Hybrid Intelligence: Filtering and Deep Learning for Handwritten Text Recognition

Arshad Iqbal
K. A. Nizami Centre for-Quranic Studies
Aligarh Muslim University
Aligarh-202002, India

Shabbir Hassan
Department of Computer-Science
Aligarh Muslim University
Aligarh-202002, India

Maria Qamar
Radha Govind University
Department of Computer Science-
and Technology
Ramgarh – 829122, India

ABSTRACT

Handwritten text recognition systems have gained substantial attention in pattern recognition and artificial intelligence due to their applications in digitizing historical documents and automated reading tasks. This study proposes a novel preprocessing approach for recognizing Kufic manuscripts, aiming to improve the accuracy of handwritten text segmentation. The approach consists of three stages: noise removal, thresholding, and additional noise removal after thresholding. Hybrid filtering, using Wiener and Median filters, effectively reduces noise in the input images. The u-Net deep learning model is employed for precise thresholding and segmentation of the handwritten text. The post-thresholding noise removal step refines the segmented regions and eliminates residual noise artifacts. The proposed approach contributes to enhancing the Quranic Kufic Manuscripts Recognition System (QKMRS) and advancing the digitization of historical documents and the preservation of cultural heritage. The performance measures for noise removal include Mean Square Error (MSE) and Peak Signal Noise Ratio (PSNR). For segmentation and thresholding Intersection over union (IoU), entropy and dice coefficient are used.

Keywords

Filtering, Kufic Manuscript, Median Filter, Thresholding, U-net, Wiener Filter.

1. INTRODUCTION

Handwritten text recognition systems have gained substantial attention in pattern recognition and artificial intelligence, primarily due to their potential applications in various fields, including digitization of historical documents, automatic postal mail sorting, and automatic bank cheque reading. Particularly, the recognition of Kufic manuscripts holds significant value due to their historical and cultural importance. In this context, digital image preprocessing becomes paramount for effectively analysing and manipulating the Kufic manuscript images, significantly facilitating the subsequent feature extraction and pattern recognition stages.

The ancient Kufic Quranic manuscripts are more challenging to read by the computer than normal Arabic handwritten texts. The Kufic manuscripts have many problems like bleed-through, blur, uneven illumination and deterioration, which makes them difficult to read. Kufic scripts have great dissimilarities in the letterforms and their proportions. These problems do not exist in normal Arabic handwritten texts. These make the machine interpretation of the Kufic manuscripts more challenging [1]. Authors devoted more time to preprocessing to remove all the above problems, which makes the recognition task fast and accurate.

The proposed preprocessing approach consists of three key stages: noise removal, thresholding, and additional noise removal after thresholding. The initial stage focuses on reducing the noise and enhancing the quality of the input images. In this stage, a hybrid filtering technique is employed, combining the strengths of both Wienerand Median filtering methods. The Wiener filter, known for its optimal noise reduction capabilities, effectively reduces Gaussian noise, while the Median filter eliminates salt-and-pepper noise. By integrating these two filters, the proposed approach robustly addresses the image noise, resulting in cleaner and more understandable inputs for subsequent processing steps. Following the noise removal stage, thresholding transforms the images into binary form, separating the foreground (handwritten text) from the background. Thresholding helps segment text regions and facilitates further analysis and recognition tasks. This study employs a state-of-the-art deep learning model, U-Net, for thresholding [2]. Designed initially for biomedical image segmentation, U-Net has proven its efficacy in highly precise applications, including document image binarization. The application of U-Net in this context provides precise segmentation of the handwritten text, significantly improving the effectiveness of subsequent recognition processes.

In addition to the initial noise removal and thresholding stages, the proposed approach also incorporates an additional noise removal step after thresholding. This post-thresholding noise removal step further refines the segmented regions and eliminates any residual noise artifacts. By applying noise removal both before and after thresholding, the proposed approach aims to maximize noise reduction and improve the accuracy of the segmentation results, ultimately enhancing the overall performance of the Quranic Kufic Manuscripts Recognition System (QKMRS). This proposed method intends to provide an enhanced Quranic Kufic Manuscripts Recognition System (QKMRS). The proposed approach has the potential to contribute significantly to the digitization of historical documents, preservation of cultural heritage, and facilitation of scholarly research in this domain. Moreover, the proposed preprocessing methodology could be generalized and adapted for other similar tasks, thereby contributing to the broader field of handwritten text recognition systems.

1.1 Contribution

a. *The paper proposes a novel preprocessing approach designed for Quranic Kufic manuscript images. It combines hybrid filtering, thresholding using U-Net,* *and additional noise removal techniques to enhance the quality and accuracy of the recognition system.*

b. *To conduct research a dataset (Quranic Kufic Dataset) that is made publicly available on Kaggle https://www.kaggle.com/datasets/sheikhburhanulhaque/quranic-kufic-dataset*
c. *The proposed preprocessing methodology can be generalized and adapted for other similar tasks in the field of handwritten text recognition systems. This offers potential applications in various domains, expanding the scope of the research and providing valuable insights for advancements in the broader field.*
d. *The paper's novel combination of filtering, thresholding, and noise removal techniques contributes to advancing preprocessing methods for handwritten text recognition systems. It offers a comprehensive and effective noise reduction and segmentation solution, addressing the challenges of degraded manuscript images.*

2. RELATED WORK

There are various methods available for image filtering and thresholding, which some researchers have already used.

The most popular global thresholding method has been proposed by Otsu [4]. Otsu's method gives better results where image illumination is not uneven. However, Otsu's method does not work fine in old images, especially in handwriting images with various kinds of degradations (e.g., non-uniformity in the foreground and background degraded text regions, noise and stain).

In the context of noise removal, several studies have explored the use of filters for noise reduction. For example, [3] proposed the use of adaptive filters for denoising handwritten text images.

In terms of image enhancement, contrast enhancement techniques have been widely used to improve the readability of handwritten text [4] and proposed an adaptive histogram equalization method for enhancing low-contrast handwritten text images.

Binarization is an important step in segmenting text regions from the background. Various thresholding techniques have been studied for binarization, including Otsu's [5] and Sauvola's [6]. These methods aim to find an optimal threshold for separating foreground and background pixels.

Normalization techniques have also been explored to address variations in illumination and background in handwritten text images. For instance, [7] proposed a background normalization method based on local average filtering to improve text recognition accuracy.

For filtering or denoising in the spatial domain, the authors [8] have proposed using mean, median and Wiener filters; for the frequency domain, Gaussian and Butterworth filters were suggested. The authors have classified historical images into six categories: good condition, shadows, spots, transparent page, broken characters and coloured characters. The Wiener filtering has produced the best result in all the above categories. Niblack and Sauvola thresholding methods have produced efficient results in all categories except broken characters.

The authors in [9] have proposed a hybrid method to enhance old manuscript images with a degraded background. The hybrid method comprises denoising by Wiener filter, thresholding by the Niblack technique, and improving the quality of text regions by post-processing.

Recognizing Kufic manuscripts presents unique challenges due to their distinctive characteristics, including variations in letterforms, proportions, and degradation over time. Researchers have proposed specific preprocessing approaches tailored for Kufic manuscripts to address these challenges. For instance, [10] developed a preprocessing pipeline for Kufic manuscripts, including noise removal, binarization, and text line segmentation. They utilized filters such as the Gaussian and median filters for noise reduction and Otsu's method for thresholding.

Another study [11] addressed the challenges of uneven illumination and ink fading in Kufic manuscripts. They proposed an illumination correction technique based on morphological operations and adaptive histogram equalization to enhance the quality of the manuscript images.

Furthermore, [12] explored the restoration of degraded Kufic manuscripts using inpainting techniques and image enhancement methods. Their approach aimed to fill in missing or damaged parts of the manuscripts while preserving the original characteristics of the text.

The authors of [13] present a comprehensive study on using soft computing techniques for the machine reading of Quranic Kufic manuscripts. The research addresses the unique challenges associated with Kufic scripts, such as letterform variations and deterioration. It aims to develop an automated system for accurately interpreting and analysing these historical documents. By integrating fuzzy logic, genetic algorithms, and neural networks, the proposed framework demonstrates promising results in accurately interpreting and understanding the complex structure and content of Kufic manuscripts. The findings contribute to advancing machine reading techniques specifically tailored for Kufic scripts, offering innovative solutions for digitizing, preserving, and analysing historical documents. The manuscript provides valuable insights into pattern recognition and artificial intelligence, presenting a comprehensive framework that combines soft computing techniques to enhance the efficiency and accuracy of machine reading systems for Quranic Kufic manuscripts.

The authors have proposed a local thresholding method in Palm-leaf Manuscript images in [14]. They have applied the decision-based median filter to remove the noises and adaptive histogram equalization to enhance the image. They have used the Sauvola technique for thresholding purposes. They evaluated their technique using time complexity and Shannon entropy. In [5][8][9][14], the authors have used a single method, like Wiener or Median, as filtering, while Sauvola has been used for thresholding for recognition of ancient manuscripts. These methods do not perform well with old, degraded, complex Quranic Kufic manuscripts.

In [15], the authors employed the grayification algorithm, which is used as a colour-to-grayscale conversion approach that improves contrast. It uses both brightness and colour information to enhance the difference between the background and the foreground.

In [16], the authors employed a binarization approach based on background estimation and energy minimization for degraded historical document images.

In [17], the authors presented a completely automated segmentation technique. It is based on the Harris corner detector, which evaluates and segments the content from the background and noise in manuscripts. The Poisson approach is also used to propose a manuscript reconstruction method from the gradient field.

These related studies highlight the significance of preprocessing techniques in improving the accuracy of handwritten text recognition, particularly in the context of Kufic manuscripts. The proposed research builds upon these previous works by introducing a novel preprocessing approach that combines hybrid filtering and deep learning for more accurate segmentation of handwritten text in Kufic manuscripts.

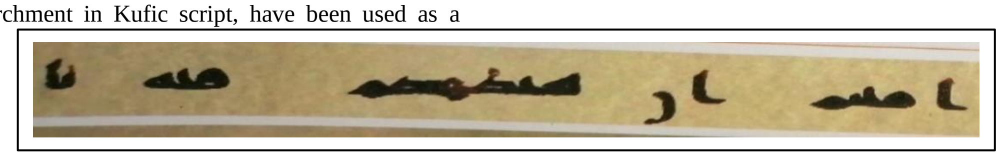
3. DATASET DESCRIPTION

The Quranic Kufic Manuscripts Recognition System dataset comprises a rare 7th-century-old Quranic manuscript written in Kufic script and additional resources of Kufic Quranic Scripts. The dataset is publicly available on the Kaggle platform as https: //www.kaggle.com/datasets/sheikhburhanulhaque/ quranic-kufic-dataset', with the initial version (V1) consisting of 82 manuscript images and an updated version (V2) containing 492 images. The updated version incorporates images created through data augmentation techniques. Data augmentation techniques have been applied using two distinct approaches: creating separate augmented images with specific variations in scaling, brightness, contrast, and blurring and combining these augmented properties into a single image. These approaches enhance the diversity and learning capabilities of the training dataset, enabling machine learning models to better recognize and analyse Kufic Quranic manuscripts.

In [18], the author has compiled a visual catalog and record, a collation of seventeen Quranic manuscripts of Surah Isra (17th chapter of The Holy Qura'an). Rare resources of nine manuscripts available in various renowned libraries, namely, Sana, Samarqand, Topkapi, Ali-Rampur, Tunis 114, Tunis 41, Oglu 2, Nurosmaniye 27 and King Fahd Library, which are written on parchment in Kufic script, have been used as a dataset for this research work. This surah contains 1559 Kufic words. This primary dataset, i.e., the KUFIC dataset, has been prepared mainly with the help of this surah. Early Kufic script used horizontal strokes. In some manuscripts, the script used dots for diacritical marks. Red dots were placed below or above characters for vowels. The Kufic script has dissimilarities in the letterforms and their proportions. Round characters have extremely small counters. The *alif* has a curved shape on its lower right side. These are some unique characteristics of the early Kufic script. Figure 1 is a sample of such a KUFIC dataset.

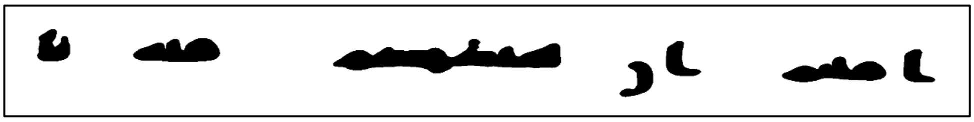
Fig 1: Samples of early Kufic Script

By employing LabKit, an annotation tool specializing in semantic segmentation, for the training image annotation process. In the annotation process, a label of 0 was assigned to denote the text regions, while a label of 1 represented the background. This annotation strategy allowed for the clear differentiation between text and background, enabling the accurate identification and segmentation of handwritten text in the Kufic Quranic manuscript images. The image masks generated through LabKit served as ground truth references, providing essential training data for the development of recognition system. Figure 2 (A-C) shows the corresponding annotated and masked image.



A. Original sample from the dataset

B. Represents the annotated image created using the *labkit* tool



C. Ground truth bitmap mask of the annotated image where one is for the background and 0 represents text

Fig 2 (A-C): Annotated and masked images

4. METHODOLOGY

The role of filtering and thresholding is to prepare the Kufic manuscript image for the further recognition stages. The main purpose of filtering is to reduce the noise from images and keep only the desired information for thresholding. In this study, a novel preprocessing approach for Quranic Kufic Manuscripts Recognition System (QKMRS) is proposed and illustrated in Figure 3. The following is a detailed description of the methodology applied.

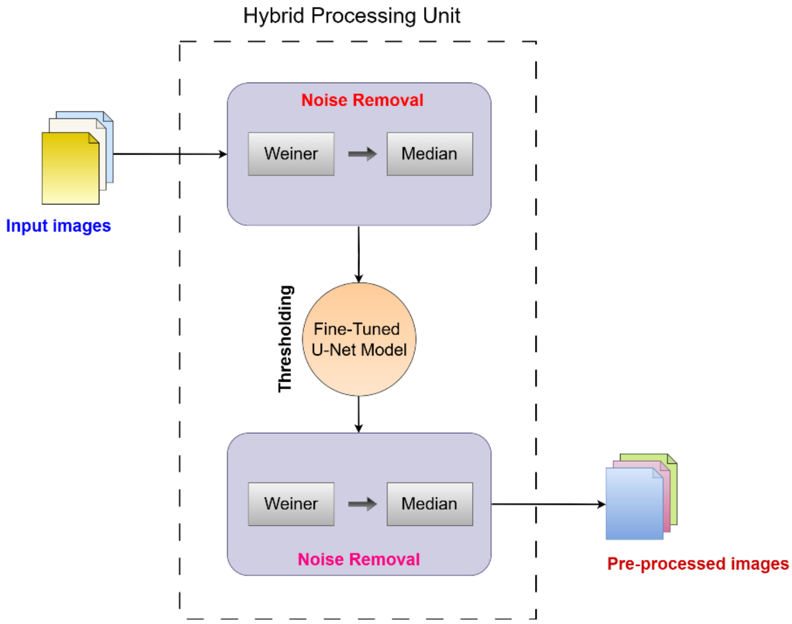


Fig 3: Proposed hybrid approach of Kufic qur'anic manuscript pre-processing

The initial preprocessing stage involves noise removal to enhance the quality of the input images. This is achieved through a hybrid filtering technique combining Wiener and Median filters, effectively reducing Gaussian and salt-and-pepper noise. The filtered images are then thresholder using the U-net deep learning model. The U-net architecture, known for its success in biomedical image segmentation, is trained to accurately binarize the images, separating the foreground (text) from the background. To evaluate the performance of the approach, the noise reduction capability of the hybrid filtering is assessed using Mean Square Error (MSE) and Peak Signal Noise Ratio (PSNR). Additionally, the performance of U-net thresholding is measured by comparing the segmented images with the manually annotated ground truth. Performance metrics such as IoU (Intersection over Union) and Dice coefficient are utilized for this evaluation.

4.1 Why Wiener and Median Filtering

The choice to use Weiner and Median filters to preprocess these images is rooted in the unique strengths these methods offer.

Wiener Filter: This filter is selected because it's more than just a noise reduction tool; it also works to undo the blurring (degradation) that might have happened during the image-capturing process [19]. It estimates the local mean and variance around each pixel, considering the image and the noise as random processes, to adapt its behaviour on a per-pixel basis. This approach makes it an excellent choice for dealing with images that contain a lot of detail, such as the intricate calligraphy in the Kufic Quranic Manuscripts. It's particularly efficient at reducing Gaussian noise, a common type of noise in photographs that could be present due to aging or the digitization process.

The Wiener filter is better than the Inverse filter. Statistical characteristics of noise, as well as degradation function, are incorporated into the restoration process. In this method, noises and images are considered as a random process and the main motive is to find a $\hat{f}$ of the unspoiled image f, such that the mean square error between them can be reduced. Equation (*1*) is expressed as an error measure.

$$e^2 = E * \{(f - \hat{f})^2\} \quad (1)$$

where, E = expected value

It is supposed that the noise and the image are not correlated; none of them has zero mean and gray levels in the estimate are a linear function of the degraded image. Based on these assumptions, the least error function is shown in the frequency domain, as represented in Equation (*2*).

$$\hat{F}(k,l) = W(k,l) * G(k,l) \quad (2)$$

Where Equation (3) is called Wiener Filter [11]. It eliminates distorted and additive noise simultaneously.

$$W(k,l) = \frac{|H(k,l)|^2}{H(k,l)\left[|H(k,l)|^2 + \frac{S_\eta(k,l)}{S_f(k,l}\right]} \quad (3)$$

Where, $H(k,l)$ = degradation function:

$$|H(k,l)|^2 = H^*(k,l) * H(k,l) \quad (4)$$

$H^*(k,l)$: complex conjugate of $H(k,l)$

$S_\eta(k,l)$: power spectrum of the noise

$S_f(k,l)$:power spectrum of un-degraded image

$G(k,l)$: degraded image

Median Filter: On the other hand, the median filter is a non-linear method known for its effectiveness in removing 'salt-and-pepper' noise while preserving edges. This type of noise, characterized by sparsely occurring white and black pixels, can be caused by image transmission errors, analog-to-digital converter errors, or bit errors in memory. This kind of noise could be present in Kufic manuscripts due to deterioration, stains, or ink spills. By selecting the median value in the neighborhood of each pixel rather than the mean, the filter is less influenced by outliers, thus effectively reducing this type of noise without overly blurring the image. This property is important in preserving the detailed calligraphic strokes in Kufic script.

The median filter is a type of non-linear filter. This filter is applied to remove noise from an image and the median of the surrounding area of that pixel changes the intensity value of any pixel.

For, $X\{x1, \quad x2, \quad x3, \cdots \quad xn\}$ and $x1 \le x2 \le x3 \le \cdots \le xn \in R$

The new value of a pixel (x, y) of Image I is calculated in Equation (5).

$$median(X) = \begin{cases} x_{n+1/2}, if \quad n \quad is\, odd \\ \frac{1}{2}\left(x_{n/2} + x_{(n/2)+1}\right), if \; n \quad is\, even \end{cases} \tag{5}$$

The combination of Wiener and Median filters effectively reduces noise and improves image quality for Kufic Quranic Manuscripts, thereby enhancing the overall accuracy of the recognition system.

4.2 Thresholding Techniques for Quranic Kufic Manuscript Images

Thresholding is a very powerful image-segmenting technique. It is used to distinguish foreground from background. Thresholding is used to convert grayscale images into binary images. Thresholding techniques like Otsu's [5], Bradley's [20], Niblack's [21], and Sauvola's [6] have been widely used for image binarization. However, it utilized the U-Net deep learning-based approach for superior thresholding on complex and varied images.

4.2.1 Fine Tuning of U-Net Model

The U-Net model, celebrated for its "U" shape, is a convolutional neural network (CNN) that has proven effective for biomedical image segmentation tasks. Comprising an encoding (contraction) path and a decoding (expansion) path, it constructs intricate abstractions from input data. The encoding pathway is a traditional succession of convolutional and max pooling layers, while the decoding pathway leverages these abstracted features to reconstruct the segmented structures of the original image. One significant feature of U-Net's architecture is the skip connections between layers at equivalent depth in the encoder and decoder, enhancing the learning of localized features and enabling a more precise output.

Utilizing this potent architecture, the paper propose a transfer learning approach tailored to the task of thresholding Kufic Quranic manuscript images. For this, a pre-trained U-Net model initially trained on a large, diverse dataset is adapted to the specific characteristics of images, set at an input size of 128x128 pixels.

In this approach, the weights of the pre-trained U-Net are frozen during training, preserving the features it has learned and enabling the model to retain its general feature detection capabilities. On appending additional layers to the base U-Net architecture to suit the specific task. This extension includes a convolutional layer with a 3x3 kernel, a batch normalization layer for training stability, and a ReLU (Rectified Linear Unit) activation function for non-linearity. The purpose of these additional layers is to enable more refined feature extraction specific to the manuscript images. A 1x1 convolution is applied at the end to map these refined features to the desired output, i.e., the thresholded image. The steps for the thresholding using U-net are illustrated in Algorithm 1.

Since, at this point only one type of object to predict thus, it is a binary task. By predicting a mask, i.e., where the object of interest is present, thus assign the following mapping:

a. **1 for the background class.**
b. **0 for the object of interest (text)**

The proposed model blends the power of a pre-trained U-Net and custom layers to perform effective thresholding on Kufic Quranic manuscript images. The model leverages the learned image features from diverse datasets and further refines them to the specific task's demands, thus presenting a promising approach for manuscript image thresholding. The proposed model combines the strengths of a pre-trained U-Net and additional custom layers to perform thresholding on Kufic Quranic manuscript images effectively. This architecture efficiently leverages learned image features from diverse datasets and further refine them to the specifics task.

Algorithm 1: Pre-processing of Kufic Quranic Manuscript Images

```
Input: Kufic Quranic manuscript image, I
Output: Thresholded and noise-reduced image, I'

STEP 1:  Load Image
         Action: Load the Kufic Quranic manuscript image, I, from the source file location,
         F
STEP 2:  Apply Hybrid Filtering
         Input: Loaded Image, I
         Output: Noise-reduced Image, In
         Action: Apply a two-step filtering process on the loaded image
Step 2.1: Apply Wiener Filter: Let W() represent the Wiener filter operation. Apply the Wiener filter
         to reduce additive noise, especially Gaussian noise: In = Wx(I)
Step 2.1: Apply Wiener Filter: Let W() represent the Wiener filter operation. Apply the Wiener filter
         to reduce additive noise, especially Gaussian noise: In = Wx(I)
         Step 2.2:    Apply Median Filter: Let M() represent the Median filter operation
STEP 3:  Threshold image using U-Net
         Input: Noise-reduced Image, In
         Output: Thresholded Image, It
         Action: Utilize the U-Net model, U(), to perform image thresholding.

Step 3.1: Preprocess Image for U-Net: Resize the noise-reduced Image, In, to match the input dimensions
         expected by the U-Net model (e.g., 128x128 pixels), yielding the resized Image, Ir.
```

Step 3.2: Apply U-Net Thresholding: Pass the preprocessed image through the U-Net model to perform thresholding, segregating the handwritten text from the background: $I_t = U(I_r)$.

STEP 4: Apply Noise Removal After Thresholding

Input: Thresholded Image, I_t
Output: *Final* processed Image, I'
Action: Further refine the thresholded image to remove any remaining noise artifacts.

Step 4.1: Apply Noise Removal: Apply a noise removal technique, such as Median and Gaussian filtering, to the thresholded Image, I_t, *to eliminate any residual noise and smoothen the segmented regions:* $I' = N(I_t)$.

STEP 5: Post-process Thresholded Image

Input: Thresholded Image, I_t
Output: Final processed Image, I'
Action: Refine the thresholded image and prepare it for further processing.

Step 5.1: Resize Thresholded Image: If necessary, resize the thresholded image back to its original dimensions, yielding the final processed image, I'.

Step 5.2: Save the Thresholded Image: Save the processed image, I', in the required format for further analysis or text recognition.

Why U-Net? U-Net is a convolutional neural network initially designed for biomedical image segmentation. It has since been widely adopted in various fields due to its excellent performance in image segmentation tasks. U-Net has some distinct advantages compared to traditional thresholding methods, especially when dealing with complex images such as Kufic Quranic Manuscript Images.

- ***Learning-based Approach:*** *Traditional thresholding techniques, such as Otsu's, Bradley's, Niblack's, and Sauvola's methods, use predetermined rules or mathematical models to make the binarization. In contrast, U-Net learns the best way to perform the task directly from the data during the training phase. This makes it a more flexible and adaptive approach, capable of handling complex scenarios.*
- ***Contextual Information****: While traditional thresholding techniques consider only local pixel values to determine the threshold, U-Net, with its convolutional structure, can consider both local and more global contextual information when making predictions. This makes U-Net especially effective when dealing with images with varying backgrounds and noise levels.*
- ***Ability to Handle Uneven Illumination:*** *One of the biggest challenges in thresholding old manuscript images is uneven illumination. Traditional thresholding techniques often struggle with this issue, resulting in poor segmentation. U-Net, on the other hand, can effectively handle uneven illumination due to its learning-based approach.*
- ***Better Handling of Complex Structures****: Kufic Quranic Manuscript Images can have complex structures and patterns that are hard for traditional thresholding methods to handle. With its ability to learn hierarchical features through its deep architecture, the U-Net model can better deal with these complexities.*
- ***End-to-End Training****: U-Net is trained end-to-end, which means it learns to map raw input images to their corresponding segmented images directly. This eliminates the need for manual feature engineering, which can be a complex and time-consuming task in traditional thresholding methods.*

In conclusion, while Otsu's, Bradley's, Niblack's, and Sauvola's methods can provide reasonable results in some situations, they often struggle with uneven illumination, noise, and complex structures. On the other hand, the U-Net approach, despite requiring substantial training data and computational resources, can handle a wide variety of complex scenarios more effectively due to its learning-based nature and ability to consider both local and global contextual information. A U-Net-based thresholding approach is a powerful tool that can overcome many of the challenges traditional thresholding techniques face when working with complex and degraded images such as Kufic Quranic Manuscript Images. In section 4.2.2, the most widely used thresholding techniques have been discussed.

4.2.2 Other Thresholding Methods

Otsu's Method

Otsu's thresholding method is applied to automatically determine the binarization level from the histogram [4].

This method exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes separated by threshold T.

$$\sigma_w^2(T) = \omega_0(T)\sigma_0^2(T) + \omega_1(T)\sigma_1^2(T) \quad (6)$$

Where, ω_0 and ω_1 = weights, the probabilities of the two classes and σ_0^2 , σ_1^2 = variances of two classes.

Otsu's method is the most efficient method among the global methods [13].Global thresholding does not perform well in low-quality images i.e., uneven illumination and low contrast. So, it is not suitable for degraded old manuscript images.

When uneven illumination, contrast, or background noise exists in degraded old document images, local thresholding techniques are used. The local thresholding techniques used in this work are Bradley, Niblack and Sauvola methods.

Bradley's Method

Bradley thresholding method is used for adaptive or local thresholding. In this method, the input image is used for the integral image and a threshold value is calculated for the local mean intensity of the neighbourhood of each pixel. The adaptive method provides a more precise output as compared to global thresholding. In MATLAB, the *adapt thresh* function is available for the implementation of this method [22].

Niblack's Method

This method was proposed by Niblack in 1986. For each pixel, a threshold is computed based on the local mean and standard deviation in local thresholding. In this method, the threshold value of each pixel (i, j) is calculated by the window size of MxM blocks as:

$$T(i, j) = m(i, j) + K * \delta(i, j) \tag{7}$$

Where, $m(i, j)$ = local mean and,

$\delta(i, j)$=local standard deviation.

K is the bias and the default value of K =-0.2.

Sauvola's Method

Sauvola's method enhances Niblack's method, mainly for ancient uneven illuminated documents. The window size of MxM blocks calculates a threshold value of each pixel in Sauvola's method as:

$$T(i, j) = m(i, j) * \left[1 + K\left(\frac{\delta(i, j)}{R} - 1\right)\right] \tag{8}$$

Where, $m(i, j)$ = local mean,

$\delta(i, j)$ = local standard deviation and,

R=Maximum value of standard deviation

K is the bias, and this value ranges from 0.2 to 0.5.

The complete steps involved in this pre-processing phase is shown in Algorithm 1.

5. EXPERIMENT AND RESULTS

The proposed preprocessing approach and performance evaluation were conducted using the Google Colaboratory platform with a Tesla T4 GPU allocated for computation. The Python 3.0 platform with Keras and Tensorflow libraries was employed for programming. The images were processed using the allocated GPU, which accelerated the computations and improved the efficiency of the preprocessing pipeline. The performance evaluation was performed on the pre-processed images using various metrics, including PSNR, MSE, entropy, IoU, and Dice coefficient.

5.1 Experimental Setup

Table 1 presents the parameters used for the noise removal filters in the proposed preprocessing approach. The Weiner and Median filters are applied with specific window sizes to target different types of noise commonly present in the Kufic manuscript images. These parameters were selected based on empirical analysis and their effectiveness in reducing specific types of noise.

Table 1: Parameters for noise removal filters

Filter	Window Size	Description
Weiner	3x3 pixels	Reduce Gaussian noise
Median	5x5 pixels	Reduce Salt-and-pepper Noise

Table 2 presents the parameters used in training the U-Net-based thresholding model for Kufic Quranic manuscript images. The Adam optimizer, a batch size of 16 and 40 epochs, was chosen for efficient and effective training. A learning rate of 0.001 balanced convergence speed and stability. The Binary Cross-Entropy loss function, ReLU activation function, and additional layers enhanced feature extraction and mapping. The specific values for the parameters used in the training of the U-Net-based thresholding model were carefully selected based on empirical experimentation and chosen to optimize the model's performance in precise segmentation of handwritten text in Kufic Quranic manuscripts.

Table 2: Parameters for the training of U-Net based thresholding model

Parameter	Value
Input size	128x128
Optimizer	Adam
Batch size	16
epoch	40
Learning rate	0.001
Loss function	Binary Cross-Entropy
Activation function	ReLU
Additional layers	Conv2D (3x3), BatchNormalization, ReLU, Conv2D (1x1)

5.2 Evaluation Metric

The following metrics are used as a performance measure for measuring the quality of the recreated image as compared to the original image.

a. *Mean Square Error (MSE)*
b. *Peak Signal to Noise Ratio (PSNR)*
c. *Entropy*
d. *Intersection over Union (IOU)*
e. *Dice coefficient (F1 score)*

5.2.1 MSE

MSE and PSNR are the two most commonly used error sensitivity measures [23].

The square of differences in the pixel values between the corresponding pixels of the two IxJ monochrome images. The MSE of IxJ size image is defined in Equation (9).

$$MSE = \frac{1}{I * J} \sum_{m=1}^{I} \sum_{n=1}^{J} \left(X_{mn} - \hat{X}_{mn}\right)^2 \tag{9}$$

Where,

J= number of rows

I= number of column

X_{mn}= *intensity of* mn^{th} *pixel in original image*

$\hat{X}_{mn}$ Intensity of mxn^{th} pixel in the reconstructed Image

5.2.2 PSNR

PSNR is defined as the ratio between the signal's maximum possible power and the corrupted noise's power. The quality of a reconstructed image depends on the PSNR value. The higher PSNR value gives better quality of a reconstructed image. PSNR is inversely proportioned to MSE. Lower the value of MSE gives a higher PSNR value, i.e., higher the quality of a reconstructed image. The unit of PSNR is dB and defined in Equation (10).

$$PSNR = 10 \log_{10}\left(\frac{R^2}{MSE}\right) \tag{10}$$

Where,

R=255, for 8-bit images

MSE, Mean Square Error

5.2.3 Entropy

The randomness of pixel information is measured in the image by entropy. Lower the value of entropy gives less randomness in the image information [24]. This measure is very useful for deciding a thresholding algorithm. Shannon entropy is defined in Equation (11):

$$E = -\sum_{i=1}^{m} p_i \log_2 p_i \tag{11}$$

Where,

p_i , is probability ith gray level intensity of the pixel

m, is possible outcomes

5.2.4 Intersection over Union (IoU)

It measures the overlap between the predicted segmented region (output) and the ground truth (reference). It is calculated as the ratio of the intersection of the predicted and ground truth regions to the union of these regions. IOU measures how accurately the predicted segmentation aligns with the true segmentation, with a value of 1 indicating a perfect match.

IoU = Intersection Area / Union Area (12)

Where,

Intersection Area is the area of overlap between the predicted segmentation mask and the ground truth mask. While Union Area is the total area encompassed by both the predicted and ground truth masks.

Mathematically, the equations for calculating the intersection area and union area are as follows:

Intersection Area = | *Predicted Mask* ∩ *Ground Truth Mask* |

Union Area = | *Predicted Mask* ∪ *Ground Truth Mask* |

5.2.5 Dice Coefficient

Dice score, also known as the Dice-Sørensen coefficient (it has many more names), is another name for the F1 score that you will find a lot more in the context of computer vision.

Here is how it is defined when applied to two sets of pixels A and B, where A is the set of true pixels and B the set of predicted ones:

DSC(A, B) =2|A∩B|/ |A|+|B| (13)

The dice score ranges from 0 to 1, with a value of 1 indicating perfect overlap and a value of 0 indicating no overlap.

5.3 Performance Evaluation

The performance evaluation of the transfer learning U-Net model on the dataset of 492 Kufic Quranic manuscript images demonstrated its effectiveness in the precise segmentation of handwritten text. The work utilized 472 images for training the model and the rest 20 images were used to validate the model. The results shown in Figure 4 shows that the U-Net model achieved an average IoU of 0.88, indicating a high overlap between the segmented text regions and the ground truth. The Dice score averaged 0.87, demonstrating the model's overall accuracy in handwritten text segmenting. Additionally, a visual inspection of the segmented images revealed that the U-Net model effectively separated the handwritten text from the background, resulting in clear and well-defined text regions. The model successfully preserved the intricate details and shapes of the Kufic script, indicating its ability to capture the unique characteristics of these manuscript images accurately.

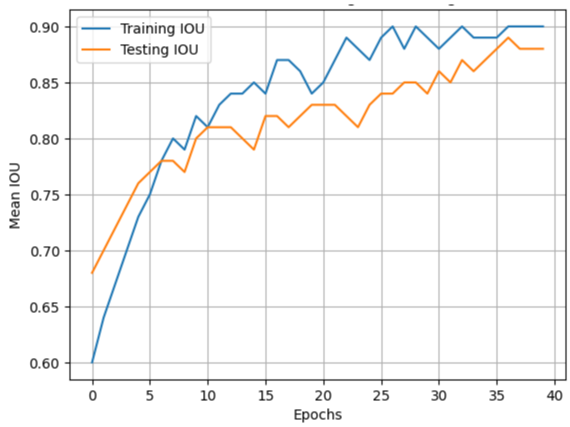


Fig 4: The flat behaviour at the end of the graph indicates that further training would not have yielded much improvement of the model. The best-performing model was found after training for 34 epochs (IOU of 88.6 %)

The lower entropy value indicates less randomness in image information and efficient thresholding. The low entropy measure indicates that the proposed approach is efficient for Kufic manuscript images. Figure 5 shows entropy comparison results in thresholding algorithms. The proposed method achieves lower entropy values than other thresholding techniques, including Otsu's Bradley and Sauvola thresholding methods. This suggests that the proposed method preserves more information and produces higher-quality segmented text regions.

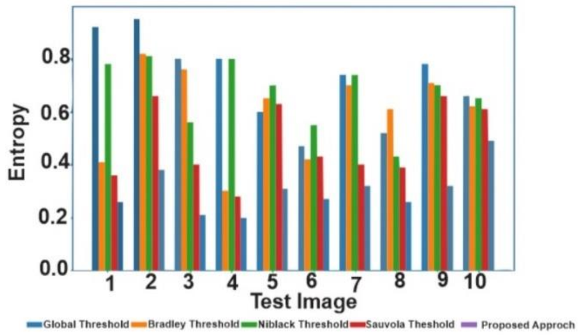


Fig 5: Comparison of Entropy Results of Thresholding Algorithms

MSE and PSNR image quality metrics evaluated the efficiency of the filtering method. Figures 6 and Figure 7 show the proposed method's average MSE and PSNR values compared to the average and Gaussian filter techniques. It can be observed that the proposed method achieves higher PSNR values and lower MSE values compared to the other technique [25]. This indicates that the proposed method effectively reduces noise and enhances image quality in Kufic manuscript images.

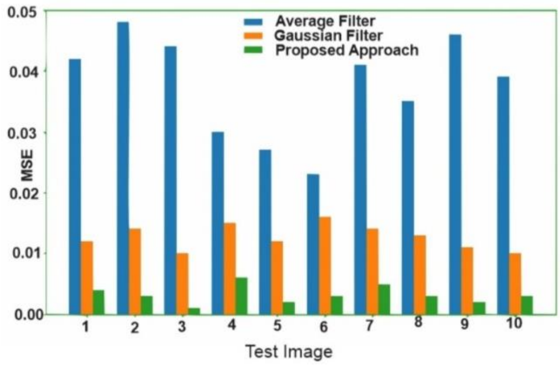


Fig 6: MSE comparison of Noise removal (filtering) methods

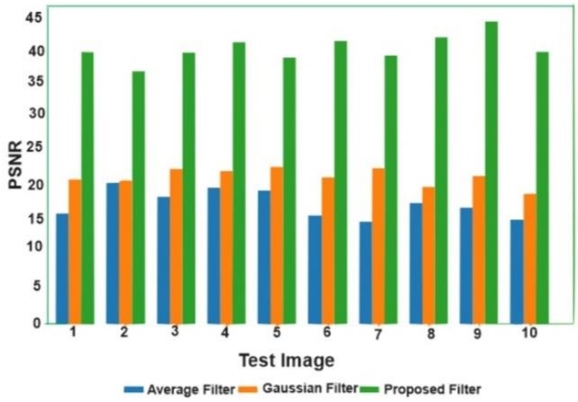


Fig 7: PSNR comparison of Noise removal (filtering) methods

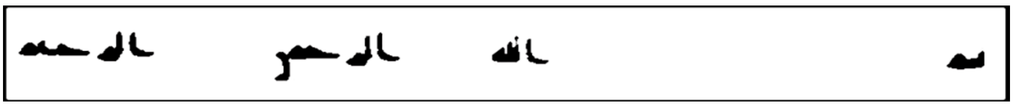
The result of the proposed approach is shown in Figure 8 (A-E). The top row shows the original Kufic manuscript images, while the bottom row displays the images processed using the proposed method.

A. Test Image (Kufic 4)

B. Grayscale Image

C. Wiener and Median Filtered Image

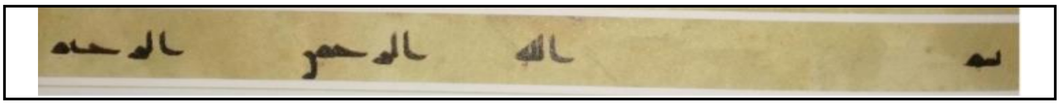
D. U-Net Based thresholding



E. Post threshold Filtering images

Fig 8 (A-E): Results of the proposed model in each step

Figure 9 (A-G) shows the result using the other threshold methods as well. It can be observed that the proposed method produces cleaner and more understandable images, with reduced noise artifacts and improved segmentation of the handwritten text.

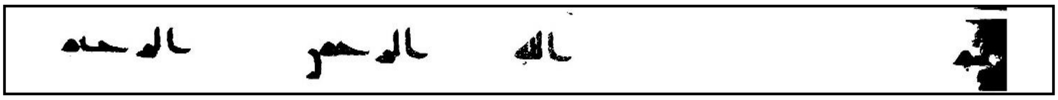


A. Original image

B. Grey scale

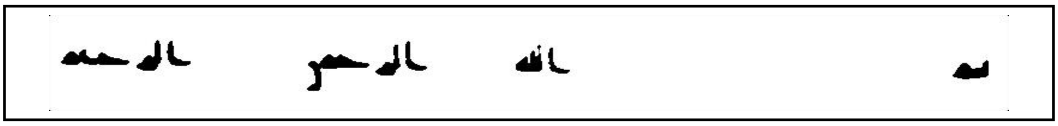
C. Weiner filter

D. Otsu's (Global Thresholding) Image



E. Bradley thresholding Image

F. Niblack thresholding Image



G. Sauvola thresholding Image

Fig 9 (A-G): Result using the other threshold methods

6. CONCLUSION

This study presents a novel pre-processing approach for recognizing Kufic manuscripts, significantly improving handwritten text segmentation accuracy. The proposed approach combines hybrid filtering, thresholding using the U-Net model, and additional noise removal techniques. Furthermore, the study provides a publicly available dataset curated for Quranic Kufic manuscripts, facilitating further research in this domain. Experimental results demonstrate that the approach effectively reduces noise, precisely segments handwritten text, and enhances the overall image quality of Kufic manuscript images. The evaluation performs better than traditional filtering techniques, with higher PSNR values and lower MSE values. The proposed approach contributes to advancing the Quranic Kufic Manuscripts Recognition System. It holds promise for digitizing historical documents, preserving cultural heritage, and future handwritten text recognition systems advancements.

While the proposed pre-processing approach for recognizing Kufic manuscripts has demonstrated promising results, several avenues for future work can further enhance the system and contribute to the field of handwritten text recognition.

a. ***Exploration of Advanced Deep Learning Models:*** *Deep learning models like U-Net have shown precise thresholding and segmentation effectiveness. Future work can explore and evaluate the performance of other advanced deep learning models, such as Mask R-CNN [26] or Transformer-based models, to improve the accuracy and efficiency of handwritten text recognition.*
b. ***Integration of Contextual Information****: Besides pre-processing, contextual information can improve the recognition system's performance. Future research can explore techniques to leverage surrounding textual and structural information, such as document layout analysis, language modelling, or semantic understanding, to enhance the recognition and interpretation of Kufic manuscripts.*
c. ***Robustness to Variations in Manuscript Quality:*** *Kufic manuscripts may exhibit variations in quality due to factors such as degradation, ink fading, or uneven background. Future work can focus on developing robust pre-processing techniques that can handle these variations effectively, ensuring accurate recognition even in challenging conditions.*
d. ***Dataset Expansion and Evaluation:*** *The proposed approach has been evaluated on a dataset of Kufic Quranic manuscript images. Future work can involve expanding the dataset by collecting more diverse and representative samples from different historical periods and regions. Additionally, comparative evaluations with other existing recognition systems or benchmark datasets can provide further insights into the performance and effectiveness of the proposed approach.*
e. ***Application to Other Historical Manuscripts:*** *While the proposed approach focuses on Kufic manuscripts, its principles and techniques can be extended to other types of historical manuscripts with similar characteristics. Future research can explore the application of the proposed pre-processing approach to different styles of handwritten text and languages, contributing to the preservation and digitization of a wider range of cultural heritage materials.*

DECLARATIONS

Ethical Approval
Not Applicable

Competing Interests
It is declared that the authors have no conflicts of interest.

Funding
Not Applicable

Availability of data and materials
Implementation of the model is done on local machine with following details:

Core Technology	:	Python 3.0
Environment	:	Google Colaboratory
Libraries	:	Keras and Tensorflow
GSimulator	:	Tesla T4 GPU
Windows	:	11 Professional

To enhance the pre-processing capabilities of the mode, authors have designed a Quranic Kufic Dataset uploaded on Kaggle (link to the dataset: https://www.kaggle.com/datasets/sheikhburhanulhaque/quranic-kufic-dataset).

7. REFERENCES

[1] Iqbal, A., Zafar, A.: Offline Handwritten Quranic Text Recognition: A Research Perspective. In: 2019 Amity International Conference on Artificial Intelligence (AICAI), pp. 125-128. IEEE (2019). https:/doi.org/10.1109/AICAI.2019.8701404

[2] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Lecture Notes in Computer Science, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

[3] Smith, R. G., Sridharan, S., & Wang, S. (2006). Handwritten Word Recognition Using Adaptive Noise Filtering. Proceedings of the 18th International Conference on Pattern Recognition, 1050-1053.

[4] Liu, Z., Liu, Z., Luo, J., & Wang, G. (2017). Contrast Enhancement for Handwritten Document Images Based on Adaptive Histogram Equalization. Proceedings of the 3rd International Conference on Industrial Engineering, 141-147

[5] Otsu, N.: A threshold selection method from gray level histograms. IEEE Trans. on Systems, Man and Cybernetics (1979) 9:62-66.

[6] Sauvola, J., Seppanen, T., .Haapakoski, S., Pietikainen, M.: Adaptive Document Binarization. 4th Int. Conf. On Document Analysis and Recognition, Ulm, Germany (1997), pp. 147-152.

[7] Aghdam, H. H., Faez, K., Gheissari, N., & Dengel, A. (2017). A Robust Text-Line Extraction Method for Historical Handwritten Documents. International Journal on Document Analysis and Recognition, 20(3), 129-142.

[8] Nikolaos, N., Dimitrios, V.: A Binarization Algorithm for Historical Manuscripts. 12th Wseas International Conference on Communications. pp. 41-51 (2008). ISSN: 1790-5117.

[9] Yahya, S. R.,Sheikh Abdullah, S. N. H.,Omar, K., Zakaria, M. S., Liong, C. Y. .: Review on image enhancement methods of old manuscript with the damaged background. International Journal on Electrical Engineering and Informatics. 2(1), 62 - 67 (2010).

[10] Al-Harahsheh, M., Al-Saleh, N., & Abuhaiba, I. (2018). Kufic Manuscript Text Line Segmentation Using Modified Projection Profile. International Journal of Computer Science and Network Security, 18(6), 227-234.

[11] Abuhaiba, I., Al-Harahsheh, M., & Al-Saleh, N. (2020). Enhancing the Quality of Kufic Manuscripts Using Image Processing Techniques. International Journal of Computer Science Issues, 17(2), 366-375.

[12] Almazroi, I., Alquran, M., Azeemi, A., & Shafait, F. (2019). Inpainting Degraded Manuscripts for Arabic Script Recognition. Proceedings of the 8th International Conference on Pattern Recognition Systems, 1-6.

[13] Zafar, A., Iqbal, A.: Application of Soft Computing Techniques in Machine Reading of Quranic Kufic Manuscripts. Journal of King Saud University - Computer and Information Sciences (2020). https://doi.org/10.1016/j.jksuci.2020.04.017.

[14] Fred, A. L., Kumar, S.N., Kumar, A. H., Daniel, A. V., Abisha, W.: Evaluation of local thresholding techniques in Palm-leaf Manuscript images. International Journal of Computer Sciences and Engineering. 6(4), 124-131 (2018). https:/doi.org/10.26438/ijcse/v6i4.124131.

[15] Bouillon, M., Ingold, R., Liwicki, M.: Grayification: A meaningful grayscale conversion to improve handwritten historical documents analysis, Pattern Recognition Letters, 2019.,

[16] Baig, A.,Al-Ma'adeed, S. A.S., Bouridane, A.,Cheriet, M.: Automatic segmentation and reconstruction of historical manuscripts in gradient domain. IET Image Process (2018). 12(4), pp. 502-512

[17] Xiong, W.,Zhou, L., L.,Yue, L., Li, L., Wang, S.: An enhanced binarization framework for degraded historical document images. EURASIP Journal on Image and Video Processing (2021). 2021:13, pp. 1-24. https://doi.org/10.1186/s13640-021-00556-4

[18] Zafar, A, Iqbal, A.: Machine Reading of Arabic Manuscripts using KNN and SVM Classifiers. In: Proceedings of the 14th INDIACom, 2020 7th International Conference on Computing for Sustainable Global Development (2020). pp. 83-87. https://doi.org/10.23919/INDIACom49435.2020. 9083696.

[19] Isa, I. S., Sulaiman, S. N., Mustapha, M., Darus, S.: Evaluating Denoising Performances of Fundamental Filters for T2- Weighted MRI Images. 19th International Conference on Knowledge Based and Intelligent Information and Engineering Systems (2015), pp. 760-768.

[20] Michalak, H., Okarma, K.: Improvement of Image Binarization Methods Using Image Preprocessing with Local Entropy Filtering for Alphanumerical Character Recognition Purposes. Entropy 2019, 21, 562. https://doi.org/10.3390/e21060562

[21] Niblack, W.: An Introduction to Digital Image Processing, Prentice Hall (1986), pp. 115-116.

[22] Michalak, H., Okarma, K.: Improvement of Image Binarization Methods Using Image Preprocessing with Local Entropy Filtering for Alphanumerical Character Recognition Purposes. Entropy 2019, 21, 562. https://doi.org/10.3390/e21060562

[23] Zhang, Y., Cheng, H. D., Huang, J. Tang, X..: An effective and objective criterion for evaluating the performance of denoising filters. Pattern Recognit (2012). 45(7): 2743–2757

[24] Westpha, F., Grahn, H., Lavesson, N.: Efficient document image binarization using heterogeneous computing and parameter tuning. International Journal on Document Analysis and Recognition (IJDAR) (2018), 21:41–58. https://doi.org/10.1007/s10032-017-0293-7

[25] Karim, A., Kadhm, M S.: An Efficient Image Thresholding Method for Arabic Handwriting Recognition System. Eng & Tech. Journal (2016). 34(1):26-34.

[26] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2961-2969). https://doi.org/10.1109/ICCV.2017.322.