou. Page frame detection for double page document images. in Proceedings of the 9th IAPR International Workshop on Document Analysis Systems. 2010. ACM.
22. Chen, K., et al. Page segmentation for historical handwritten document images using color and texture features. in Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on. 2014. IEEE.
23. Chakraborty, A. and M. Blumenstein. Preserving Text Content from Historical Handwritten Documents. in Document Analysis Systems (DAS), 2016 12th IAPR Workshop on. 2016. IEEE.
24. Farahmand, A., A. Sarrafzadeh, and J. Shanbehzadeh, Document image noises and removal methods. 2013.
25. Naz, S., et al., Segmentation techniques for recognition of Arabic-like scripts: A comprehensive survey. Education and Information Technologies, 2016. **21**(5): p. 1225-1241.
26. Nagy, G., S. Seth, and M. Viswanathan, A prototype document image analysis system for technical journals. Computer, 1992. **25**(7): p. 10-22.
classification. CVGIP: Graphical models and image processing, 1992. **54**(6): p. 484-496.
28. Wang, D. and S.N. Srihari, Classification of newspaper image blocks using texture analysis. Computer Vision, Graphics, and Image Processing, 1989. **47**(3): p. 327-352.
29. O'Gorman, L., The document spectrum for page layout analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993. **15**(11): p. 1162-1173.
30. Wong, K.Y., R.G. Casey, and F.M. Wahl, Document analysis system. IBM journal of research and development, 1982. **26**(6): p. 647-656.
31. Baird, H.S., Background structure in document images. International Journal of Pattern Recognition and Artificial Intelligence, 1994. **8**(05): p. 1013-1030.
32. Breuel, T.M. Two geometric algorithms for layout analysis. in International workshop on document analysis systems. 2002. Springer.
33. Kise, K., A. Sato, and M. Iwata, Segmentation of page images using the area Voronoi diagram. Computer Vision and Image Understanding, 1998. **70**(3): p. 370-382.
34. Jelodar, M.S., et al. A Persian OCR System Using Morphological Operators. in WEC (2). 2005.
35. Bukhari, S.S., F. Shafait, and T.M. Breuel, Layout analysis of Arabic script documents, in Guide to OCR for Arabic scripts. 2012, Springer. p. 35-53.
36. Otsu, N., A threshold selection method from gray-level histograms. Automatica, 1975. **11**(285-296): p. 23-27.
37. Sauvola, J. and M. Pietikäinen, Adaptive document image binarization. Pattern recognition, 2000. **33**(2): p. 225-236.
38. El Moubtahij, H., A. Halli, and K. Satori, HMM Off-line Arabic Handwriting Recognition Using the Components of the Frame Binary Intensity, its Grayscale Transform and its Horizontal and Vertical Derivatives. International Journal of Imaging and Robotics™, 2016. **16**(2): p. 78-90.
39. Kef, M., L. Chergui, and S. Chikhi, A novel fuzzy approach for handwritten Arabic character recognition. Pattern Analysis and Applications, 2016. **19**(4): p. 1041-1056.
40. Naz, S., et al., The optical character recognition of Urdu-like cursive scripts. Pattern Recognition, 2014. **47**(3): p. 1229-1248.
41. Saberi, A., et al., Evaluating the legibility of decorative arabic scripts for Sultan Alauddin mosque using an enhanced soft-computing hybrid algorithm. Computers in Human Behavior, 2016. **55**: p. 127-144.
42. Su, B., S. Lu, and C.L. Tan. Binarization of historical document images using the local maximum and minimum. in Proceedings of the 9th

IAPR International Workshop on Document Analysis Systems. 2010. ACM.
43. Su, B., S. Lu, and C.L. Tan, Robust document

algorithm using stochastic computing. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2016. **24**(2): p. 808-812.



processing, 2013. **22**(4): p. 1408-1417.
44. Chen, Y. and L. Wang, Broken and Degraded Document Images Binarization. Neurocomputing, 2017.
45. Agrawal, M. and D. Doermann. Clutter noise removal in binary document images. in 2009 10th International Conference on Document Analysis and Recognition. 2009. IEEE.
46. Agrawal, M. and D. Doermann, Clutter noise removal in binary document images. International Journal on Document Analysis and Recognition (IJDAR), 2013. **16**(4): p. 351-369.
47. Toh, K.K.V., N.A.M. Isa, and N. Ashidi, Noise adaptive fuzzy switching median filter for salt-and-pepper noise reduction. IEEE signal processing letters, 2010. **17**(3): p. 281-284.
48. Chinnasarn, K., Y. Rangsanseri, and P. Thitimajshima. Removing salt-and-pepper noise in text/graphics images. in Circuits and Systems, 1998. IEEE APCCAS 1998. The 1998 IEEE Asia-Pacific Conference on. 1998. IEEE.
49. Lee, W.-L. and K.-C. Fan, Document image preprocessing based on optimal Boolean filters. Signal processing, 2000. **80**(1): p. 45-55.
50. Ping, Z. and C. Lihui, Document filters using morphological and geometrical features of characters. Image and Vision Computing, 2001. **19**(12): p. 847-855.
51. Bataineh, B., S.N.H.S. Abdullah, and K. Omar, Adaptive binarization method for degraded document images based on surface contrast variation. Pattern Analysis and Applications, 2017. **20**(3): p. 639-652.
52. Hadjadj, Z., et al., A new efficient binarization method: application to degraded historical document images. Signal, Image and Video Processing, 2017: p. 1-8.
53. Jino, P. and K. Balakrishnan. Combined approach for binarization of offline handwritten documents. in Electronics and Communication Systems (ICECS), 2017 4th International Conference on. 2017. IEEE.
54. Kashef, S., H. Nezamabadi-pour, and E. Rashedi, Adaptive enhancement and binarization techniques for degraded plate images. Multimedia Tools and Applications, 2017: p. 1-17.
55. Lu, D., et al. Binarization of degraded document image based on contrast enhancement. in Control Conference (CCC), 2016 35th Chinese. 2016. IEEE.
56. Najafi, M.H. and M.E. Salehi, A fast fault-tolerant architecture for sauvola local image thresholding

the entropy of the histogram. Computer vision, graphics, and image processing, 1985. **29**(3): p. 273-285.
58. Kittler, J. and J. Illingworth, Minimum error thresholding. Pattern recognition, 1986. **19**(1): p. 41-47.
59. Ntirogiannis, K., B. Gatos, and I. Pratikakis, Performance evaluation methodology for historical document image binarization. IEEE Transactions on Image Processing, 2013. **22**(2): p. 595-609.
60. Liu, M., et al., Genetic algorithm and mathematical morphology based binarization method for strip steel defect image with non-uniform illumination. Journal of Visual Communication and Image Representation, 2016. **37**: p. 70-77.
61. Hassanien, A.E., et al. Historic handwritten manuscript binarisation using whale optimisation. in Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on. 2016. IEEE.
62. Singh, B.M., et al., Adaptive binarization of severely degraded and non-uniformly illuminated documents. International Journal on Document Analysis and Recognition (IJDAR), 2014. **17**(4): p. 393-412.
63. Wen, J., S. Li, and J. Sun, A new binarization method for non-uniform illuminated document images. Pattern recognition, 2013. **46**(6): p. 1670-1690.
64. Guo, Y., et al., Deep learning for visual understanding: A review. Neurocomputing, 2016. **187**: p. 27-48.
65. Ahmed, R., W.G. Al-Khatib, and S. Mahmoud, A survey on handwritten documents word spotting. International Journal of Multimedia Information Retrieval, 2017. **6**(1): p. 31-47.
66. Elleuch, M., N. Tagougui, and M. Kherallah, A novel architecture of CNN based on SVM classifier for recognising Arabic handwritten script. International Journal of Intelligent Systems Technologies and Applications, 2016. **15**(4): p. 323-340.
67. Hafemann, L.G., R. Sabourin, and L.S. Oliveira, Learning features for offline handwritten signature verification using deep convolutional neural networks. Pattern Recognition, 2017. **70**: p. 163-176.
68. Jin, L., et al., Online Handwritten Chinese Character Recognition: From a Bayesian Approach to Deep Learning, in Advances in Chinese Document and Text Processing. 2017, World Scientific. p. 79-126.

69. Pattern Recognition. 2016.
70. Sheng, F., et al. End-to-End Chinese Image Text Recognition with Attention Model. in International Conference on Neural Information Processing. 2017. Springer.
71. Chen, L., et al. Cascading Training for Relaxation CNN on Handwritten Character Recognition. in Frontiers in Handwriting Recognition (ICFHR), 2016 15th International Conference on. 2016. IEEE.
72. Xiao, X., et al., Design of a Very Compact CNN Classifier for Online Handwritten Chinese Character Recognition Using DropWeight and Global Pooling. arXiv preprint arXiv:1705.05207, 2017.
73. Amrouch, M. and M. Rabi, The Relevant CNNs Features Based HMM for Arabic Handwriting Recognition. International Journal of Imaging and Robotics™, 2017. **17**(3): p. 98-109.
74. Maitra, D.S., U. Bhattacharya, and S.K. Parui. CNN based common approach to handwritten character recognition of multiple scripts. in Document Analysis and Recognition (ICDAR), 2015 13th International Conference on. 2015. IEEE.
75. Sun, Y. and Q. Liu, Attribute Recognition from Clothing Using a Faster R-CNN based Multi-task Network. International Journal of Wavelets, Multiresolution and Information Processing, 2018.
76. Wu, C., et al. Handwritten character recognition by alternately trained relaxation convolutional neural network. in Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on. 2014. IEEE.
77. Xiao, X., et al., Building Fast and Compact Convolutional Neural Networks for Offline Handwritten Chinese Character Recognition. arXiv preprint arXiv:1702.07975, 2017.
78. Zhong, Z., L. Jin, and Z. Xie. High performance offline handwritten chinese character recognition using googlenet and directional feature maps. in Document Analysis and Recognition (ICDAR), 2015 13th International Conference on. 2015. IEEE.

View publication stats