Intelligent Model for Detecting GAN-Generated Images Based on Multi-Classifier and Advanced Data Mining Techniques

Areen Mohammad Arabiat
Department of Communication and Computer Engineering, Al-Ahliyya Amman University, Amman, Jordan
Email: a.arabiat@ammanu.edu.jo (A.M.A.)

Manuscript received February 16, 2025; revised April 9, 2025; accepted May 4, 2025

***Abstract*—The ability of Generative Adversarial Networks (GANs) to produce images that closely resemble real ones has raised concern. This requires the creation of efficient detection techniques because it has significant ramifications for digital media, security, and ethics. In order to demonstrate the growing difficulties of attaining authenticity in the rapidly developing field of Artificial Intelligence (AI), this study introduces this critical issue by leveraging the "Detect AI-Generated Faces: High-Quality Dataset," obtained from Kaggle which contains 3,203 images of real human faces and AI-generated faces. However, the Orange3 data mining framework is used to analyze these images, focusing on extracting essential features such as shape attributes, texture descriptors, and color histograms. The dataset was divided into a training set (70%) and a testing set (30%) to evaluate our models effectively. Also, four machine learning algorithms were employed: K-Nearest Neighbor (KNN), Artificial Neural Network (ANN), Adaptive Boosting (AdaBoost), and Gradient Boosting (GB). The results revealed that KNN and AdaBoost achieved impressive accuracies of 99.4% and 97.07%, respectively, while GB and ANN reached even higher accuracies of 99.8% and 99.9%. These results underscore the effectiveness of advanced machine learning techniques in accurately distinguishing between AI-generated and real faces.**

***Index Terms*—artificial intelligence-generated images, Kaggle, adaptive boosting, decision trees, gradient boosting, and random forest**

I. Introduction

Technological developments in computer vision and Artificial Intelligence (AI) have made it feasible for machines to simulate human characteristics, facilitating jobs like booking flights and diagnosing illnesses [1]. With the use of Machine Learning (ML) and Deep Learning (DL) models, researchers are creating methods to more precisely identify manipulated images by detecting frauds more effectively than traditional methods, which extract simple information and features. Better resilience is essential for real-world AI applications since DL models are susceptible to adversarial attacks, despite their demonstrated good performance in computer vision tasks [2, 3]. However, AI-generated images can be generated using Generative Adversarial Networks (GANs), which are unsupervised ML models consisting of a discriminator and generator network that separate genuine and erroneous data and produce high-quality synthetic data, offering speed, efficiency, and adaptability to various data types [4, 5].

Usually, a binary classifier is trained using a huge collection of GAN images from pre-trained models in order to create fake GAN images. However, access to the particular model that the attacker employed is typically not available in real-world systems. Two approaches are investigated in order to train a classifier with fewer fake images by locating the important up-sampling component and showcasing distinct frequency-domain artifacts by creating an emulator framework to replicate the standard generating pipeline [6]. With an emphasis on artificial face image synthesis, the researchers in [7] offered a novel method for separating GAN-generated images from real ones using spectral band differences. Through the use of cross-band and spatial co-occurrence matrices, face images are digitally preserved and subsequently fed into an architecture of Convolutional Neural Networks (CNN). In a variety of post-processing settings, the performance increase is greater than 92%.

Mercaldo *et al*. used a dataset of retinal images, this research proposed a technique to evaluate the distinguishability of bioimages produced by a generative adversarial network. To find out how well the models can distinguish between real and fraudulent retinal images, the researchers trained a number of supervised ML models. Using a deep convolutional generative adversarial network, they showed that a classifier with an F-measure greater than 0.95 can accurately identify created images, even when they were not visually noticeable as fakes [4]. According to [8], a study on GAN-based image-to-image translation detection, some detectors exhibited improvements on original photos but deteriorating on compressed images similar to those found on Twitter. Even when training-test mismatching occurred, deep networks, in particular XceptionNet, maintained resilience better than other detectors. According to [9], detection accuracy of up to 95% can be attained by both conventional and DL detectors, with DL offering a high accuracy of up to 89%. On the other hand, Rossler *et al.* [10] employed CNNs to identify them. In addition to establishing a benchmark dataset, the study showed how well CNNs differentiate between authentic and manipulated images. CNN architecture achieved a 91.83% forgery detection rate, highlighting the need for reliable detection techniques in the face of developing AI

doi: 10.18178/ijeetc.14.3.147-157

capabilities.

In order to enhance medical signal processing and diagnosis, namely in the classification of epilepsy, the researcher in [11] investigates the use of AI-generated content approaches. Also, the researcher presents a system for creating synthetic EEG signals using generative adversarial networks to solve data imbalance and scarcity. The model effectively processes and categorizes EEG signals using a temporal convolutional network model that is attention-based. The outcomes demonstrate the potential of AI-generated content in medical signal processing, with a 98.89% accuracy and a 98.91% F1 score.

Baraheem *et al*. in [12] offered an approach for employing Convolutional Neural Networks (CNNs) to distinguish between actual and AI-generated photos. The method entails integrating Class Activation Maps (CAM), using transfer learning, and gathering GAN-generated images from diverse tasks and structures. The method outperform other datasets and setups and obtained 100% accuracy on the Real or Synthetic Images dataset. EfficientNetB4, which has been pre-trained and optimized on the dataset using Adam as an optimizer, was the best detector.To identify GAN-generated face images.

Tao and his colleagues in [13] provided a blind method based on hand-crafted features. GAN-generated images and natural texture and sensor sounds are used as hints in this method. Additionally, subtractive pixel adjacency matrix (SPAM) features are retrieved from real and generated images using uniform Local Binary Pattern (LBP) features. The SVM classifier confirms that the images are correct. The method has a 97.60% accuracy rate in identifying GAN-generated fraudulent face images.
However, soft computing neural network models (Shallow-FakeFaceNets) with an effective facial manipulation detection pipeline were employed in [14] using a dataset of Handcrafted Facial Manipulation (HFM) images. Shallow-FakeFaceNet (SFFN), a neural network classifier, uses altered facial landmarks to identify fraudulent images. For handcrafted fake facial images, the method's best Area Under the Receiver Operating Characteristic (AUROC) performance is 72.52%, whereas for tiny GAN-generated fake photos, it reaches 93.99%.

The researchers in [15] focused on artificial face image synthesis and suggests a novel method for distinguishing real photos from GAN-generated ones using spectral band differences. The method uses a cross-band and spatial co-occurrence matrix to digitally preserve face images. In a variety of post-processing settings, the convolutional neural network architecture achieves a performance improvement of more than 92% when utilized for identifying genuine faces.

In this work, we provide a novel approach that builds several classification models by extracting features from images using the Orange3 data mining tool. These features form the basis for training a number of fundamental classifiers, such as AdaBoost, ANN, GB, and KNN. The study emphasizes Orange3's potential as an efficient image categorization tool and emphasizes the need for strong models to counter AI-generated material. This result show that model performance in image analysis can be significantly enhanced by integrating a variety of ML approaches with effective feature extraction. This strategy not only improves detection techniques but also offers insightful information about the persistent problem of recognizing AI-generated images. With everything considered, the study highlights how incorporating cutting-edge tools and methods can improve image-classifying employment accuracy. In the other hand, the main goal of this research is to create strong classification models that can distinguish between actual and AI-generated images. This will enhance detection methods and offer important insights into the properties of AI-generated content.

II. Methodology

Using the Orange3 data mining framework, this paper proposes a thorough ML model for identifying whether images of human faces are (AI)-generated or real, as shown in Fig. 1, which illustrates the proposed model's structure. To ensure accurate evaluation, the dataset's 3,203 images—which were sourced from Kaggle—are split into 70% for training and 30% for testing. However, to increase generalization, decrease dimensionality, and boost computing performance, the SqueezeNet embedder model is restricted to 1000 features. This choice reduces the possibility of noisy or insignificant information, enhances interpretability, and lessens overfitting. This method produces better machine learning results by achieving a compromise between efficiency and information. However, actual and AI-generated images are included in this category. To improve image quality, preprocessing operations include scaling, normalization, and feature extraction using methods like histogram equalization. From raw pixel data, pertinent visual elements, including color histograms, texture patterns, and facial landmarks, are converted into valuable features. AdaBoost, KNN, GB, and ANN are among the classifiers used in the model. During model evaluation, a 10-fold cross-validation process is used to avoid overfitting. The classification process is evaluated using performance indicators including F1-score, accuracy, sensitivity, and precision. The model's implementation and visualization are made easier by Orange3's attractive interface, which also makes it possible to analyze feature importance and decision limits. The goal of the research is to improve visual content classification by tackling the difficulties presented by AI-generated images. Consequently, these procedures in Orange3 will efficiently extract features and get the dataset ready for reliable image classification as (AI) or real. The ability to create a thorough ML model is improved by the user-friendly interface, which makes it simple to visualize and examine each step of the process. Fig. 1 shows the structure proposed model while Fig. 2 shows training model using Orange 3.

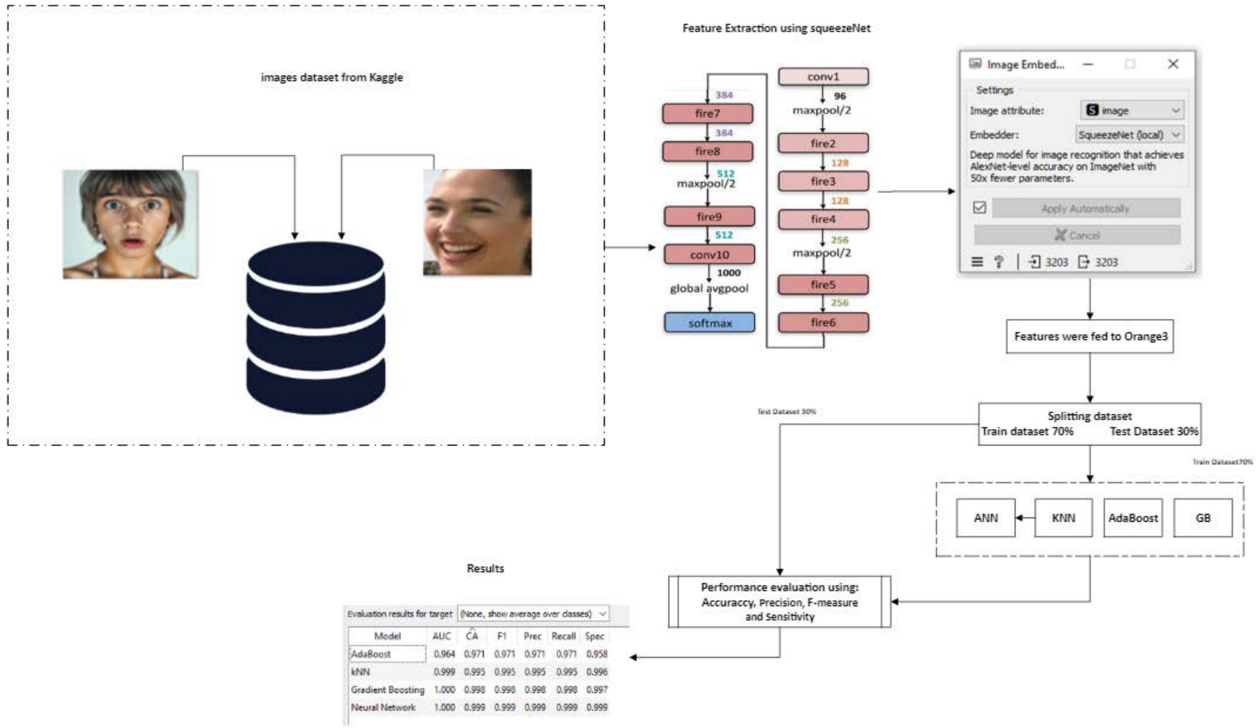


Fig. 1. Proposed model structure.

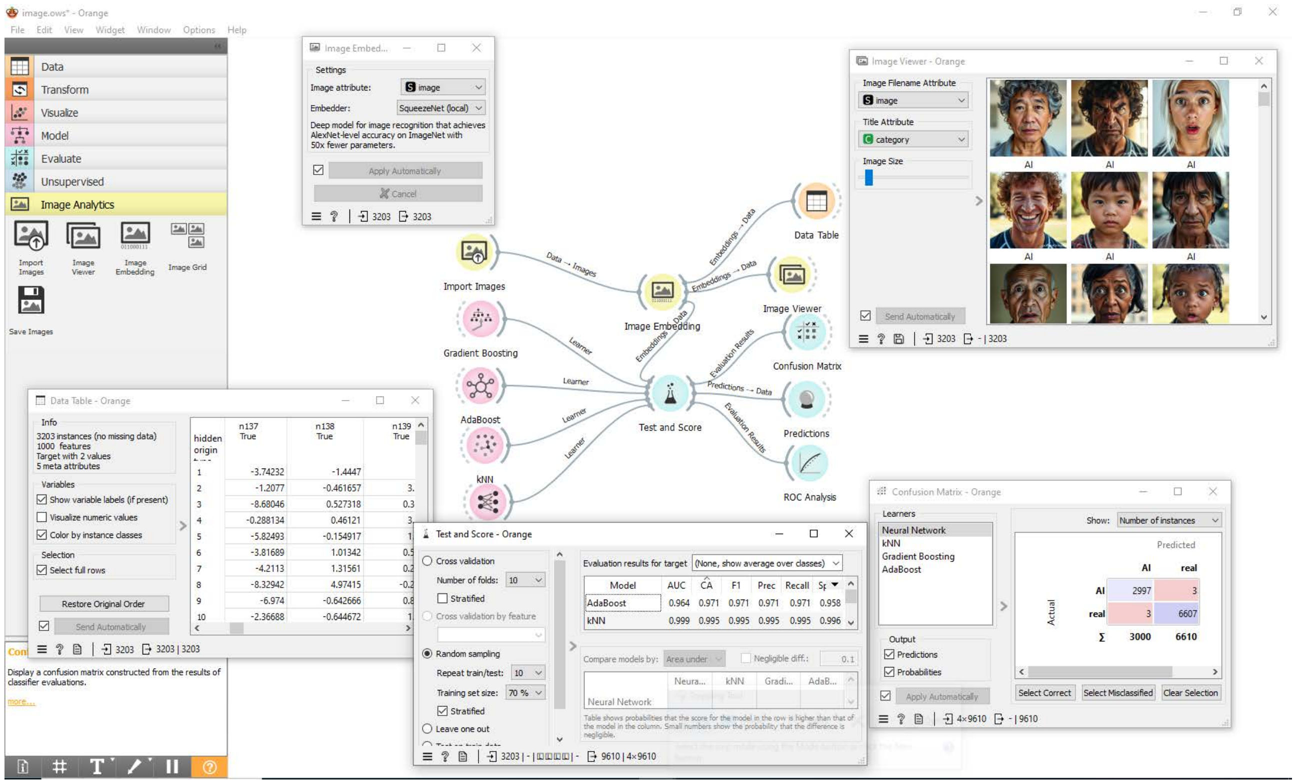
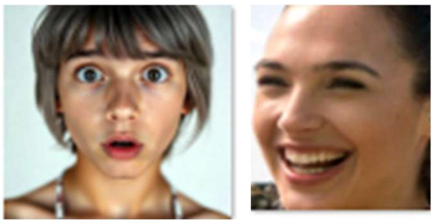


Fig. 2. Orange 3 training model.

A. Dataset

The dataset used in this study consists of 3,203 Kaggle images that are divided into two categories: AI-generated faces and real human faces. This dataset, created by Shahzaib Ur Rehman, includes images of AI-generated synthetic faces as well as genuine human faces, intended for use in DL and ML applications. It offers a useful tool for creating and evaluating classifiers that can recognize real faces from AI-generated ones. This dataset is carefully selected to support state-of-the-art research and applications, making it perfect for tasks like facial image analysis, deep fake identification, and image authenticity verification [16]. 30% of the dataset is used for testing, while 70% is used for training, to guarantee a thorough assessment. Each image goes through preprocessing, which includes feature extraction, scaling, and normalization. To improve image quality, methods like histogram equalization are used. To convert raw pixel data into useful features for analysis, 1,000 features are retrieved, including visual qualities like color histograms, texture patterns, and facial landmarks. For the ML model created with the Orange3 data mining framework, this systematic methodology provides the groundwork. Fig. 3 (a) shows the AI-generated face image, while Fig. 3 (b) shows the real face image.



(a) (b)

Fig. 3. Face images: (a) Image generated by AI and (b) real image.

B. Feature Extraction

In order to minimize the amount of resources needed, feature extraction is an essential procedure in computer vision, data mining, image retrieval, and image processing. In order to solve issues and communicate significant elements, it entails removing the visual components of an image [17–19]. A modularized CNN known as the Fire module serves as the foundation for SqueezeNet, a thin and powerful CNN model. However, SqueezeNet's architecture allows for great precision and few parameters, which makes it effective for tools with limited resources. It employs "fire modules" to extract characteristics from input images, which are then processed by SoftMax, convolutional, and global average pooling layers. The architecture of SqueezeNet consists of eight fire modules, three max-pooling layers, two convolution layers, one global average pooling layer, and one SoftMax output layer, making up the model's fifteen layers [20–22]. SqueezeNet's construction is depicted in Fig. 4.

Essential elements such as texture descriptors, color histograms, and shape attributes can be extracted from images using the Orange3 tool [23]. For classification tasks, the "Image Analytics" add-on enables users to examine geometric attributes such as perimeter, area, and contour. Local Binary Patterns and Gabor filters are examples of texture descriptors that capture textural features, whereas shape attributes are defined for accuracy [24, 25]. Additionally, Orange3 can effectively extract more than 1,000 characteristics from images using SqueezeNet, which improves the feature extraction procedure even more. The study's clarity and depth are greatly increased by this thorough integration, which also helps to improve feature relationship analysis and increases model robustness and classification accuracy.

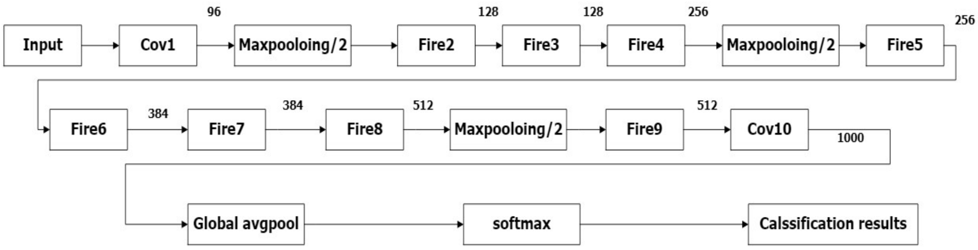


Fig. 4. SqueezNet structure [26].

C. Classification Methods

A branch of AI called ML creates supervised and unsupervised learning algorithms that can learn from and predict data that has not been observed previously [27]. However, ML algorithms learn to perform tasks on their own. ML, a technology that teaches machines to handle data better, is becoming more and more common in data mining, image processing, and predictive analytics [28]. In statistics and ML, classification techniques are used to group or classify data into specific categories. They are essential components of data mining and ML. Data analysis, process classification, and correct decision-making based on patterns and visions retrieved from the dataset are all frequently accomplished through the application of DL and ML [29]. Because of their better experimental results, the study selected AdaBoost, KNN, GB and ANN. AdaBoost handles imbalanced datasets and improves weak classifiers by lowering bias. KNN is a straightforward and efficient method for proximity-based instance classification, especially in complex decision boundaries. GB minimizes overfitting across data types and has a high prediction accuracy. ANN effectively learns features and models' complex relationships, particularly in image classification problems. These classifiers offer an original approach to the classification problem because they haven't been widely applied in similar scenarios before, and the studies' empirical data supports their inclusion as the best classifiers for this investigation. Fig. 5 illustrates classification techniques that employed in this study.

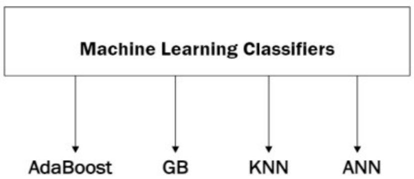


Fig. 5. ML classification method.

1) Adaptive Boosting (AdaBoost)

The AdaBoost algorithm is a supervised learning-based ensemble technique that creates a robust classifier distinguishes between positive and negative scenarios by combining the training of multiple weak classifiers as shown in Fig. 6 [30, 31]. AdaBoost uses single-level or partitioned decision trees to merge weak classifiers into a powerful classifier. This approach assigns equal weights to each data point, giving points with inaccurate classifications more weight. In succeeding models, points with higher weights are given more relevance as the model trains until a smaller error is achieved. Following each step, a number of new models are produced based on the prior model's sampling error rate. In order to make some undesirable samples stand out sufficiently to be seen in the sample guarantee, the sample weight is increased in accordance with the sampling error rate. Following the generation of each model, weak learners gain knowledge by repeated preparation, as per the interaction stated. As a result, models 1, 2, 3, and *N* are all separate models, sometimes referred to as decision trees. To produce an efficient learner, a combination that is balanced is finally executed [32–36].

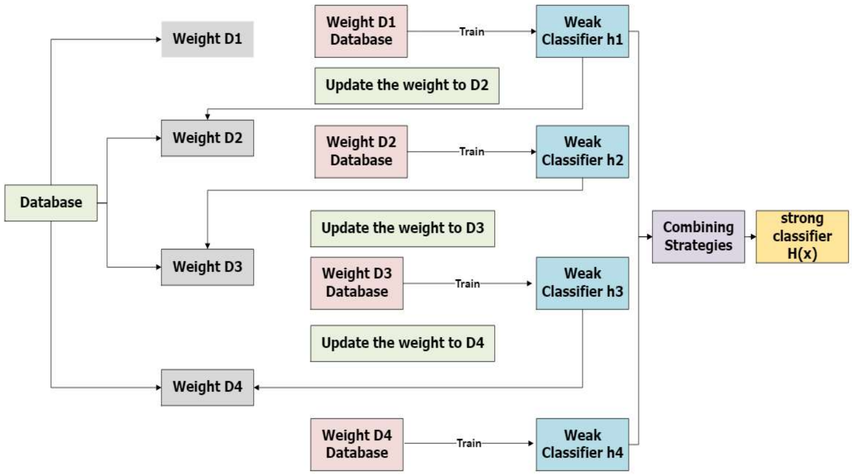


Fig. 6. Structure of AdaBoost [37].

2) K-Nearest Neighbor (KNN)

KNN algorithm is a straightforward, flexible, and incredibly effective method for classifying data according to how closely it resembles the training dataset. Identifying classes according to the number of *k* values that are closest to the training data is another usage for it [38–40]. In order to identify the *k* closest samples to unknown samples in a dataset, the KNN classification technique leverages the distance between eigenvectors. "Weights" and "*n*_neighbors" are the two main factors that are used to identify the label of the unknown sample. With inverse distances and homogenous weights, the weight function forecasts potential values [41]. Fig. 7 shows the structure of the KNN algorithm.

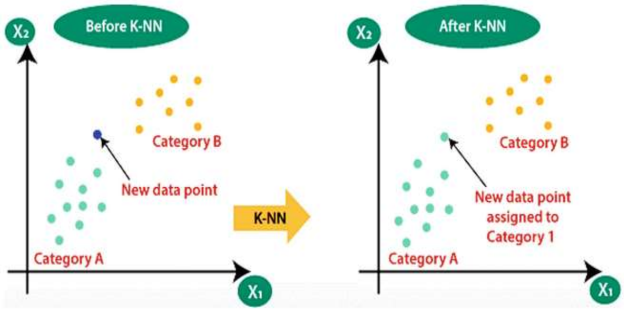


Fig. 7. Structure of the KNN algorithm [42].

3) Gradient Boosting (GB)

GB is a ML technique used for regression and classification that creates a prediction model as an ensemble of weak prediction models, usually decision trees. One of the boosting processes, it constructs the model step-by-step and reduces the bias error of the model [43]. In order to increase prediction accuracy, GB trains weak classifiers, such as decision trees. It improves performance in regression and classification tasks by minimizing bias error and maximizing prediction accuracy through the application of a cost function known as Mean Square Error (MSE) [44]. However, by fixing the errors of the earlier models, the basic goal is to progressively increase the ensemble's predictive capability, the model is considered a weak learner if it completes a task just a little bit better than random guessing. The ensemble consists of M trees [45]. However, an explanation of the GB model is provided as

$$F_m(x) = \sum_{m=1}^{M} \gamma_m h_m(\mathrm{x}) \tag{1}$$

where $F_m(x)$ is the cumulative prediction after mmm models have been added to the ensemble, M is the maximum number of weak learners (e.g., decision trees) that will be added to the ensemble, γ_m determines how much the predictions from the new model hm(*x*)h_m(*x*)hm(*x*) are scaled before being added to the cumulative prediction Fm(*x*)F_m(*x*)Fm(*x*), $h_m(x)$ is typically a function or model that is trained to predict the residuals (errors) of the previous predictions. In many implementations, this is often a decision tree [46].

As shown in Fig. 8 the context of GB, includes the variables *r*, *x*, and *y*, where *r* indicates the errors or residuals from the prior iteration, these residuals are iteratively corrected by the model in GB to enhance predictions; *x* stands for the predictors or input features that the model uses, the model makes predictions based on these variables; *y* indicates the actual output or target variable that the model is attempting to forecast, in the context of categorization, this might stand in for class labels.

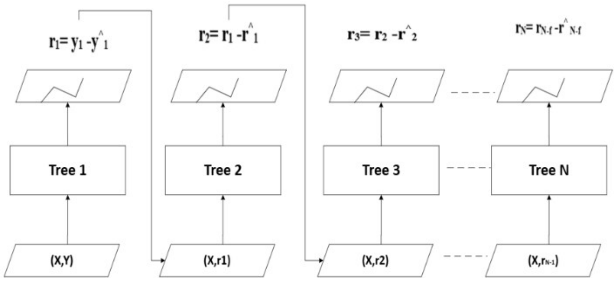


Fig. 8. Gradient boosting structure [45].

4) Artificial Neural Network (ANN)

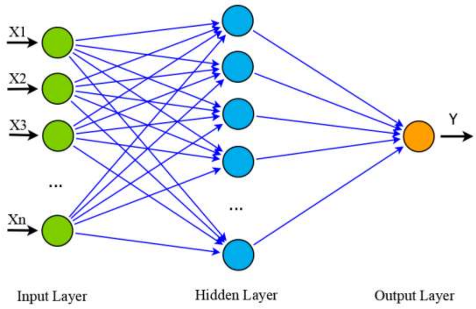


Fig. 9. ANN classifier structure [49].

Hundreds of individual units, or artificial neurons, joined by coefficients (weights) that make up the neural structure, create an artificial neural network, a computer model inspired by biology as shown in Fig. 9 [47]. Additionally, ANN is composed of three layers: the input, hidden, and output layers, which is a feed-forward neural network. It can handle nonlinear functions and learning weights and is useful for addressing problems involving text, images, and tabular data. ANNs are capable of learning universal approximations, which are complex relationships between input and output data. ANNs are used by researchers to resolve complicated issues like WiFi and cellular networks coexisting in unlicensed spectrum [48].

III. Results and Discussion

Once the primary model assumptions have been tested and validated, it is crucial to examine how well the proposed model predicts. Therefore, the suitability of the suggested models was assessed using assessment measures. Because of this, the confusion matrix is a useful instrument for evaluating the performance of the classification and prediction algorithms. In accuracy rate computations, it calculates the proportions of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [50–55]. This study examined the effectiveness of the suggested model using the F-measure, accuracy, precision, and sensitivity, as indicated in (2), (3), (4), and (5):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+F} \tag{2}$$

$$Sensitivity = \frac{TP}{TP+} \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$F-measure = \frac{2\times Precision\times Sensitivity}{Precision+Sensitivity} \tag{5}$$

However, Fig. 10 illustrates the performance matrices evaluation by using Orange3 to assess the effectiveness of classifiers: Fig. 10 (a) accuracy, Fig. 10 (b) the F-measure, Fig. 10 (c) precision and Fig. 10 (d) sensitivity.

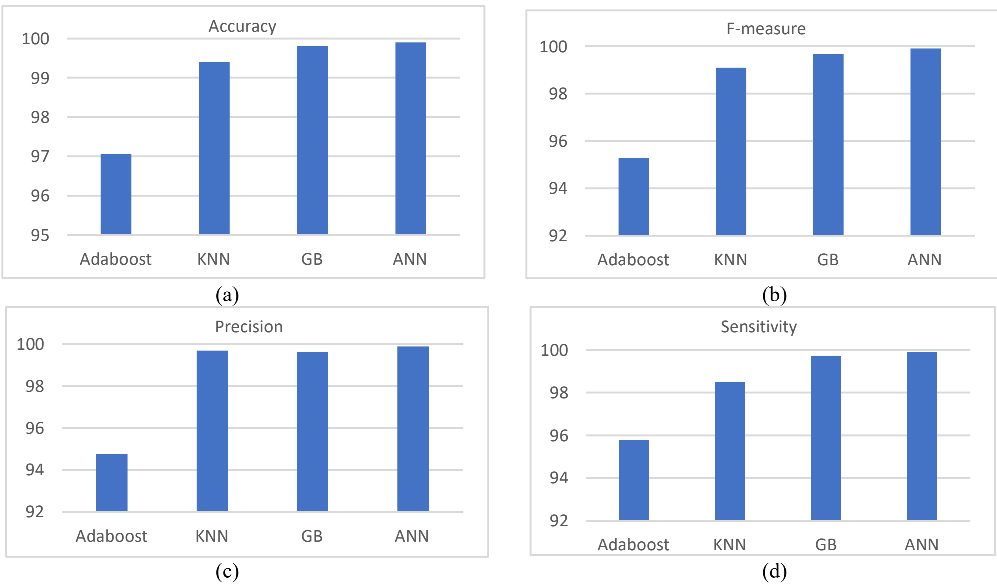


Fig. 10. Performance matrices evaluation: (a) Accuracy of all classifiers, (b) F-measure of all classifiers (c) precision of all classifiers, and (d) sensitivity of all classifiers.

Several methods, including AdaBoost, ANN, GB, and KNN, were employed in this study to classify real vs. AI-generated images.

The confusion matrix is used to evaluate each algorithm's performance, Fig. 11 describes the confusion matrix of classifiers using Orange3: (a) confusion matrix of ANN, (b) confusion matrix of KNN, (c) confusion matrix of GB, and (d) confusion matrix of AdaBoost, as shown in Fig. 11 (a), (b), (c), and (d), respectively.

The study examined how well different classification models performed in differentiating between actual and artificial intelligence-generated images of human faces. AdaBoost classifier achieved 97.07% accuracy rate while the accuracy of the KNN model was 99.4%, high-dimensional data and computational resources may limit its effectiveness. With a remarkable accuracy of 99.8%, the GB model proved its capacity to capture the complexities of the dataset. The remarkable 99.9% accuracy of the ANN model was ascribed to its capability to identify intricate in the data. However, in real-time applications, the complexity and resource requirements could provide difficulties. The study indicated that while AdaBoost offers a strong basis, KNN, GB, and ANN greatly improve classification accuracy and reliability. To overcome the weaknesses of individual models and capitalize on the advantages of several classifiers, future studies could investigate ensemble approaches. The study emphasizes how crucial it is to choose the right categorization methods when dealing with AI-generated content because incorrect classification can have serious implications in a number of areas, such as media integrity and security. Fig. 12 depicts a comparison of performance matrices of all classifiers, while Table I shows classification results and performance evaluation. Also, Table I presents the findings of the comparison of the classifier's performance evaluation of the suggested model with earlier research.

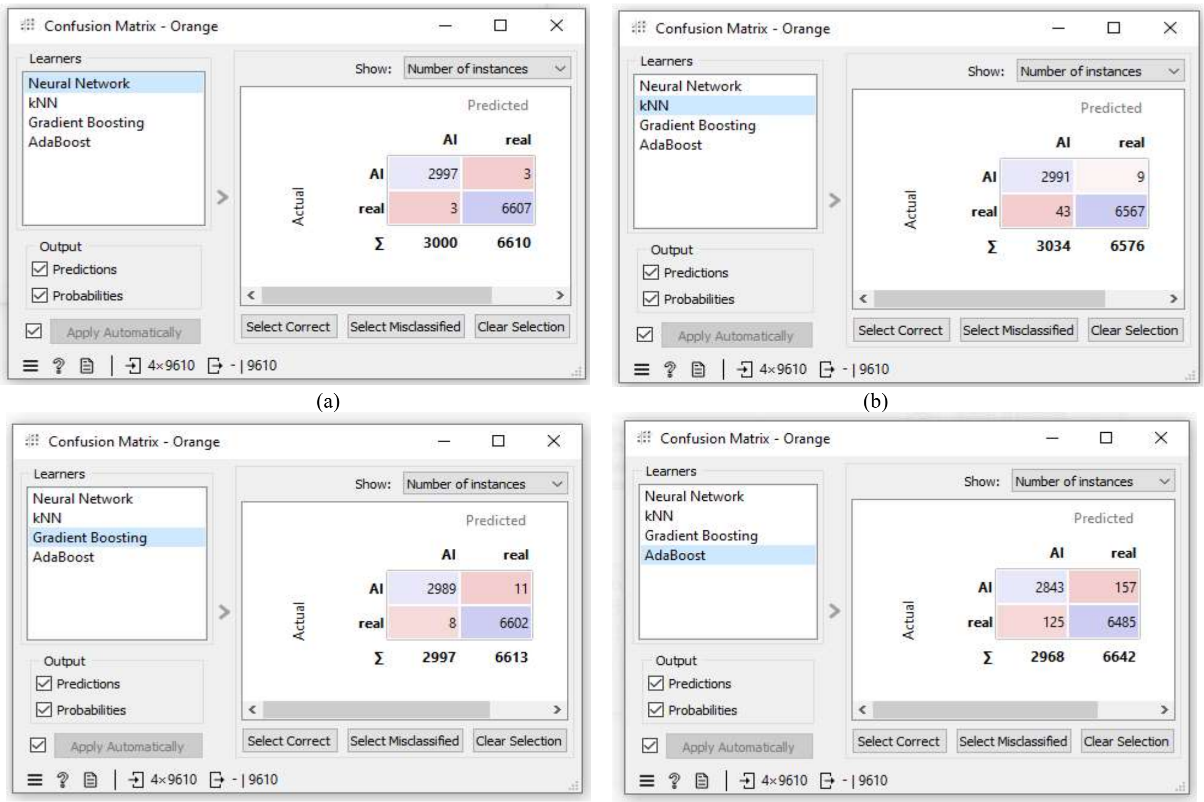


Fig. 11. Confusion matrix evaluation: (a) confusion matrix of ANN, (b) confusion matrix of KNN, (c) confusion matrix of GB, and (d) confusion matrix of AdaBoost.

TABLE I: CLASSIFICATION RESULTS AND PERFORMANCE EVALUATION

Model	AUC	Accuracy	F-measure	Precision	Sensitivity
AdaBoost	96.4	97.07	95.27	94.77	95.79
KNN	99.9	99.4	99.1	99.7	98.5
GB	100	99.8	99.68	99.63	99.73
ANN	100	99.9	99.9	99.9	99.9

TABLE II: COMPARISON OF THE CLASSIFIER'S PERFORMANCE ASSESSMENT FROM EARLIER RESEARCH AND PROPOSED MODEL

Study/Year	Methods\Model	Performance metrics
[4] /2023	Supervised machine learning models	F-measure > 0.95
[7] /2023	CNNs with cross-band and spatial co-occurrence matrices	Performance > 92%
[9] /2023	Convolutional and deep learning detectors	accuracy up to 89%
[ ]/rosseler	Convolutional neural networks (CNN).	Accuracy 91.83%
[10]/2024	Convolutional neural networks (CNN).	F1-measure 98.89 %
[12] 2022	Hand-crafted features with SVM classifier	97.60%
[13] 2021	Neural network classifier using altered facial landmarks	AUC 72.52% - 93.99%
Proposed	AdaBoost, ANN, GB and KNN using Orange3 data mining tool	Accuracy 99.9 %

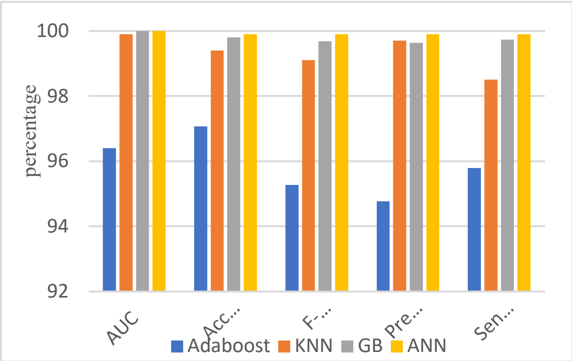
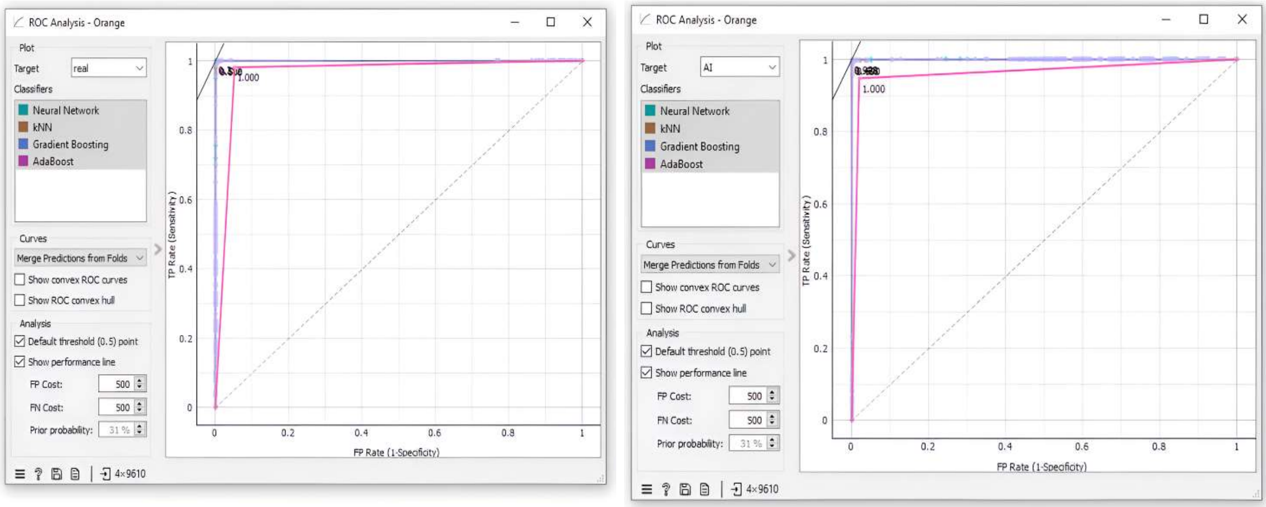


Fig. 12. Performance matrices vs. classifiers.

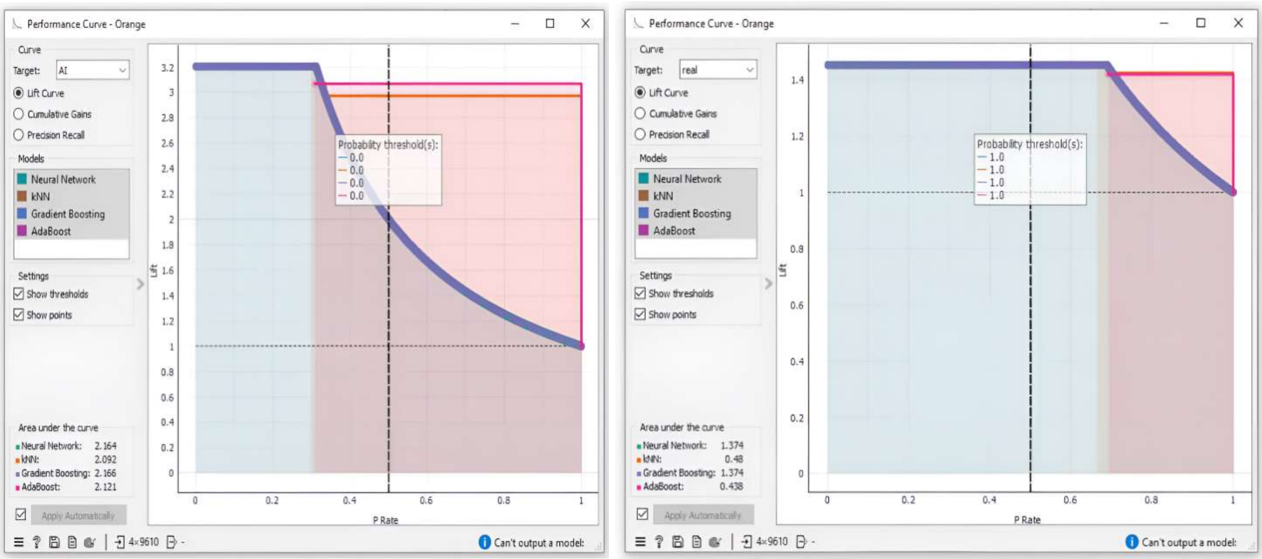
In the other hand, this work used a visual representation of a binary classifier's performance is the Receiver Operating Characteristic (ROC) curve [56]. In computational statistics and ML, it is now the accepted metric for assessing binary classifications. The area under the ROC curve (AUC) measures the classifier's overall performance, whereas the ROC curve efficiently illustrates the trade-offs between true positive and false positive rates. This tool is crucial for evaluating binary classifiers' performance in a range of applications [57–60]. However, Fig. 13 illustrates the ROC curve that is utilized to assess the effectiveness of classifiers using Orange3: (a) ROC curve analysis according to AI and (b) ROC curve analysis according to real target.

Additionally, the performance curve is employed to assess classifier performance. Lift curves, cumulative gains, and precision-recall curves are the three types of curves that compare the fraction of genuine positive data occurrences to the classifier's threshold [61]. The model is improved with a longer flatness and a larger initial curve, as Fig. 14 illustrates. The performance curve is also utilized to assess the effectiveness of classifiers using Orange3. AI target performance curves are shown in Fig. 14 (a); real target performance curves are shown in Fig. 14 (b).



(a) (b)

Fig. 13. ROC curves analysis of all classifiers using orange3 according to target: (a) ROC curve analysis according to AI and (b) ROC curve analysis according to real target.



(a) (b)

Fig. 14. Lift curves analysis of all classifiers using orange3 according to target: (a) lift curve analysis according to AI and (b) lift curve analysis according to real target.

IV. Conclusion

Creating a ML model that can distinguish between actual and artificial intelligence-generated human face images is the main objective of the work. The model makes use of a collection of 3,203 photos from Kaggle, which are separated into images created by AI and actual human faces, using the Orange3 data mining framework. The study extracted visual features from the images using feature extraction, which forms the basis of classification systems. AdaBoost, K-NN, GB, and ANNs are among the techniques used in the model. F1-score, recall, accuracy, and precision are some of the measures was used to evaluate how well these classifiers work. To determine which of the four algorithms was the best classifier, a comparative analysis was conducted. The results showed that the ANN classifier had the highest accuracy, at 99.9%. By addressing the consequences of AI-generated images in a variety of fields, such as media, entertainment, security, and privacy, the work seeks to advance the conversation regarding the authenticity of visual information and improve the capacity to distinguish between real and artificial faces. Additionally, Limitations of the classifier study include a particular dataset that might not accurately represent real-world situations, concentrating on a small number of classifiers without investigating other algorithms or hybrid models, and failing to consider each aspect of classifier performance. Making decisions might be difficult due to the ANN model's interpretability, particularly in high-stakes industries like healthcare and finance. Furthermore, the study did not take into consideration the computational resources needed for practical use, which can have an impact on their viability. Understanding these limitations can motivate further study for better classification methods.

V. Future Work

Future image classification research can increase robustness and accuracy by creating hybrid models with different algorithms, utilizing transfer learning for small datasets, refining real-time classification systems for dynamic environments, improving explain ability, adding adversarial training to make them insensitive to manipulation, and growing datasets to include a variety of AI-generated images. Ethical issues and cross-domain applications can increase these classifiers' effectiveness. However, this study explores the potential applications of classifiers like AdaBoost and ANN in various industries, including supply chain management, healthcare, finance, retail, transportation, and environmental monitoring. These models can improve demand forecasting, diagnostic accuracy, credit scoring, customer experience customization, autonomous vehicle decision-making, and environmental monitoring. By transforming empirical data into insights, these classifiers can stimulate innovation, enhance decision-making, and advance society by fostering the development of more effective systems.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgment

The author wishes to thank Al Ahliyya Amman University for their support.

References

[1] T. Arora and R. Soni, "A review of techniques to detect the GAN-generated fake images," *Elsevier eBooks*, 2021, pp. 125–159. doi: 10.1016/b978-0-12-823519-5.00004-x

[2] E.-C. Chen and C.-R. Lee, "Towards fast and robust adversarial training for image classification," *Lecture Notes in Computer Science*, 2021, pp. 576–591. doi: 10.1007/978-3-030-69535-4_35

[3] R. Agarwal, D. Khudaniya, A. Gupta, and K. Grover, "Image forgery detection and deep learning techniques: A review," in *Proc. 2022 6th Int. Conf. on Intelligent Computing and Control Systems (ICICCS)*, May 2020, pp. 1096–1100.

[4] F. Mercaldo, L. Brunese, F. Martinelli *et al.*, "Generative adversarial networks in retinal image classification," *Applied Sciences*, vol. 13, no. 18, 10433, Sep. 2023.

[5] A. M. Almars, "DeepFakes detection techniques using deep learning: A survey," *Journal of Computer and Communications*, vol. 9, no. 5, pp. 20–35, Jan. 2021.

[6] X. Zhang, S. Karaman, and S. F. Chang, "Detecting and simulating artifacts in GAN fake images (extended version)," arXiv:1907.06515v2, 2019.

[7] E. Nowroozi and Y. Mekdad, "Detecting high-quality GAN-generated face images using neural networks," *River Publishers eBooks*, 2023, pp. 235–252. doi: 10.1201/9781003373384-12

[8] F. Marra, D. Gragnaniello, D. Cozzolino, and L. Verdoliva, "Detection of GAN-generated fake images over social networks," in *Proc. 2019 IEEE Conf. on Multimedia Information Processing and Retrieval*, 2018. doi: 10.1109/mipr.2018.00084

[9] G. Aggarwal, A. K. Srivastava, K. Jhajharia, N. V. Sharma, and G. Singh, "Detection of deep fake images using convolutional neural networks," in *Proc. 2023 3rd Int. Conf. on Technological Advancements in Computational Sciences (ICTACS)*, 2023. doi: 10.1109/ictacs59847.2023.10389973

[10] A. Rössler, D. Cozzolino, L. Verdoliva *et al.* "FaceForensics++: Learning to detect manipulated facial images," arXiv:1901.08971v3. doi: 10.48550/arxiv.1901.08971

[11] L. Fang, Y. Li, M. Shao *et al.*, "Enhancing medical signal processing and diagnosis with AI-generated content techniques," *IEEE Journal of Biomedical and Health Informatics*, Jul. 2024. doi: 10.1109/JBHI.2024.3429560

[12] S. S. Baraheem and T. V. Nguyen, "AI vs. AI: Can AI detect AI-generated images?" *Journal of Imaging*, vol. 9, no. 10, 2023. doi: 10.3390/jimaging9100199

[13] T. Fu, M. Xia, and G. Yang, "Detecting GAN-generated face images via hybrid texture and sensor noise based features," *Multimedia Tools and Applications*, vol. 81, pp. 26345–26359, 2022. doi: 10.1007/s11042-022-12661-1

[14] S. Lee, S. Tariq, Y., Shin, and S. S. Woo, "Detecting handcrafted facial image manipulations and GAN-generated facial images using Shallow-FakeFaceNet," *Applied Soft Computing*, vol. 105, 2021. doi: 10.1016/j.asoc.2021.107256

[15] E. Nowroozi and Y. Mekdad, "Detecting high-quality GAN-generated face images using neural networks," *River Publishers eBooks*, 2023. doi: 10.1201/9781003373384-12

[16] S. Shahzaib, "Detect AI-generated faces: High-quality dataset," *Kaggle*, vol. 1, no. 1, pp. 1–1, Dec. 2024.

[17] M. Altayeb and A. Arabiat, "Crack detection based on mel-frequency cepstral coefficients features using multiple classifiers," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 3, pp. 3332–3341, 2024.

[18] G.-F. Fan, Y.-Y. Han, J.-W. Li *et al.*, "A hybrid model for deep learning short-term power load forecasting based on feature extraction statistics techniques," *Expert Systems With Applications*, vol. 238, 2024. doi: 10.1016/j.eswa.2023.122012

[19] S. Kılıçarslan and S. Kılıçarslan, "A comparative study of bread wheat varieties identification on feature extraction, feature selection and machine learning algorithms," *European Food*

Research and Technology, vol. 250, no. 1, pp. 135–149, 2023.

[20] N. Sharmili, S. Yonbawi, S. Alahmari *et al.*, "Earthworm optimization with improved SqueezeNet enabled facial expression recognition model," *Computer Systems Science and Engineering*, vol. 46, no. 2, pp. 2247–2262, 2023.

[21] M. Hao, Q. Sun, C. Xuan *et al.* "SqueezeNet: An improved lightweight neural network for sheep facial recognition," *Applied Sciences*, vol. 14, no. 4, 2024. doi: 10.3390/app14041399

[22] T. H. B. Nguyen, E. Park, X. Cui *et al.*, "FPADNET: Small and efficient convolutional neural network for presentation attack detection," *Sensors*, vol. 18, no. 8, 2018. doi: 10.3390/s18082532

[23] H. Yu, W. Yang, G.-S. Xia, and G. Liu, "A color-texture-structure descriptor for high-resolution satellite image classification," *Remote Sensing*, vol. 8, no. 3, Mar. 2016. doi: 10.3390/rs8030259

[24] H. Li and J. Wang, "Texture analysis using Gabor filters and local binary patterns." *International Journal of Computer Vision*, vol. 131, no. 4, pp. 456–470, 2023.

[25] T. Vijayan, M. Sangeetha, A. Kumaravel, and B. Karthik, "Feature selection for simple color histogram filter based on retinal fundus images for diabetic retinopathy recognition," *IETE Journal of Research*, vol. 69, no. 2, pp. 987–994, 2023.

[26] A. Agnihotri and N. Kohli, "A novel lightweight deep learning model based on SqueezeNet architecture for viral lung disease classification in X-ray and CT images," *International Journal of Computational and Experimental Science and Engineering*, vol. 10, no. 4, 2024. doi: 10.22399/ijcesen.425

[27] L. Xu, X. Chen, and X. Yang, "Tourism image classification based on convolutional neural network SqueezeNet—Taking Slender West Lake as an example," *PLoS ONE*, vol. 19, no. 1, 2024, e0295439. doi: 10.1371/journal.pone.0295439

[28] M. Karimzadeh, D. Basvoju, A. Vakanski, I. Charit, F. Xu, and X. Zhang, "Machine learning for additive manufacturing of functionally graded materials," *Materials*, vol. 17, no. 15, 2024, 3673. doi: 10.3390/ma17153673

[29] A. Arabiat, M. Hassan, and O. Almomani, "WEKA-based machine learning for traffic congestion prediction in Amman City," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 4, pp. 4422–4434, 2024.

[30] A. Akgundogdu, "Detection of pneumonia in chest X-ray images by using 2D discrete wavelet feature extraction with Random Forest," *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, pp. 82–93, 2020.

[31] J. Zheng, D. Lin, Z. Gao, S. Wang, M. He, and J. Fan, "Deep learning assisted efficient AdaBoost algorithm for breast cancer detection and early diagnosis," *IEEE Access*, vol. 8, pp. 96946–96954, 2020.

[32] Y. Xie, T. Wang, M. K. Jeong, and G. T. Lee, "Multi-source adaptive thresholding AdaBoost with application to virtual metrology," *International Journal of Production Research*, 2024. doi: 10.1080/00207543.2024.2314151

[33] T. R. Mahesh, V. D. Kumar, V. V. Kumar, J. Asghar, O. Geman, G. Arulkumaran, and N. Arun, "AdaBoost ensemble methods using K-fold cross validation for survivability with the early detection of heart disease," *Computational Intelligence and Neuroscience*, 2022. https://doi.org/10.1155/2022/9005278

[34] S. Usha and S. Kanchana, "Revived ant colony optimization-based AdaBoost algorithm for heart disease and diabetes (HDD) prediction," *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 4, pp. 1552–1567, Feb. 2023.

[35] P. Thamilselvan, "Lung cancer prediction and classification using AdaBoost data mining algorithm," *International Journal of Computer Theory and Engineering*, vol. 14, no. 4, pp. 149–154, 2022.

[36] S. Gamil, F. Zeng, M. Alrifaey, M. Asim, and N. Ahmad, "An efficient AdaBoost algorithm for enhancing skin cancer detection and classification," *Algorithms*, vol. 17, no. 8, 2024. doi: 10.3390/a17080353

[37] W. Ni, L. Zhao, L. Zhang, K. Xing, and J. Dou, "Coupling progressive deep learning with the AdaBoost framework for landslide displacement rate prediction in the Baihetan Dam Reservoir, China," *Remote Sensing*, vol. 15, no. 9, 2023. doi: 10.3390/rs15092296

[38] S. Masturoh, R. L. Pratiwi, M. R. R. Saelan, and U. Radiyah, "Application of the k-nearest neighbor (KNN) algorithm in sentiment analysis of the OVO e-wallet application," *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, vol. 8, no. 2, pp. 78–83, 2023.

[39] H. Babbar, S. Rani, D. K. Sah, S. A. AlQahtani, and A. K. Bashir, "Detection of Android malware in the internet of things through the k-nearest neighbor algorithm," *Sensors*, vol. 23, no. 16, 2023. doi: 10.3390/s23167256

[40] C. Zhang, Y. Liu, and N. Tie, "Forest land resource information acquisition with Sentinel-2 image utilizing support vector machine, k-nearest neighbor, random forest, decision trees and multi-layer perceptron," *Forests*, vol. 14, no. 2, 2023. doi: 10.3390/f14020254

[41] S. Mohsen, A. Elkaseer, and S. G. Scholz, "Human activity recognition using k-nearest neighbor machine learning algorithm," in *Sustainable Design and Manufacturing*, S. G. Scholz, R. J. Howlett, R. Setchi, Ed., Singapore: Springer, 2021, pp. 304–313. doi: 10.1007/978-981-16-6128-0_29

[42] C. Feng, B. Zhao, X. Zhou, X. Ding, and Z. Shan, "An enhanced Quantum K-Nearest Neighbor classification algorithm based on polar distance," *Entropy*, vol. 25, no. 1, 2023, p. 127. doi: 10.3390/e25010127

[43] A. A. Mahamat, M. M. Boukar, N. Leklou *et al.*, "Decision tree regression vs. gradient boosting regressor models for the prediction of hygroscopic properties of Borassus fruit fiber," *Applied Sciences*, vol. 14, no. 17, 2024, p. 7540. doi: 10.3390/app14177540

[44] R. Sibindi, R. W. Mwangi, and A. G. Waititu, "A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices," *Engineering Reports*, vol. 5, no. 4, 2023. https://doi.org/10.1002/eng2.12599

[45] S. B. Nadkarni, G. S. Vijay, and R. C. Kamath, "Comparative study of random forest and gradient boosting algorithms to predict airfoil self-noise," *Engineering Proceedings*, vol. 24, 2023. doi: 10.3390/engproc2023059024

[46] O. Lyashevska, F. Malone, E. MacCarthy, J. Fiehler, J. H. Buhk, and L. Morris, "Class imbalance in gradient boosting classification algorithms: Application to experimental stroke data," *Statistical Methods in Medical Research*, vol. 30, no. 3, pp. 916–925, 2021.

[47] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 22, no. 5, pp. 717–727, 2000.

[48] M. G. M. Abdolrasol, S. M. S. Hussain, T. S. Ustun *et al.*, "Artificial neural networks based optimization techniques: A review," *Electronics*, vol. 10, no. 21, 2021. doi: 10.3390/electronics10212689

[49] P. Tsikas, A. Chassiakos, V. Papadimitropoulos, and A. Papamanolis, "BIM-based machine learning application for parametric assessment of building energy performance," *Energies*, vol. 18, no. 1, 2025. doi:10.3390/en18010201

[50] A. Almomani, I. Akour, A. M. Manasrah *et al.*, "Ensemble-based approach for efficient intrusion detection in network traffic," *Intelligent Automation & Soft Computing*, vol. 37, no. 2, 2023. doi: 10.32604/iasc.2023.039687

[51] M. A. Cifci, S. Hussain, and P. J. Canatalay, "Hybrid deep learning approach for accurate tumor detection in medical imaging data," *Diagnostics*, vol. 13, no. 6, 2023. doi: 10.3390/diagnostics13061025

[52] D. Božić, B. Runje, D. Lisjak, and D. Kolar, "Metrics related to confusion matrix as tools for conformity assessment decisions," *Applied Sciences*, vol. 13, no. 14, p. 8187, 2023. doi: 10.3390/app13148187

[53] A. Arabiat and M. Altayeb, "Assessing the effectiveness of data mining tools in classifying and predicting road traffic congestion," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, no. 2, pp. 1295–1303, 2024.

[54] A. Arabiat and M. Altayeb, "Enhancing internet of things security: Evaluating machine learning classifiers for attack prediction," *International Journal of Power Electronics and Drive Systems*, vol. 14, no. 5, pp. 6036–6046, 2024.

[55] S. Nafaa, K. Ashour, R. Mohamed *et al.*, "Advancing roadway sign detection with YOLO models and transfer learning," in *Proc. 2024 IEEE 3rd Int. Conf. on Computing and Machine Intelligence*. doi: 10.1109/ICMI60790.2024.10586105

[56] O. Turanlı and Y. B. Yakut, "Classification of faults in power system transmission lines using deep learning methods with real, synthetic, and public datasets," *Applied Sciences*, vol. 14, no. 20,

2024. doi: 10.3390/app14209590

[57] Y. Yang, W. Zhou, I. M. Jiskani, X. Lu, Z. Wang, and B. Luan, "Slope stability prediction method based on intelligent optimization and machine learning algorithms," *Sustainability*, vol. 15, no. 2, 2023. doi: 10.3390/su15021169

[58] G. ALMahadin, H. A. Owida, J. A. Nabulsi, N. Turab, and N. A. Hawamdeh, "Automated detection of kidney masses lesions using a deep learning approach," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 3, pp. 2862–2869, 2024.

[59] A. Arabiat and M. Altayeb, "An automated system for classifying types of cerebral hemorrhage based on image processing techniques," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 2, pp. 1594–1603, 2024.

[60] A. Arabiat and M. Altayeb, "Driving behavior analytics: An intelligent system based on machine learning and data mining techniques," *Bulletin of Electrical Engineering and Informatics*, vol. 14, no. 3, pp. 2055–2065, May 2025.

[61] *Performance Curve—Orange Visual Programming 3 Documentation*. (Dec. 29, 2024). [Online]. Availble: https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/evaluate/performancecurve.html

Copyright © 2025 by the authors. This is an open access article distributed under the Creative Commons Attribution License (CC BY 4.0), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

Areen M. Arabiat obtained the B.Sc. degree in computer engineering in 2005 from Al-Balqa' Applied University (BAU), Jordan, and her M.Sc. degree in Intelligent Transportation Systems (ITS) from Al-Ahliyya Amman University (AAU), Jordan in 2022. She is currently a part-time lecturer and computer lab supervisor at the Faculty of Engineering, AAU since 2013. Her research interests are focused on the following areas: machine learning, data mining, artificial intelligence, IoT, and image processing.