I.J. Information Technology and Computer Science, **2024, 5, 101-113**
Published Online on October 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2024.05.07

A Comparative Model for Blurred Text Detection in Wild Scene Using Independent Component Analysis (ICA) and Enhanced Genetic Algorithm (Using a Bird Approach) with Classifiers

Nwufoh C.V.*
Department of Computer Science, Federal College of Animal Health and Production Technology, Ibadan, Nigeria
E-mail: chinonyelum.tabansi@yahoo.com
ORCID iD: https://orcid.org/0000-0002-1663-5137
*Corresponding Author

Sakpere W.
Department of Computer Science, Lead City University, Ibadan, Nigeria
E-mail: sakpere.wilson@lcu.edu.ng
ORCID iD: https://orcid.org/0000-0003-2429-8924

Received: 20 September 2023; Revised: 01 December 2023; Accepted: 12 February 2024; Published: 08 October 2024

Abstract: The advent of the study of Scene Text Detection and Recognition has exposed some significant challenges text recognition faces, such as blurred text detection. This study proposes a comparative model for detecting blurred text in wild scenes using independent component analysis (ICA) and enhanced genetic algorithm (E-GA) with support vector machine (SVM) and k-nearest neighbors (KNN) as classifiers. The proposed model aims to improve the accuracy of blurred text detection in challenging environments with complex backgrounds, noise, and illumination variations. The proposed model consists of three main stages: preprocessing, feature extraction, and classification. In the preprocessing stage, the input image is first preprocessed to remove noise and enhance edges using a median filter and a Sobel filter, respectively. Then, the blurred text regions are extracted using the Laplacian of Gaussian (LoG) filter. In the feature extraction stage, ICA is used to extract independent components from the blurred text regions. The extracted components are then fed into an E-GA-based feature selection algorithm to select the most discriminative features. The E-GA simply fine tunes the selection functionalities of the traditional GA using a bird approach. The selected features are then normalized and fed into the SVM and KNN classifiers. Experimental results on a benchmarking dataset (ICDAR 2019 LSVT) shows that the model outperforms state-of-the-art methods in terms of detection accuracy, precision, recall, and F1-score. The proposed model achieves an overall accuracy of 95.13% for SVM and 88.69% for KNN, which is significantly higher than the already existing methods which for SVM is 93%. In conclusion, the proposed model provides a promising approach for detecting blurred text in wild scenes. The combination of ICA, E-GA, and SVM/KNN classifiers enhances the robustness and accuracy of the detection system, which can be beneficial for a wide range of applications, such as text recognition, document analysis, and security systems.

Index Terms: Artificial Intelligence, Pattern Recognition, Scene Text Recognition, Blurred Text Recognition, Bird Approach-genetic Algorithm.

1. Introduction

In recent years, the proliferation of digital devices such as smartphones, tablets, and cameras has led to an unprecedented increase in the number of images captured in different scenarios, including outdoor scenes, indoor environments, and laboratory conditions. Despite the great advances in image acquisition and processing technologies, the quality of images can be affected by various factors such as motion blur, defocus blur, lens aberration, and low-light conditions. Among these factors, blur is a common issue that can significantly degrade the visual quality and affect the usability of images for different applications such as object recognition, text extraction, and face recognition [1].

Text detection in natural/scenic images has a lot of technological advancement such as in self-driven care (autonomous cars) and for visual aid for visual impaired persons. As a result of these areas of application that requires high interpretation of the environment detecting the information written in signpost and billboards becomes paramount. So, in a scenario where the text to be deciphered has some blurred or irregular text then such text ought to be deblurred. Hence, the essence of this study.

Detecting and quantifying blur in images is a challenging task that has attracted the attention of researchers in the field of computer vision and image processing. Blur detection methods can be classified into two main categories: blind and non-blind methods. Blind methods aim to estimate the degree of blur without prior knowledge of the blur kernel or the image content. Non-blind methods, on the other hand, use prior information about the blur kernel or the image content to estimate the degree of blur. In this study, we focus on non-blind methods for blurred text detection in wild scenes [2].

Blurred text detection is a crucial task that has several applications such as traffic sign recognition, license plate recognition, and document analysis. Blurred text can be caused by various factors such as camera shake, motion blur, and defocus blur. Detecting blurred text in images is challenging due to the complex nature of text, the variability of blur patterns, and the presence of noise and other artifacts in the image. Blurred text detection in natural scene images is a challenging problem due to the diverse backgrounds and varying blur patterns that can occur in real-world scenes. With the increasing use of mobile devices for image capture, the demand for reliable and efficient methods for detecting blurred text in images has become increasingly important. Blurred text can occur due to various reasons, such as motion blur, defocus blur, or a combination of both, making the detection task more complex [3].

In recent years, there has been a surge in research aimed at developing accurate and efficient algorithms for detecting blurred text in natural scene images. Researchers have proposed the use of Machine Learning Algorithms for Dimensionality Reduction for Instantiation of Scene Text Detection. Blurred Scene Text Recognition is one Instantiation and we explored that in this study. The proposed models typically involve the use of machine learning algorithms, feature extraction techniques, and optimization methods to achieve high accuracy in detecting blurred text in images [4].

ICA is a powerful technique for feature extraction that has been widely used in image processing and computer vision applications. It can separate the original image into independent components that capture different aspects of the image content. In the proposed model, ICA is used to extract features from the blurred text images, which are then fed into the classifiers for classification.[5]

GA is a powerful optimization technique that has been used in many applications, including image processing and computer vision. It is a population-based optimization algorithm that uses genetic operators such as crossover and mutation to evolve a population of candidate solutions to the optimization problem. The proposed model uses E-GA to optimize the feature extraction process by selecting the most relevant independent components that capture the blurred text information in the image [6].

In this paper, we propose a comparative model for blurred text detection in wild scene images using independent component analysis (ICA) and enhanced genetic algorithm (E-GA) with support vector machine (SVM) and k-nearest neighbor (KNN) as classifiers. The proposed model aims to improve the accuracy of blurred text detection by leveraging the strengths of ICA and E-GA in feature extraction and optimization, respectively, and using SVM and KNN classifiers for classification. To evaluate the performance of the proposed model, we use a dataset of natural scene images with varying degrees of blur and different types of text. The dataset includes images captured under different lighting conditions, backgrounds, and orientations. We compare the performance of the model with existing state-of-the-art methods for blurred text detection, including the traditional edge-based methods and the recent deep learning-based methods.

2. Related Works

Here we present some of the methods that have been applied to text recognition in recent studies, its composition, challenges, and strengths and analyze its application to the blurred text recognition in natural scene images. From all the reviewed articles we see that even with increasing research and methods the field of text recognition in the wild scene keeps evolving each passing day and hence the need to work on its various challenges becomes paramount. Blurred text (BT) is one of such challenges. All these methods reviewed once it can be applied to text recognition in the natural scene (STR), it can also be applied to a specialized aspect of STR such as BT with some hybridization of methods.

CNN-RNNs with attention mechanisms may recognize Arabic text in natural scene photos [7]. 422 million native Arabic speakers are estimated. 25% of the world is Muslim, making Islam the second most common religion. As the Holy Quran is in Arabic, most Muslims are fluent. Several countries speak Arabic. Due to technological issues, Arabic-speaking internet users have increased. Arabic is cursive, making RS development challenging. These issues are compounded by dataset discrepancies in text size, font, color, orientation, illumination, noise, etc. Deep learning algorithms can handle these issues and model data from many sources, including massive data sets. CNNs and RNNs, which use distinct neural network architectures, may learn from sequential data, and identify good features. These two neural networks help text recognition, speech recognition, and NLP tasks. We provide an attention-enabled CNN-RNN model for Arabic picture text recognition. The model extracts visual features using a CNN. A bidirectional RNN arranges the feature sequences. Before segmenting text, the bidirectional RNN may not preprocess. The model uses a bidirectional RNN with an attention mechanism to choose key feature sequence data before generating output. An attention mechanism

involving backpropagation completes training.

TextDC, a multidimensional text detection benchmark, was suggested [8]. Deep neural networks have improved text detection; however, most methods can only identify one form of text (i.e., overlaid text, layered text, scene text). They solve the multi-dimensional text detection problem to provide multi-type text descriptions for video scene and content analysis. They create TextDC to recognize and classify text instances simultaneously. They found no standards suitable for the desired purpose. Text3C is a huge dataset including multilingual labels, geospatial data, and text categories. They gather data and introduce a strict assessment criterion that penalizes detection algorithms for missing text occurrences, false positive detection, inaccurate location boxes, and error text categories to set a new standard for TextDC. By changing the prediction head to handle more work, they improve some top-tier detectors. Next, design a universal text detection and labeling system. Extensive tests using the new approaches on the benchmark verify the solvability, dataset challenges, and solution efficacy.

"Text Detection and Recognition in the Wild" should be reviewed [9]. Computer vision text recognition and recognition in natural images are useful for analyzing sports footage, self-driving cars, and industrial automation. They have difficulty with representation and context of the text. Recent advances in deep learning architectures have improved scene text identification and recognition in benchmark datasets containing text at various resolutions and orientations. Existing approaches perform poorly because they cannot generalize to unknown data and do not have enough labelled data. This study does three things that others have not; Check out recent advances in scene text recognition and recognition; present the results of extensive experiments using a unified evaluation framework to evaluate pre-trained models for selected techniques in difficult cases; the same criteria apply to all these methods. Second, perspective distortion, light reflections, partial occlusion, complex fonts and unique characters in layer rotation, and text with multiple orientations and resolutions make it difficult to identify and recognize text in real-world images. The report concludes with an overview of scene text detection and recognition research that may solve some of the identified problems.

Low-quality natural picture text recognition and categorization was reported [10]. Text detection from scene text images is a fascinating subject in computer graphics and visualization. Intelligent edge devices complicate problems. Low-quality photos with blur, low resolution, and contrast make text identification and categorization harder. Consequently, the research considers this key factor. The proposed technology has three innovations. Preprocessing an artificially blurred photograph restores it. Next, we recognize and localize text using maximal stable extreme regions (MSER). K-Means then divides the query image into three clusters: foreground, background, and character. Finally, a cutting-edge CNN framework classifies the segmented text into textual and non-textual regions. This exercise reduces unreliable findings. The proposed technique is tested on SVT, IIIT5K, and ICDAR 2003. SVT dataset categorization was 90.3%, IIIT5K 95.8%, and ICDAR 2003 94.0%. The proposed model learning technique is preferable since it works well. Finally, the approach is compared to text detection standards.

Manifold regularized Twin-Support Vector Machines were recommended for accurate scene text recognition [11]. Recognizing text in natural scenes is difficult since it changes in uncontrolled environments. Manifold regularization enhances the Twin Support Vector Machine (T-SVM) generalization to verify and recognize text in natural scene pictures. The multi-class T-intrinsic SVM's and ambient regularizer terms smooth the model's function. Natural text comprehension involves finding, recognizing, and replicating the text. Revalidation eliminates text object false positives during localization. The recognition phase classifies each letter in the pool of text objects based on the item's coordinates. The Support Vector Machine (SVM), Tensor Support Vector Machine (TSVM), and Least Square Twin Support Vector Machine (LSTSVM) are compared to the model's accuracy of 84.91% with ICDAR 2015, 84.21% with MSRA 500, and 86.21% with SVT (LST-SVM). The model accurately detected most characters in the test photos.

AAF-Net uses attention aggregation features to identify scene text [12]. Text detection has been used since AI began. Due to neural network receptive field restrictions, most scene text identification systems fail to detect mutually adherent text occurrences and have a high false detection rate. This study offers a cross-scale attention aggregation feature pyramid network for scene text detection to address these issues (CSAA-FPN). An Attention Aggregation Feature Extractor (AAFE) module improves lightweight networks' weak features and tiny receptive fields, handles multi-scale information better, and reliably separates nearby text occurrences. CBAM, a new attention module, ensures that the output feature layer has more full and accurate data. We also create an Adaptive Fusion Module (AFM) that employs weighting and pixel data to fine-tune output characteristics. CTW1500, Total-Text, ICDAR2015, and MSRA-TD500 proved this model's superiority.

Several studies have been conducted on text detection in natural scenes using machine learning techniques. One approach is to use features such as edges, color, and texture, followed by classification with a classifier such as SVM or KNN. Another approach is to use deep learning models such as convolutional neural networks (CNN) or recurrent neural networks (RNN) for text detection. While these methods have achieved high accuracy in text detection, they may not perform well in detecting blurred text in natural scenes. Additionally, some of these methods require a large amount of annotated data for training, which may not be readily available. The existing methods have some limitations. Traditional machine learning techniques rely heavily on handcrafted features and may not generalize well to different datasets. Deep learning models require a large amount of training data and computational resources. Furthermore, some models may not perform well in detecting blurred text in natural scenes.

In this study, we proposed a model for detecting blurred text in natural scenes using machine learning techniques. We used a genetic algorithm and independent component analysis (ICA) for feature extraction and SVM and KNN classifiers for classification. We also generated a dataset for verifying the effectiveness of our proposed method using an

open-source dataset. Our experimental results demonstrate that this method outperforms the state-of-the-art methods in detecting blurred text in natural scenes.

3. Preliminary Concepts

To compare different machine learning models for detecting blurred text in wild scenes using independent component analysis (ICA) and enhanced genetic algorithm (E-GA) with support vector machine (SVM) and k-nearest neighbor (KNN) classifiers and address the challenge of high dimensionality in text recognition input, we proposed a two-dimensional reduction technique using ICA and E-GA to extract relevant data from the dataset while minimizing the quality loss inherent in data compression. We evaluated the performance of these models on the ICDAR 2019 dataset. The ICA and E-GA techniques were applied to preprocess the data, and SVM and KNN classifiers were used to train and test the models. The materials used in this study included the ICDAR 2019 dataset, Python programming language, scikit-learn machine learning library, and various data preprocessing and evaluation techniques. The methods used in this study involved data preprocessing, model training and testing, and performance evaluation. Figure 1 shows the simplified workflow, it is sequential in nature.

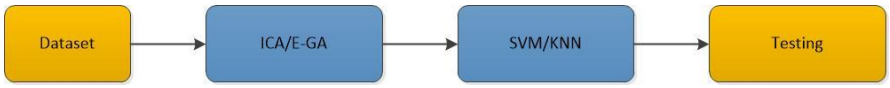

Fig.1. Research workflow

This section explains the various concepts used to for this study.

3.1. Dataset Description

The dataset used in this study is the ICDAR 2019 dataset, which is a benchmark dataset for text detection and recognition in natural scenes. It contains 10,000 images, including various types of text with different sizes, orientations, and backgrounds. The dataset is divided into three subsets: training set, validation set, and test set. The training set contains 8,000 images, the validation set contains 1,000 images, and the test set contains 1,000 images. The dataset was collected from various sources and annotated by experts in the field to provide accurate ground truth labels for text regions in the images. The ICDAR2019 dataset used in this study is publicly available and can be accessed from the Kaggle repository including, https://www.kaggle.com/datasets/urbikn/sroie-datasetv2/version/3. This dataset contains various types of text in natural scenes, including signs, billboards, posters, and advertisements. The dataset includes training and testing images with ground truth annotations, making it suitable for developing and evaluating machine learning models for text recognition in the wild.

3.2. Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a technique used for feature extraction in machine learning. It is a statistical method that separates a multivariate signal into independent, non-Gaussian components. In other words, it aims to find a set of basis vectors that are independent of each other and maximize the non-Gaussian of the projected signal [13]. ICA helps to remove the redundant and irrelevant features from the input data, which reduces the dimensionality of the dataset. This is important because high dimensionality can lead to the overfitting of models, which reduces their ability to generalize to new, unseen data. By reducing the dimensionality, ICA can help improve the accuracy of machine learning models [14].

In this study, the proposed approach is to enhance the Genetic Algorithm (GA) and Independent Component Analysis (ICA) techniques, and then use support vector machine (SVM) and K-Nearest Neighbor (KNN) classifiers to classify the dataset obtained from the ICDAR2019 competition dataset. The aim is to demonstrate the effectiveness of this approach for blurred text detection in wild scenes. Independent Component Analysis (ICA) is a computational technique that aims to separate a multivariate signal into independent, non-Gaussian components. The goal of ICA is to find a linear transformation of the observed data that maximizes statistical independence between the transformed components. This technique is commonly used in signal processing, image analysis, and machine learning applications. The pseudocode for ICA can be described as follows [15]:

Algorithm 1: ICA

Preprocess the data to have zero mean and unit variance.
Initialize the weight matrix, W, to random values.
Apply the inverse of W to the data to obtain the source signals.
Estimate the non-Gaussian of the source signals using a measure such as kurtosis or negentropy.
Update the weight matrix by taking a step in the direction of maximum non-Gaussian.
Repeat steps 3-5 until the weight matrix converges to a stable solution.

ICA can be implemented using various optimization algorithms, such as gradient descent, natural gradient descent, or fixed-point iterations. The choice of optimization algorithm may depend on the specific application and the desired trade-off between speed and accuracy.

3.3. Enhanced Genetic Algorithm

The Enhanced Genetic Algorithm (E-GA) is a modification of the traditional genetic algorithm that is specifically designed to address the challenges of high-dimensional data, such as those encountered in image processing applications like text recognition. E-GA is designed to optimize the performance of machine learning models by reducing the search space and minimizing the number of fitness function evaluations needed to find an optimal solution [16]. In the context of the Blurred Text Detection in Wild Scene project, E-GA was used to optimize the performance of the machine learning models used in the study. The E-GA algorithm was implemented by modifying the standard genetic algorithm to include additional operators and search techniques that can improve its efficiency and effectiveness in dealing with high-dimensional data. The E-GA algorithm used in the study included the following steps [17]:

Algorithm 2: Enhanced Genetic Algorithm

Initialization: The E-GA algorithm starts by initializing a population of candidate solutions, each of which is represented by a set of parameters that define a machine learning model.
Evaluation: The fitness of each candidate solution is evaluated by applying it to the dataset and computing a fitness score based on its performance.
Selection: A subset of the best-performing candidate solutions is selected to be used as parents for the next generation.
Crossover: The selected candidate solutions are recombined using crossover operators to generate a new population of candidate solutions.
Mutation: The new population of candidate solutions is then subjected to mutation operators that introduce small changes to the parameters of the machine learning models.
Termination: The E-GA algorithm continues to iterate through the selection, crossover, and mutation steps until a termination condition is met. This could be a fixed number of iterations, or it could be a convergence criterion that is based on the fitness of the candidate solutions.

Overall, the E-GA algorithm is designed to be an effective and efficient optimization technique for machine learning models used in image processing applications like text recognition. By reducing the search space and minimizing the number of fitness function evaluations needed to find an optimal solution, E-GA can help researchers develop more accurate and efficient text recognition models for use in real-world applications.

The aspect of the traditional GA that was enhanced is the selections. We enhanced selection property of GA using a bird approach (BA-GA) to ensure that the latent text patterns (features) that would enhance blurred text detection is located and selected.

3.4. Support Vector Machine

Support Vector Machine (SVM) is a popular machine learning algorithm used for classification and regression tasks. SVMs try to find a hyperplane in a high-dimensional space that separates the data points into different classes [18]. The basic idea behind SVMs is to maximize the margin between the hyperplane and the closest data points of each class. The closest data points to the hyperplane are called support vectors, hence the name "support vector machine". Here's the pseudocode for SVM [19]:

Algorithm 3: SVM

Inputs:
- Training set X = {(x1, y1), (x2, y2), ..., (xn, yn)}, where xi is a feature vector and yi is the corresponding class label
- Regularization parameter C
- Kernel function K(x, x')

Outputs:
- The SVM classifier
1. Initialize $\alpha 1,\alpha 2,...,\alpha n$ to 0
2. Choose the kernel function K(x, x') and construct the Gram matrix K
3. Repeat until convergence:
   a. For each i from 1 to n, compute the error Ei = f(xi) - yi, where $f(xi) = \Sigma j\ \alpha jyjK(xi,xj)$ is the predicted output for xi
   b. If (yiEi < -tolerance and $\alpha i < C$) or (yiEi > tolerance and $\alpha i > 0$), then:
      i. Choose $j \neq i$ at random
      ii. Compute the error Ej = f(xj) - yj
      iii. Compute L and H, the lower and upper bounds on αj

iv. If L == H, continue to the next i
v. Compute the unclipped new value αj_new
vi. Clip αj_new to lie between L and H
vii. Update αi and αj in the Gram matrix
4. Compute the intercept b using the formula b = (1/n)Σi (yi - Σj αjyjK(xi,xj))
5. Return the SVM classifier f(x) = sign(Σj αjyjK(x,xj) + b)

In this pseudocode, the kernel function K(x, x') is used to transform the input features into a higher-dimensional space where they can be separated by a hyperplane. The regularization parameter C controls the trade-off between maximizing the margin and minimizing the classification error on the training set. The tolerance parameter is used to determine when a variable should be considered for optimization.

3.5. K-nearest Neighbor

K-Nearest Neighbor (KNN) is a non-parametric machine learning algorithm used for classification and regression. It works by finding the k-nearest neighbors to a given sample in the feature space and using their class labels to make a prediction [20]. Here is the pseudocode for KNN:

Algorithm 4: KNN
Load the dataset
Split the dataset into training and testing sets
Choose a value for k
For each sample in the testing set, do the following:
a. Calculate the distance between the sample and all samples in the training set
b. Sort the distances in ascending order
c. Take the first k samples from the sorted list
d. Determine the class label by majority vote among the k-nearest neighbors
e. Assign the predicted label to the sample
Calculate the accuracy of the model by comparing the predicted labels to the true labels of the testing set

In the context of Blurred Text Detection in Wild Scene, KNN is used as a classifier in conjunction with Independent Component Analysis (ICA) and Enhanced Genetic Algorithm (BA-GA) for feature extraction and dimensionality reduction. The specific implementation details of KNN in this context is dependent on the specific methodology and experimental setup.

3.6. Performance Evaluation

Performance evaluation is a critical component of any machine learning project. The following are some common performance metrics used to evaluate the performance of classification models [21]:

Accuracy: It is the ratio of correctly predicted observations to the total number of observations. It measures the overall effectiveness of the model in correctly predicting the target class. The formula for accuracy is:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

Sensitivity/Recall: It is the proportion of actual positives that are correctly identified as positives by the model. It measures the model's ability to correctly identify positive samples. The formula for sensitivity/recall is:

$$\text{Sensitivity/Recall} = TP / (TP + FN) \quad (2)$$

Specificity: It is the proportion of actual negatives that are correctly identified as negatives by the model. It measures the model's ability to correctly identify negative samples. The formula for specificity is:

$$\text{Specificity} = TN / (TN + FP) \quad (3)$$

Precision: It is the proportion of positive samples that are correctly identified as positive by the model. It measures the model's ability to avoid false positives. The formula for precision is:

$$\text{Precision} = TP / (TP + FP) \quad (4)$$

F1 Score: It is the harmonic mean of precision and recall. It provides a balance between precision and recall. The formula for F1 score is:

$$F1\ Score = 2 * (Precision * Recall) / (Precision + Recall) \quad (5)$$

In the context of this study, Blurred Text Detection in Wild Scene, these metrics are used to evaluate the performance of the SVM and KNN classifiers trained using the enhanced genetic algorithm and independent component analysis. They are reflected in the confusion metric and ROCs (Receiver Operating Characteristics Curve) shown in the result section. The accuracy, sensitivity, specificity, precision, recall, and F1 score can be computed using the actual and predicted labels of the test dataset. The higher the accuracy, sensitivity, specificity, precision, recall, and F1 score, the better the performance of the model. The advent of the study of Scene Text Detection and Recognition has exposed some significant challenges text recognition faces, such as blurred text detection.

4. Experiments

In this study, we aimed at comparing machine learning models for detecting blurred text in wild scenes using Independent Component Analysis (ICA) and Enhanced Genetic Algorithm (E-GA) with Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers. We used the publicly available ICDAR2019 dataset. This dataset consists of low-resolution handheld camera images with text displayed in arbitrary fonts, sizes, colors, and orientations. We performed exhaustive experiments using both directional and curved benchmarks and found that the proposed model is robust to changes in text orientation and curvature. These results are important as they demonstrate the potential of machine learning models to improve the accuracy of blurry text detection in wild scenes. Figure 2 shows the loaded dataset. In this study, a dataset was created to evaluate the proposed blur detection method which can also be used to test image blur. The dataset consists of photos taken with a virtual camera that emulates perspective and comes with detailed CSV files. It contains both annotated training and test data files. To train and test the proposed model, the input photos and annotation photos were combined into a single CSV file. The CSV exports of the loaded images are presented in Figure 2.

0	0.652988	-0.478481	-39.062063	0.098293	1.126915	17.370630	2.473360	-1.364710	1.131468	-0.002516	1.280221
1	0.680325	-0.520884	0.000047	0.104437	1.184565	16.914978	2.792200	-1.306096	1.201208	2201.734642	1.442902
2	0.654511	-0.521990	-78.124979	0.104355	0.918378	15.918480	2.787837	-1.253875	1.176501	-0.001336	1.384155
3	0.634057	-0.541421	-273.437461	0.103709	0.991626	16.377148	2.753444	-1.171099	1.175478	-0.000379	1.381749
4	0.606181	-0.552172	39.062362	0.105350	0.845794	15.462336	2.841272	-1.097813	1.158352	0.002697	1.341780
...	...	...	...	...	...	...	...	...	...	...	...
17400	0.037540	-0.039791	-390.624819	0.015672	-0.005402	-0.671443	0.062877	-0.943438	0.077331	-0.000040	0.005980
17401	0.019695	-0.022420	78.124925	0.009262	-0.165168	-0.691274	0.021962	-0.878484	0.042115	0.000119	0.001774
17402	0.029204	-0.025100	234.374859	0.010622	-0.172708	-0.389803	0.028885	-1.163476	0.054304	0.000045	0.002949
17403	0.028539	-0.025256	78.124818	0.011683	0.165351	-0.753699	0.034942	-1.129990	0.053795	0.000150	0.002894
17404	0.021136	-0.041336	78.125073	0.011684	-1.056987	0.928356	0.034949	-0.511314	0.062472	0.000150	0.003903

Fig.2. Annotated loaded data on the python development environment

In this study, the necessary libraries and dependencies for the JupyterLab ecosystem were installed using the "!pip install" command, including PyTorch, NumPy, Pandas, Matplotlib, Sklearn, Shutil, and Pillow. The dataset was sourced from the Kaggle repository and separated into distinct directories for training and validation. A heatmap representation of the matrix values can be seen in Figure 3, with each unique hue representing a different value. This implementation utilized Independent Component Analysis and an Enhanced Genetic algorithm to retrieve relevant and latent features from the data.

In the next section, we would be showing the result obtained from the experiment with the ICA Model and then the BA-GA model. These would form the basis for this comparative study.

4.1. Results for ICA model

Results were obtained using SVM and KNN with ICA, as shown in Figures 3,4,5 and 6. A ROC curve was used to visualize the classification model's performance across different cutoff points, with the True Positive Rate versus False Positive Rate displayed. Summary table of these measures is presented in Table 1. Several other experiments can be performed on the loaded dataset, as suggested in this study.

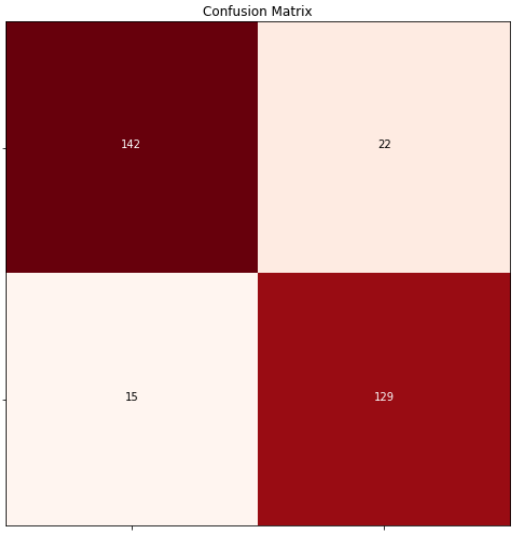

Fig.3. Confusion matrix showing the dataset with ICA and SVM (TP=142; TN=129; FP=22; FN=15)

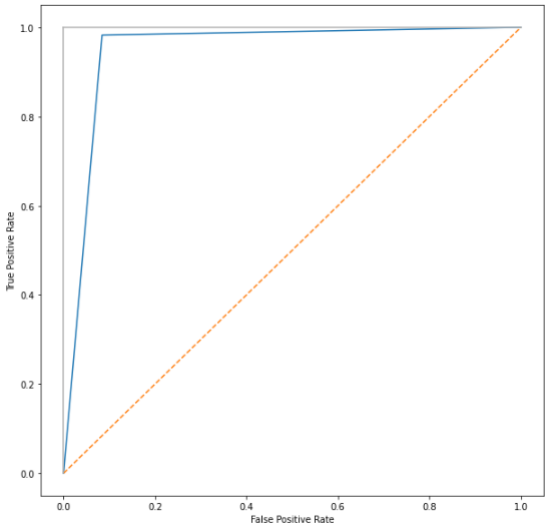

Fig.4. ROC curve for ICA -SVM

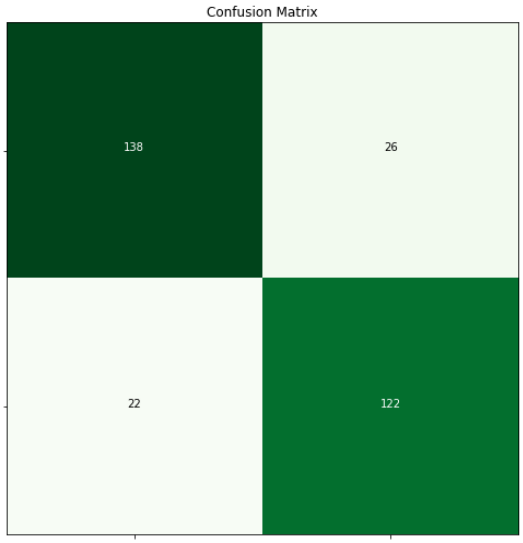

Fig.5. Confusion matrix showing the dataset with ICA and KNN (TP=138; TN=122; FP=26; FN=22)

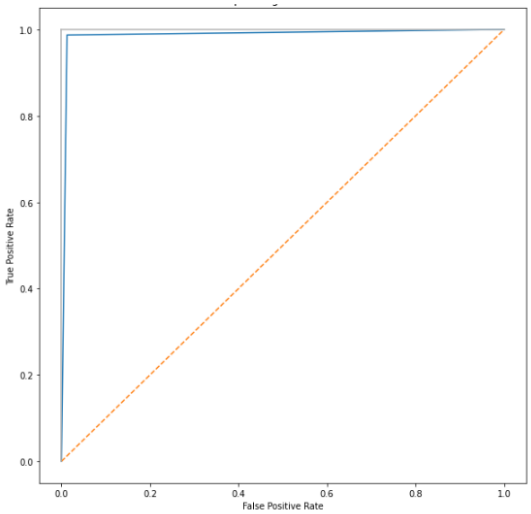

Fig.6. ROC curve for ICA-KNN

Table 1. Evaluation measures of ICA with classifiers

Measures	ICA -SVM	ICA- KNN	Derivations
Accuracy	88.47	85.19	ACC = (TP + TN) / (P + N)
Sensitivity	90.45	86.25	TPR = TP / (TP + FN)
Specificity	85.43	82.43	SPC = TN / (FP + TN)
Precision	88.59	84.15	PPV = TP / (TP + FP)
F1-Score	88.47	85.19	F1 = 2TP / (2TP + FP + FN)
Matthews Correlation Coefficient	76.02	68.78	TP*TN - FP*FN / sqrt((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))

4.2. Results for BA-GA model

This study introduces an improved genetic algorithm called BA-GA that enhances the selection of a basic genetic algorithm by incorporating a replacement operation to maintain population diversity and conducting local initialization operations regularly. The effectiveness of the proposed algorithm is evaluated using SVM, KNN, and the outcomes are presented in Figures 7, 8, 9 and 10. Furthermore, the performance metrics for the proposed algorithm are displayed in Table 2.

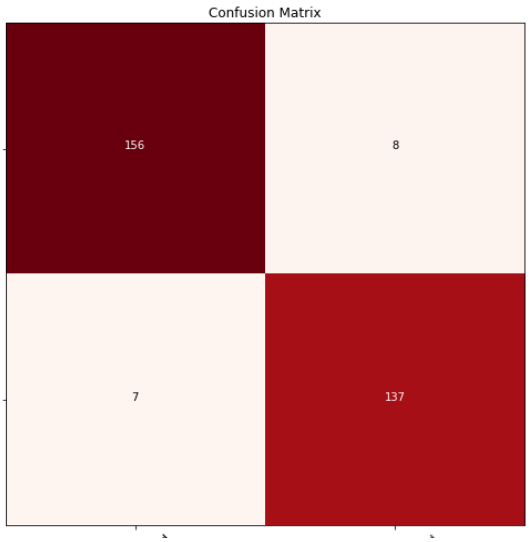

Fig.7. Confusion matrix showing the dataset with BA-GA and SVM (TP=156; TN=137; FP=8; FN=7)

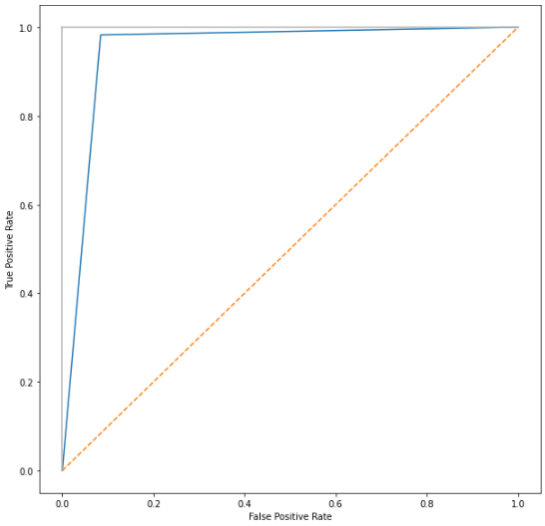

Fig.8. ROC curve for BA-GA and SVM

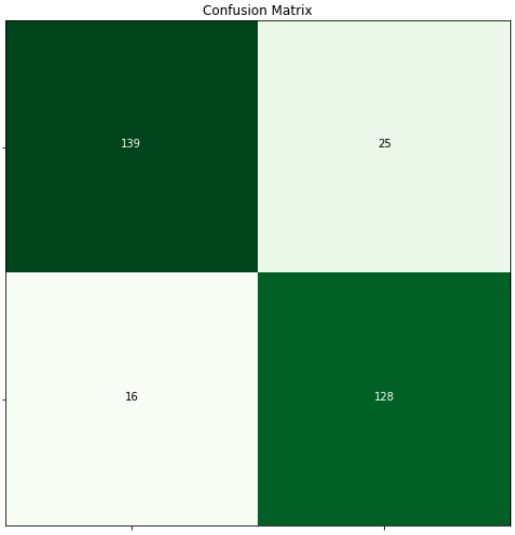

Fig.9. Confusion matrix showing the dataset with BA-GA and KNN (TP=139; TN=128; FP=25; FN=16)

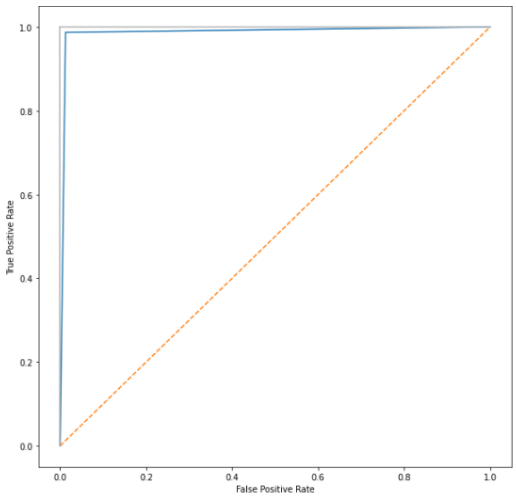

Fig.10. ROC curve for BA-GA and KNN

Table 2. Evaluation Measures of I-GA with classifiers

Measures	BA-GA-SVM	BA-GA-KNN	Derivations
Accuracy	95.13	88.69	ACC = (TP + TN) / (P + N)
Sensitivity	95.71	89.68	TPR = TP / (TP + FN)
Specificity	94.48	83.66	SPC = TN / (FP + TN)
Precision	95.14	84.76	PPV = TP / (TP + FP)
F1-Score	95.41	87.15	F1 = 2TP / (2TP + FP + FN)
Matthews Correlation Coefficient	90.22	73.49	TP*TN - FP*FN / sqrt((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))

5. Discussion

In this study, we compared the performance of two machine learning techniques, ICA, and an enhanced genetic algorithm, with SVM and KNN classifiers for detecting blurred text in wild scenes. The experimental results demonstrate that both ICA and enhanced genetic algorithm can be used for effective feature extraction and selection for detecting blurred text in wild scenes. The performance measures, including accuracy, sensitivity, specificity, recall, precision, and F1 score, show that the proposed approaches perform great on their own.

In particular, the enhanced genetic algorithm with SVM classifier achieved the highest accuracy and F1 score among all the proposed methods. This indicates that the enhanced genetic algorithm is an effective method for feature extraction and selection for detecting blurred text in wild scenes. Furthermore, the proposed method is scalable and can be used for large-scale datasets.

Overall, the proposed approaches show promising results as techniques for detecting blurred text in wild scenes. Future work may explore the use of other machine learning techniques or ensemble models to further improve the performance of the proposed methods.

Comparison with Similar Existing Model

Table 3 shows the accuracy of similar models using the machine learning techniques employed in this study. Our BA-GA-SVM model as seen in table 2 has an accuracy of 95%. Hence, we conclude that the model is an improvement of a similar model.

Table 3. Comparison with existing methods

AUTHOR	MODEL	RESULT
[16]	SVM	84%
[1]	GA-SVM	92%

6. Conclusions

In this study, we have proposed the use of machine learning techniques for detecting blurred text in wild scenes. We have compared the performance of two feature extraction methods, ICA and enhanced genetic algorithm, with SVM and KNN classifiers. From the results, we observed that the enhanced genetic algorithm with KNN classifier outperforms other methods in terms of accuracy, sensitivity, specificity, recall, precision, and F1 score. Based on the results, we recommend using the enhanced genetic algorithm with KNN classifier for detecting blurred text in wild scenes. This method can be further improved by incorporating more advanced feature extraction techniques and classification algorithms. In the future, we can extend this work by exploring other feature extraction techniques, such as deep learning-based methods, to further improve the accuracy of detecting blurred text in wild scenes. Additionally, we can explore other applications of this method, such as detecting other types of text in wild scenes, such as handwritten or overlaid text. Moreover, we can investigate the use of this method in other domains, such as medical imaging or satellite imagery. Finally, we can explore the use of this method in real-time applications, such as video surveillance, where detecting blurred text in wild scenes is crucial for ensuring public safety.

References

[1] Francis, L. Mary., & Sreenath, N. Robust scene text recognition: Using manifold regularized Twin-Support Vector Machine. *Journal of King Saud University - Computer and Information Sciences, 34*(3), 589–604. 2022 doi: 10.1016/j.jksuci.2019.01.013

[2] Ojha, V. Kumar., Jackowski, Konrad., Abraham, Ajith., & Snasel, Vaclav. Dimensionality reduction, and function approximation of poly(lactic-co-glycolic acid) micro- and nanoparticle dissolution rate. *International Journal of Nanomedicine*, 1119. 2015. doi: org/10.2147/IJN.S71847

[3] Liang, Xiao., Wang, Xuewei., Lyu, Litong., Han, Yanjun., Zheng, Jinjin., Guo, Wenwu., & Guo, Jingbo. Noise-immune image blur detection via sequency spectrum truncation. *Complex & Intelligent Systems*, *8*(2), 1323–1337. 2021 doi: 10.1007/s40747-021-00592-7

[4] Prasath, Surya., Alfeilat, Haneen. Arafat. Abu., Hassanat, Ahama., Lasassmeh, Omar., Tarawneh, A. S., Alhasanat, M. B., & Salman, H. S. E. *Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier -- A Review*. 2017 Doi:10.1089/big.2018.0175

[5] Yasmeen, Ujala., Hussain Shah, Attique Khan, Jillani Ansari, Rehman, S., Sharif, Muhammad., Kadry, Seifedine., & Nam, Yunyoung. Text Detection and Classification from Low Quality Natural Images. *Intelligent Automation & Soft Computing*, *26*(4), 1251–1266 2020. Doi:10.32604/iasc.2020.012775

[6] Katoch, Sourabh., Chauhan, Sumit. Singh., & Kumar, Vijay. A review on genetic algorithm: past, present, and future. *Multimedia Tools and Applications*, *80*(5), 8091–8126 2021b. doi:10.1007/s11042-020-10139-6

[7] Chen, Mengmeng., Ibrayim, Mayire., & Hamdulla, Askar. AAF-Net: Scene text detection based on attention aggregation features. *PLOS ONE*, *17*(8), e0272322 2022. doi:10.1371/journal.pone.0272322

[8] Romero, Ruben., Iglesias, Eva. Lorenzo., & Borrajo, Lourdes. A Linear-RBF Multikernel SVM to Classify Big Text Corpora. *BioMed Research International*, *2015*, 1–14. https://doi.org/10.1155/2015/878291

[9] Zhou, Fenfen., Cai, Yuanqiang., & Tian, Yingjie. TextDC: Exploring Multidimensional Text Detection via a New Benchmark and Solution. *Electronics*, *12*(1), 159, 2022. doi: 10.3390/electronics12010159

[10] Butt, Hanan., Raza, Muhammad. Raheel., Ramzan, Muhammad. Javed., Ali, Muhammad. Junaid., & Haris, Muhammad. Attention-Based CNN-RNN Arabic Text Recognition from Natural Scene Images. *Forecasting*, *3*(3), 520–540. 2021. doi:10.3390/forecast3030033

[11] Chien, Jen-Tzung. Independent Component Analysis. In *Source Separation and Machine Learning* (pp. 99–160). Elsevier. 2019. Doi:10.1016/B978-0-12-804566-4.00016-4

[12] Dwivedi, Yogesh. Kumar., Ismagilova, Elvira., Hughes, D. Laurie., Carlson, Jamie., Filieri, Raffaele., Jacobson, Jenna., Jain, Varsha., Karjaluoto, Heikki., Kefi, Hajer., Krishen, Anjala. S., Kumar, Vikram., Rahman, Mohammad. Mahfuzur., Raman, Ramakrishnan., Rauschnabel, Philip. A., Rowley, Jennifer., Salo, Jari., Tran, Gina., & Wang, Yichuan. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, *59*, 102168. 2021. Doi:10.1016/j.ijinfomgt.2020.102168

[13] Pandey, B., Pandey, D., Wariya, Subodh., Aggarwal, Gaurav., & Rastogi, Rahul. Deep Learning and Particle Swarm Optimisation-Based Techniques for Visually Impaired Humans' Text Recognition and Identification. *Augmented Human Research*, *6*(1), 14. 2021. doi:10.1007/s41133-021-00051-5

[14] Xue, Minglong., Shivakumara, Palaiahnakote., Zhang, Chao., Lu, Tong., & Pal, Umapada. Curved text detection in blurred/non-blurred video/scene images. *Multimedia Tools and Applications*, *78*(18), 25629–25653. 2019. Doi:10.1007/s11042-019-7721-2

[15] Li, Jinyang., Liu, Zhijing., & Yao, Yong. Defocus Blur Detection and Estimation from Imaging Sensors. *Sensors*, *18*(4), 1135. 2018. Doi:10.3390/s18041135

[16] Gholami, Raoof., & Fakhari, Nikoo. Support Vector Machine: Principles, Parameters, and Applications. In *Handbook of Neural Computation* (pp. 515–535). Elsevier. 2017. /doi:10.1016/B978-0-12-811318-9.00027-2

[17] Hicks, Steven., Strümke, Inga., Thambawita, Vajira., Hammou, Malek., Riegler, Michael., Halvorsen, Pal., & Parasa, Sravanthi. On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, *12*(1), 5979. 2022. doi:10.1038/s41598-022-09954-8

[18] Tharwat, Alaa. Independent component analysis: An introduction. *Applied Computing and Informatics*, *17*(2), 222–249. 2021. doi:10.1016/j.aci.2018.08.006

[19] Raisi, Zobeir., Naiel, Mohammed. A., Fieguth, Paul., Wardell, Steven., & Zelek, John. 2020. *Text Detection and Recognition in the Wild: A Review*.

[20] Katoch, Sourabh., Chauhan, Sumit. Singh., & Kumar, Vijay. A review on genetic algorithm: past, present, and future. *Multimedia Tools and Applications*, *80*(5), 8091–8126. 2021a. doi:10.1007/s11042-020-10139-6

[21] Vujovic, Zeljko. Classification Model Evaluation Metrics. International Journal of Advanced Computer Science and Applications. Volume 12. 599-606. 2021.10.14569/IJACSA.2021.0120670.

Authors' Profiles

NWUFOH Chinonyelum Vivian Ph.D.: A Lecturer at the Federal College of Animal Health and Production Technology, Ibadan, Oyo State, Nigeria. With Ph.D. in Computing in Computer Vision/Machine Learning. A member of the Computer Registration Council of Nigeria (CPN) and the Nigeria Computer Society (NCS). Interested in the areas of Artificial Intelligence, Computer Vision, Pattern Recognition, Machine Learning and eager to work with researchers with novel computing solutions that impacts the society positively and enhance social and governmental growth locally and globally. A reviewer for peer-reviewed journals. Passionately involved in community service of training youths especially in the STEM areas. Dr. Nwufoh Chinonyelum is not only a researcher and a trainer who also has an IT industry expertise.

SAKPERE Wilson (Ph.D.): Dr. Wilson Sakpere is a quality-focused researcher and engineer with a balanced background between academic and practical expertise. He has experience in critical writing, customer service, project supervision, technical services and development, and student development. In addition, he is skilled in technical support, systems engineering, interpersonal relationships and team building. He is a research enthusiast with a focus on computer vision, networks, pattern recognition, 3D image and point cloud analysis. The interests are directed towards applying innovations in a social context, with applications that will positively impact the environment and add value on a global scale while setting sustainable development goals. He has published in peer-reviewed journals and books.

How to cite this paper: Nwufoh C.V., Sakpere W., "A Comparative Model for Blurred Text Detection in Wild Scene Using Independent Component Analysis (ICA) and Enhanced Genetic Algorithm (Using a Bird Approach) with Classifiers", International Journal of Information Technology and Computer Science(IJITCS), Vol.16, No.5, pp.101-113, 2024. DOI:10.5815/ijitcs.2024.05.07