International Journal of Computational and Electronic Aspects in Engineering
Volume 7, Issue 4, August 2026, pp 338 - 349
https://doi.org/10.26706/ijceae.7.2.20260803

A Novel Hybrid InceptionV3-SqueezeNet Network (HIS-Net) Model for Kidney Disease Classification in CT Medical Image

Ruaa H. Ali Al-Mallah*

Department of Artificial Intelligence Engineering Technology, College of Technical Engineering for Computer and AI, Northern Technical University, Mosul 41000, Iraq

*Correspondence Author : ruaa_almallah@ntu.edu.iq

Abstract: Accurate and early diagnosis of kidney diseases is essential for improving healing processes. This study proposes a novel hybrid deep learning model, called HIS-Net (Hybrid InceptionV3 - SqueezeNet), to enhance CT images classification of kidney disease on a dataset of 5,600 colorized samples across four categories: Tumor, Cyst, Stone, and Normal. A key novelty of this study is integrating parallel feature extraction by reducing dimensions of features. This improves efficiency and reduces computational complexity while preserving practical classification power. The framework incorporates preprocessing steps, including resizing, augmentation, normalization, grayscale conversion, and dataset partitioning. Two pre-trained CNNs, InceptionV3 and SqueezeNet, are employed in parallel as balancing feature extractors. Principal Component Analysis (PCA) is applied to each feature set, reducing them to 90 components, followed by feature fusion into a compact 180-dimensional representation. The fused features are evaluated using three machine learning classifiers, which are support vector machine (SVM), logic regression (LR), and naïve bayes (NB), under 10-fold cross-validation. The results show that grayscale images perform better than pseudo-color representations. SVM achieved the best performance with 98.5% accuracy and AUC of 1.000, with tumor sensitivity of 99.2%. These findings confirm that HIS-Net is an accurate, efficient, and reliable computer-aided diagnosis system.

Keywords: Hybrid deep network, InceptionV3, SqueezeNet, kidney disease, medical image classification, Support Vector Machine.

Article – Peer Reviewed
Received: May 16, 2026
Accepted: July 23, 2026
Published: August 11, 2026

Cite this article: Ruaa H. Ali Al-Mallah, "A Novel Hybrid InceptionV3-SqueezeNet Network (HIS-Net) Model for Kidney Disease Classification in CT Medical Image", *International Journal of Computational and Electronic Aspects in Engineering*, RAME Publishers, vol. 7, issue 4, pp. 338-349, 2026.
https://doi.org/10.26706/ijceae.7.4.20260803

1. Introduction

The body needs to filter blood, control blood pressure, and get rid of waste. The kidneys are the main organs that do all of these functions. Kidney disease (KD) is a big health problem around the world. Kidney stones, tumors, and cysts, are all common kidney problems that can cause kidney failure. Discovering diseases early is very important for stopping their spread and decrease death rates. Medical imaging methods like computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and X-rays are very important for finding kidney problems [[1][2][3]]. But doing it conventionally takes a long time and is easy to make mistakes. Recent advances in Artificial Intelligence (AI), particularly deep learning (DL), have demonstrated potential for medical image analysis [[4][5][6][7]]. AI-based models can help doctors make more accurate diagnoses, lighten their workloads, and find kidney diseases earlier.

Advances in medical image processing using machine learning and deep learning have contributed to the detection and diagnosis of kidney diseases in their early stages. However, a certain percentage of error in classification using artificial intelligence techniques still exists in previous work. Efforts have been made to reduce these errors and make the methods more reliable. The analysis presented in [[8][9][10][11][12]] has shown a high level of Earlier research utilized conventional methods of machine learning. According to Attia et al. [13] (2015), the Principal Component Analysis (PCA) framework and artificial neural networks were used to classify ultrasound images into normal kidney, cyst, stone, tumor, and kidney with renal failure, achieving a 97% success rate. Martinez et al. [14] (2020), employed Random Forest (RF) and ensemble KNN models for classifying kidney stones from ureteroscopy images with 89% accuracy. In 2023, Nadeem et al. [15] used EANet, InceptionV3, and SqueezeNet networks to detect kidney stones. They concluded that the best accuracy was achieved using InceptionV3 of (96.14%) which exceeded the other networks slightly. Gill et al. [16] applied MobileNetV3 to classify X-ray images of chronic kidney disease, stones and tumors, and achieved a maximum accuracy of 97%.

Conversely, Maniyar et al. [17] used CNNs to analyze CT images to distinguish between healthy kidneys, cysts, tumors, and kidney stones and reported the highest accuracy of 92.01%. Also, in the year 2024, Yazhini et al. [18] combined machine learning and deep learning techniques for kidney tumors' classification, resulting in an optimal accuracy of 92%. Furthermore, Jain et al. [19], proposed a Convolution Neural network-Support Vector Machine (CNN-SVM) hybrid model for classifying kidney diseases, achieving 90.91% accuracy and an F1 Score of 91.95%. Moreover, in 2025, Shreif et al. [20], evaluated EfficientNetV2B0 and ResNet50V2 in classifying sponge kidney, kidney cysts, stones, and tumors. They managed to attain 97% accuracy using EfficientNetV2B0 and 96% using ResNet50V2. Shirdhone and Chede [21], compared five types of CNNs (EfficientNet, DenseNet, ResNet, InceptionNet, and VGGNet) in detecting kidney abnormalities. In their work, EfficientNet had the best performance, with 96.4% accuracy and an F1-score of 95.7%.

Motivated by the limitations of present methods, this study proposes an efficient hybrid model for kidney disease classification. The method mixes complementary deep feature extraction using InceptionV3 and SqueezeNet to capture more discriminative representations. PCA is then applied to reduce feature dimensionality and eliminate redundancy. The optimized features are subsequently evaluated using SVM, Logic Regression (LR), and Naive Bayes (NB) to determine the most effective classifier. In addition, the impact of different kidney image representations on classification performance is systematically analyzed. This framework aims to enhance accuracy, efficiency, and generalization compared to existing approaches. A comprehensive summary of the reviewed related works is presented in Table 1, it showing models, image types, classification categories, accuracies, and key techniques.

2. Methodology and Materials

The proposed classification methodology objectives to classify renal pathologies into one of four groups: normal, cyst, tumor, and stone. The suggested method utilized feature-level fusion to improve classification accuracy. This approach is based on utilizing deep features extracted from medical images using InceptionV3 and SqueezeNet architectures. The extracted features from both models are fused to complement each other and enhance the overall classification performance. In evaluation model, both grayscale and pseudo-color image datasets will be used. The pseudo-color images will be generated using a deterministic false-color technique to enhance contrast and make structural variations more visible. Dimensionality reduction of the extracted image features will be performed using PCA. SVM, NB, and LR classifiers will be evaluated.

The evaluation process will include 10-fold cross-validation using both grayscale and pseudo-color image datasets. The details of the proposed work will be explained in the following paragraphs, as illustrated in Fig. 1.

Table 1. Summary of the reviewed related works

Authors	Models	Data Type	Techniques	Accuracy
M. W. Attia et al. [13]	Neural Network, PCA	Ultrasound Normal, cyst, stone, tumor, renal failure	Wavelet features, PCA	97%
A. Martinez et al. [14]	Random Forest, KNN Ensemble	Ureteroscopy: Kidney stones	Feature encoding, ensemble models	89%
M. Nadeem et al. [15]	EANet, InceptionV3, SqueezeNet	CT stones	Deep learning image classification	96.14% (InceptionV3)
K. S. Gill et al [16]	MobileNetV3	X-ray:CKD, stones, tumors	Deep learning classifier	97%
M. F. Maniyar et al. [17]	CNN	CT: Normal, cysts, tumors, stones	Adam optimizer, deep learning	92.01%
A. Yazhini et al [18]	Decision Tree, Random Forest, KNN, CNN	CT / MRI: Tumor	Comparative ML and DL analysis	92%

R. Jain et al. [19]	CNN + SVM	CT: Kidney disease (binary)	Hybrid CNN-SVM model	90.91%
S. A. Shreif et al. [20]	EfficientNetV2B0, ResNet50V2	CT: Sponge kidney, cysts, stones, tumors	Deep learning classification	97%EfficientN etV2B0 96% ResNet50V2
T. Shirdhone et al. [21]	EfficientNet, DenseNet, ResNet, InceptionNet, VGGNet	CT: Kidney abnormalities	Data augmentation, consistent training	96.4% Efficient Net
Proposed method (HIS-Net)	Hybrid InceptionV3, SqueezeNet +SVM	CT: Cyst, stone, tumor, normal	HIS-Net + PCA	98.5% (HIS-Net) +SVM

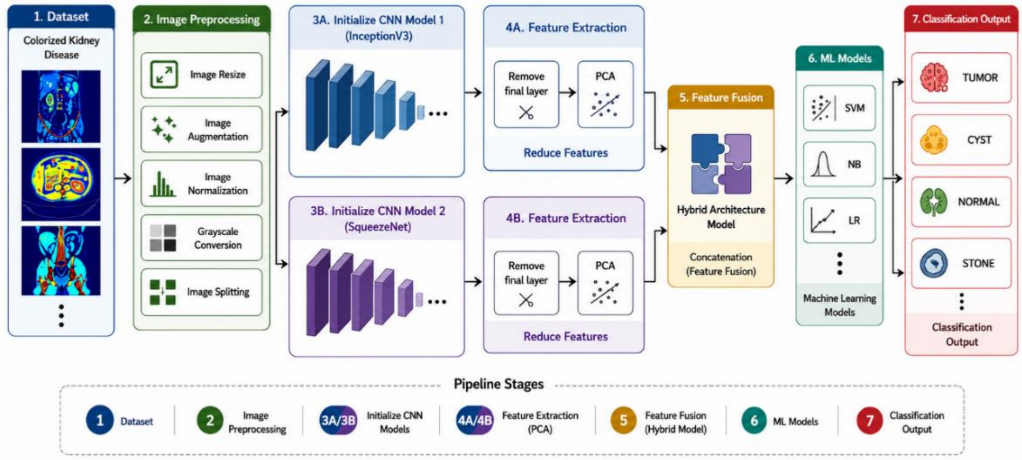

Figure 1. Flowchart of Hybrid InceptionV3-SqueezeNet (HIS-Net) model

2.1 Dataset

A colorized kidney CT dataset was used, consisting of four classes: normal, cyst, tumor, and stone [22]. The data were collected from multiple hospitals using a PACS system. This dataset involving 12,446 coronal and axial CT scans with and without contrast. To ensure balanced training and reliable evaluation, a subset of 5,600 images was selected, with an equal number of samples for each class as illustrated in Fig.2. The dataset was splitting into (1000,200,200) images per class for training, validation, and testing set respectively.

The data initially had a RGB image representation after thorough processing at the source that involved noise reduction, segmentation of lesions, and enhancement of image contrast using CLAHE [22]. Then, a deterministic pseudo colorization process was carried out using the PLASMA colormap. Under the scheme, regions with low intensity values would be colored in dark purples and blues, while regions with medium intensity will transition from magenta to orange.

Intense structures like kidney stones would be represented using yellows and whites [22] as illustrated in Fig. 3. This dataset provides diverse and well-structured samples, enabling robust evaluation of the proposed hybrid classification framework

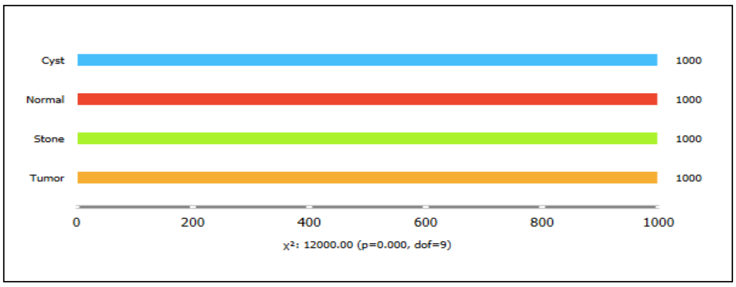

Figure 2. Class distribution of the CKD Dataset

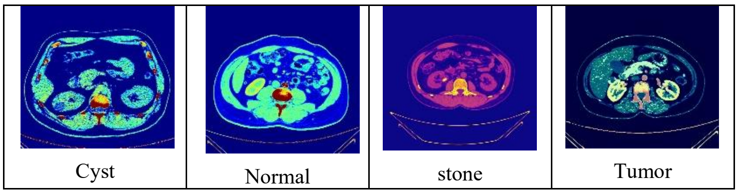

Figure 3. Random Samples of Colorized (pseudo-color) KD dataset [22]

2.2 Colorized and grayscale images:

The pseudo-color is used color mapping techniques (such as a heatmap or the Jet color map). The benefit of this colorization is enhanced photographic interpretability Pesdeuo-color is used in color mapping techniques (such as heat mapping or Jet color mapping) as illustrated in figure (3). Its benefit lies in improving the interpretability of medical images. Pseudo-coloring helps highlight subtle differences in organ density, thus facilitating visual diagnosis and easier interpretation of results. Therefore, it is medically important. In this research, these puseduo-colored images were analyzed and trained on the suggestion model, and their diagnostic accuracy was recorded. The same images were then converted to grayscale as illustrated in Figure (4). This conversion is completed using a luminance-based weighted sum of RGB channels, consistent with ITU-R BT.601 standard and fed into the same model. The following sections will explain the importance of colored and grayscale images in medical diagnosis and identify the results obtained in both cases.

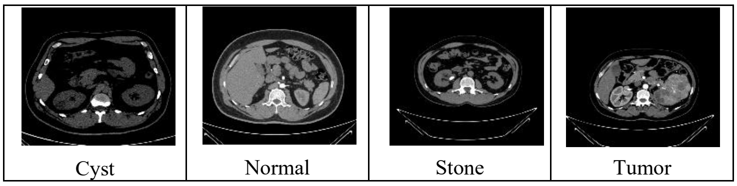

Figure. 4 Random Samples of Grayscale image KD dataset

2.3 K-fold cross validation strategy

The 10-fold cross-validation strategy was used to evaluate this proposed model. This strategy is based on the principle of dividing the data into ten equal subsets. In each subset, one part is tested, while the remaining nine parts are trained. This process is repeated ten times to ensure that all samples contribute to both the training and testing phases.

10-fold cross-validation is improved the reliability of the evaluation by reducing variance and limiting over-allocation, resulting in a more accurate and objective performance assessment [23].

2.4 Deep feature extraction

2.4.1. SqueezeNet

SqueezeNet [24] was chosen for the suggested architecture because of its lightweight design and effective use of resources. It has 1.24 million trainable parameters. This is very suitable for medical imaging because it is important to keep costs low while still getting accurate results. The core component of this model is the Fire module. It consists of two sequential stages: a squeeze and expand layers as illustrated in Figure (5). The SqueezeNet layer uses the 1 × 1 convolutional filter to decrease the number of input channels to reduce the complexity of computation. Then, the data is processed through the expand layer via two parallel streams using the 1 × 1 Filter and 3 × 3 Filters. The network receives 224 x 224 x 3 input images. it applies max-pooling are applied relatively late to preserve spatial information for as long as possible. The output is converted into a 14 x 14 feature map with 512 channels at the end of the pipeline. The feature maps are converted into a compact one-dimensional vector of size 1000-dimensions using the global average pooling (GAP) layer. It helps to model of nonlinear relationships between pixel intensities and underlying pathological patterns [24].

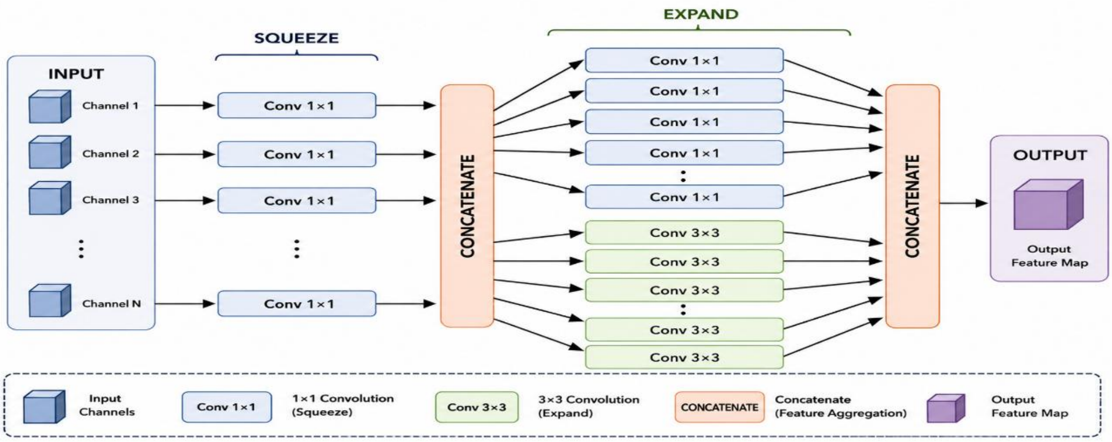

Figure 5. SqueezeNet architecture with a squeeze and expand process.

2.4.2. InceptionV3

The InceptionV3 model is used in this project as a feature extraction network [25]. It is a deep convolutional neural network introduced by Google Research that seeks to enhance the performance of CNNs in image classification and detection tasks. As illustrated in figure 6 , Input images will be preprocessed by resizing them to 299 × 299 × 3 before passing them through this network. This neural network has 42 layers and was pretrained on ImageNet with 1,000 categories of images. The network architecture is based on three types of inception blocks (inception A, B, and C) as illustrated in figure (6). These blocks have a several convolutional filters with multiple kernel sizes (1×1, 3×3, and 5×5) are applied in parallel to generate higher level representations of data (features).

The network has multiple layers such as convolutional and max-pooling layers which are utilized to extract features. Both batch normalization and ReLU activations are employed to stabilize the training process. Rather than adopting the last layer comprising of fully connected units and softmax classifier, this network is designed to extract high-level feature representations in this research. The output from the GAP layer is (2048 dimensional feature representation) will be considered as the final feature representation [25].

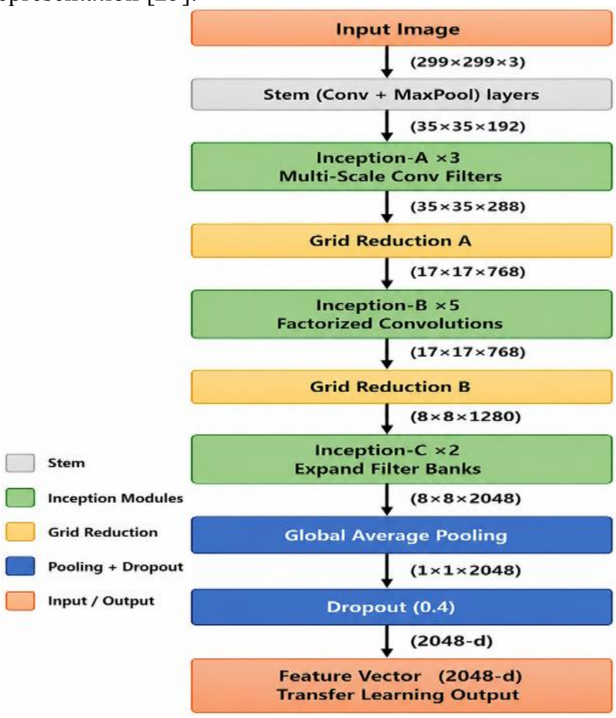

Figure 6. Inception V3 structure [25]

2.5 HIS-Net model

Hybrid (fusion-based) deep learning methods involve concatenation of feature extraction from various networks to produce more accurate and robust predictions. There are three types of fusion processes include late fusion that concatenates the results of various models, intermediate fusion that concatenates learned representations at the hidden layer stage, and early fusion that concatenates raw data inputs. The proposed HIS-Net take part deep features extracted from SqueezeNet and InceptionV3 using a feature-level fusion strategy. The concatenated feature vector is further optimized via PCA . PCA is applied to the feature vectors InceptionV3 and SqueezeNet and reducing their dimensions from (2048 and 1000) to 90 components respectively, the two feature vectors are then merged into one feature vector through concatenation. This step falls under feature level fusion, where the two vectors are merged directly without the need for linear transformations or weighing.

For mathematical illustration, if we have $f_1 \in R^{90}$ and $f_2 \in R^{90}$ as the feature vectors of the first and second branches respectively, the merged feature vector f_{fused} can be expressed as :

$$f_{fused} = f_1 \oplus f_2 = [f_{1,1}, f_{1,2}, \dots.., f_{1,90}, f_{2,1}, f_{2,2}, \dots.., f_{2,90}]^T \in R^{180} \quad (1)$$

In this case, $\oplus$ represents the concatenation operation.

With such an arrangement, there will be no loss of information from either stream of data. The ML classifier will be able to learn any relationship that exists between the two sets of features obtained from each network. Unlike other types of fusion methods that depend on averaging or summing feature vectors. concatenation does not entail the loss of any features. As such, it is well suited for cases where redundancy is low. The resulting 180-dimensional vector is then fed into classifier for final classification. this hybrid network is used in complex medical image application.

2.6 Classification algorithm in ML

There are many techniques used in data classification and prediction, the most important of which are traditional machine learning techniques and deep learning models. To improve the results of this study, traditional machine learning techniques were combined with feature extraction from deep learning models. To clarify the mechanism of these techniques, the traditional machine learning techniques used in this study are explained.

- ***SVM:*** This technique is one that classifies data in supervised learning with high-dimensional information. In this study, SVM was utilized to classify kidney images into different disease groups. The efficiency of the classification process depends on the quality of the dataset, and the results were improved using a nonlinear kernel known as a polynomial kernel [26] .
- ***LR:*** This algorithm is utilized for classification tasks. It is estimated fitting input data by sigmoid/logistic function in order to map any real number value to a range between 0 and 1[27] . The predicted probability is consider as in:

$$P(x) = \frac{e^{(b0+b1*x)}}{1+e^{(b0+b1*x)}} \quad (2)$$

Where the predicted output is $p(x)$, The intercept term is $b0$, and $b1$ the coefficient for the single input value (x)

- ***NB:*** It refers to a statistical model that uses Bayes' theorem in estimating the probability of a certain class based on given input variables. Naïve Bayes model assumes independent input variables and calculates posterior probability as in:

$$\mathrm{P}\ (\mathrm{A|B}) = \frac{\mathrm{P}\ (\mathrm{B|A})\mathrm{P}\ (\mathrm{A})}{\mathrm{P}\ (\mathrm{B})} \quad (3)$$

Consider the probability of the class to be P(A), probability of the evidence given the class to be P(B|A), and probability of the evidence to be P(B). The model can classify efficiently because it maximizes the posterior probability[28],[29].

ML models have many hyperparameters which are used to evaluate system design. In SVM model have been applied several activation functions such as Radial Basis Function (RBF), Sigmoid, and polynomial , the best result got with activation function polynomial with iteration is 1000, Cost is 0.7 and regression loss (ε) is 0.1. In LR model designed by using Regularization Ridge (L2) with cost is 10 and without weights. The best parameters for all Models which are used in this study shown in Table (2).

Table 2: Summary hyperparameters of all algorithms in our system

ML Algorithm	Parameters
SVM	Kernel Polynomial, c=0.7, ε=0.1, Iterations: 1000
LR	Regularization type: Ridge(L2), c=10, class weights=False
NB	Laplace Smoothing (alpha= 1), Gaussian/Discrete), Instance Weights = disabled

2.7 Pre-Fusion Dimensionality Reduction:

Data reduction was used to improve the classification process with DL and traditional ML algorithms (SVM, NB, and LR), ensuring accurate results during operation. Both the InceptionV3 and SqueezeNet networks extract high-dimensional feature vectors: InceptionV3 produces 2048-dimensional features, while SqueezeNet produces 1000-dimensional features, which may contain redundant information. PCA was utilized to reduce the high dimensionality of the extracted features and eliminate noise from the input data [30]. The features were reduced to 90 components using PCA. Pre-fusion dimensionality reduction helps reduce computational complexity without compromising the essential information in the data. The number 90 was chosen because it accounts for approximately 95% of the total variance in each feature set.

2.8 Evaluation performance

The performance of the proposed model for kidney disease classification was assessed using the following metrics:

- **Confusion Matrix:** Confusion matrix helping as a foundational tool in evaluation method. The confusion matrix is represented four sections:
  - True positive (TP): is predicted correct and actual is positive.
  - False positive (FP): is predicted incorrect and actually actual is positive
  - False negative (FN): is predicted incorrect and actually actual is negative.
  - True negative (TN): is predicted correct and actually actual is negative.
- **Accuracy:** determine the total amount of correct predictions as :

$$\text{Accuracy} = \frac{T_p+T_N}{T_p+T_N+F_p+F_N} \tag{4}$$

- **Recall:** refer to the number of actual positives correctly identified as:

$$\text{Recall (Sensitivity)} = \frac{T_P}{T_P+F_N} \tag{5}$$

- **Precision:** determine the proportion of positive predictions that are correct, important when false positives are costly.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{6}$$

- **F1-score:** consonant mean of precision and recall. The better balance between precision and recall when the higher value of F1-score.

$$F1 = 2\ x\ \frac{\text{Precision x Recall}}{\text{Precision+Recall}} \tag{7}$$

- **AUC-ROC:** Evaluates the predictive performance of the model to distinguish between positive and negative classes utilized the area under the curve (AUC) of the receiver operating characteristic (ROC). The ROC curve is representing true positive rate and false positive rate with range 0.5 to 1 and perfect classification, when AUC values is close to 1.

3. Results and Discussion

The aim of this study is to achieve high accuracy in classifying kidney diseases including cysts, stones, and tumors from normal kidneys. The experiments were executed and evaluated using Python on an open-source platform prepared with an Intel Core i7-13620H processor operating at 2.40 GHz and an NVIDIA GeForce RTX 4050 GPU The study utilized two types of images: grayscale and pseudo-color, for result comparison. All models were evaluated using stratified 10-fold cross-validation to ensure a statistically reliable performance assessment. The proposed model was evaluated using the ROC curve, confusion matrix, and several performance metrics to provide comprehensive insights into the effects of feature extraction and data characteristics on classification accuracy and overall model effectiveness.

3.1 Performance Analysis with Pseudo-Color Images

The evaluation results of performance classification models with Pseudo-color kidney images are illustrated in Table 3. In the first stage of the experiments, separate DL feature extractors, InceptionV3 and SqueezeNet, were combined with ML classifiers (SVM, NB, and LR). The best result was recorded using SVM with SqueezeNet features, obtaining a classification accuracy (CA) of 95.7%, while InceptionV3 achieved 94%. The better performance of SqueezeNet is attributed to its more compact and discriminative feature representation, whereas InceptionV3 produced higher-dimensional features with some redundant information. To improve kidney disease classification, the proposed HIS-Net achieved the highest accuracy of 97.9%, indicating that feature fusion and dimensionality reduction enhance classification performance and overall feature representation. Meanwhile, NB and LR showed lower performance across all experiments.

Table 3: Evaluation of ML classifiers using deep feature extraction techniques in Pseudo-color image analysis

Feature extraction	Classifier	AUC	CA(%)	F1	Recall
InceptionV3 (2048 features)	SVM	0.993	94	0.94	0.940
	NB	0.886	66.7	0.66	0.667
	LR	0.94	77.6	0.77	0.776
SqueezeNet (1000 features)	SVM	0.995	95.7	0.95	0.957
	NB	0.865	64.4	0.64	0.644
	LR	0.932	75.1	0.75	0.751
HIS-Net (180 features)	SVM	0.999	97.9	0.97	0.979
	NB	0.921	73.7	0.73	0.737
	LR	0.970	85.5	0.85	0.855

3.2 Performance with Grayscale Images

The same techniques are used with grayscale images. The pre-trained CNNs are recorded the best result with SVM classifiers. SqueezeNet got CA is 96.7% and AUC of 0.998, while InceptionV3 achieved 95.6% accuracy and 0.996 AUC as illustrated in Table 4. While the proposed model (HIS-Net) used with SVM got accuracy CA of 98.5% and a perfect AUC of 1.000. The results proved that (HIS-Net) achieved the best performance with a minimized feature set and the original grayscale images. The use of pseudo-color images offered no additional benefit, suggesting that preserving the native grayscale information is more effective for accurate classification**.** The improvement in detecting kidney diseases with grayscale images, which have true color intensity gradations in CT scans, is crucial for accurate medical interpretation. However, when pseudo-coloring techniques were used, they added color representations to the color intensity gradations, distorting the original information and affecting the reliability of the results

Table 4. Evaluation of ML classifiers using deep feature extraction techniques in grayscale image analysis

Feature extraction	Classifier	AUC	CA(%)	F1	Recall
InceptionV3 (2048 features)	SVM	0.996	95.6	0.95	0.956
	NB	0.881	67.4	0.67	0.674
	LR	0.939	76.6	0.76	0.766
SqueezeNet (1000 features)	SVM	0.998	96.7	0.96	0.967
	NB	0.886	68.3	0.68	0.683
	LR	0.941	77.5	0.77	0.775
HIS-Net (180 features)	SVM	1.000	98.5	0.98	0.985
	NB	0.928	75.3	0.75	0.753
	LR	0.971	86.1	0.86	0.861

3.3. Evaluation HIS-Net with SVM classifier

The efficiency of the proposed model (HIS-Net) was evaluated based on ROC curves and confusion matrices in the pseudo color and grayscale cases. Figure 7 and Figure 8 , illustrate ROC curves built through HIS-Net and SVM classifier with four cases (Tumor, Stone, Cyst, and Normal). The analysis shows that the performance of SVM is provided a perfect AUC values (0.9988–0.9998) and representing its superiority over LR and NB classifiers.

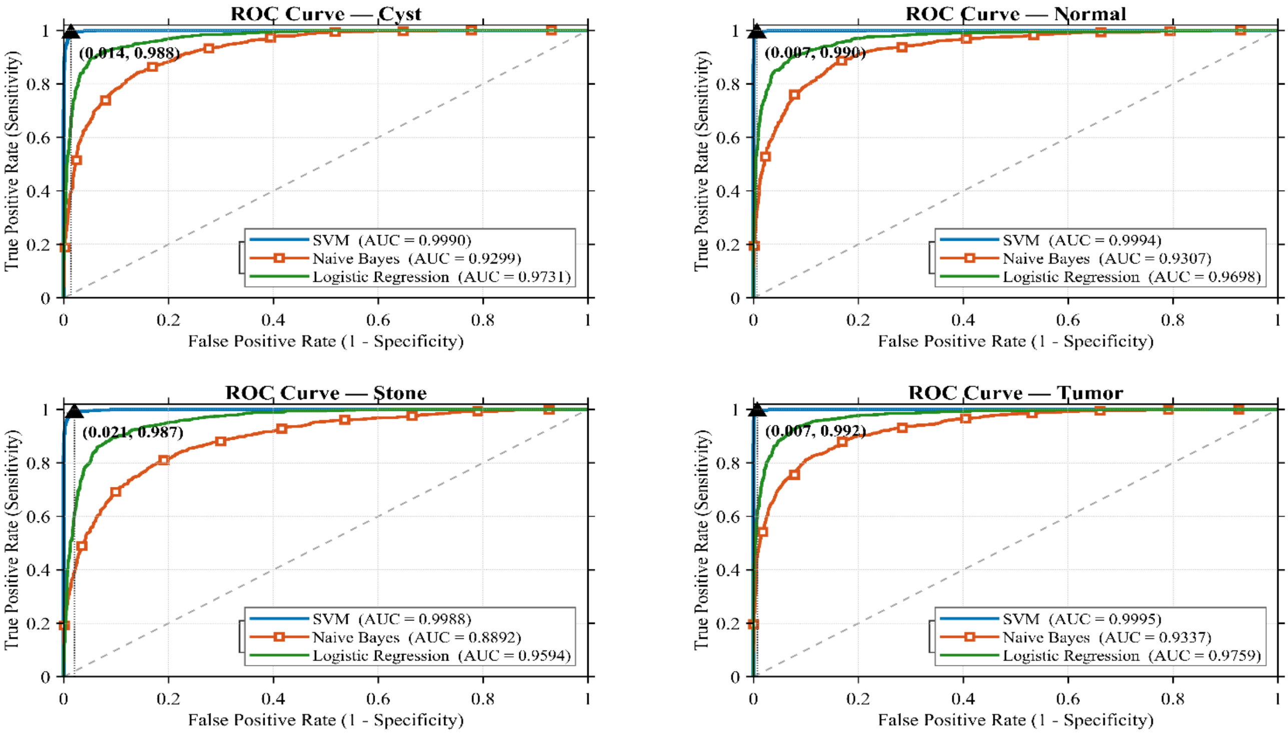

Figure 7. ROC curves for evaluated HIS-Net and SVM model in Pseudo -color

Although Pseudo-color representations were used to increase the clarify of data, the HIS-Net was able to extract important features from grayscale images with better results. The HIS-Net insensitive to color conversions as illustrated in the classification of tumors, SVM classifier built using grayscale images provided AUC at 0.9998 with the FPR of 0.003, while the one based on the input of pseudo-color images produced an AUC of 0.9995 with an FPR of 0.007. Even though the gap between those two performance values is small, the former proved to provide more precise results and suggests that color does not contribute significantly to classification accuracy, whereas morphology and textures play a much larger role.

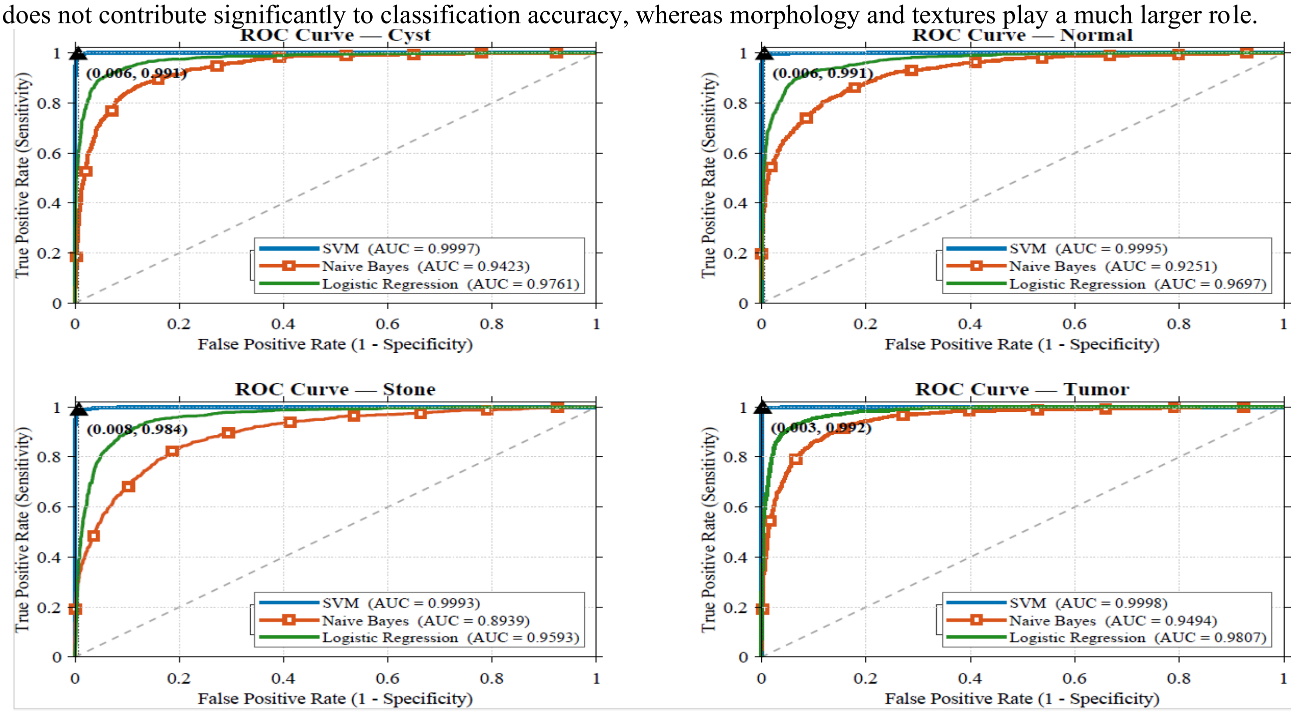

Figure 8. ROC curves for evaluated HIS-Net with SVM model in grayscale

From the analysis of the confusion matrix, it is clear that the classifier SVM outperforms all other models (Figure 9 and Figure 10). the superiority of the SVM classifier can be verified by its high TPR which was (98.7% for Cyst, 98.3% for

Normal, 97.8% for Stone, and 99.2% for Tumor) and (97.1% for Cyst, 98.3% Normal , 97.2% for Stone and 99.1% for Tumor) in grayscale images and pseudo color images respectively. The error rate is relatively low for SVM classifier and occurs only when there is a misclassification between classes with similar visual characteristics, like Cyst and Stone classes. On the other hand, Logistic Regression classifier shows reasonable performance (82.1% to 88.9%) due to a higher rate of confusion in classifying objects since it is unable to accommodate non-linearities in features.

Finally, the naïve Bayes classifier shows poor performance and lower accuracy rates (67.5% to 80.1%) and confusion among different classes with misclassification of about 32.8% of Stone as Cyst or Normal.

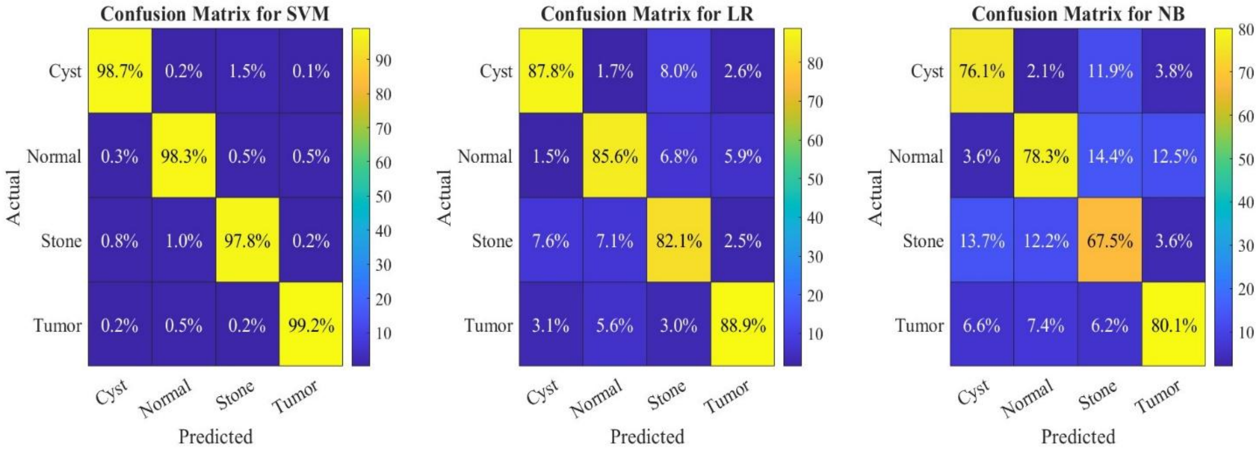

Figure 9. Confusion Matrix of HIS-Net with three models in gray scale images.

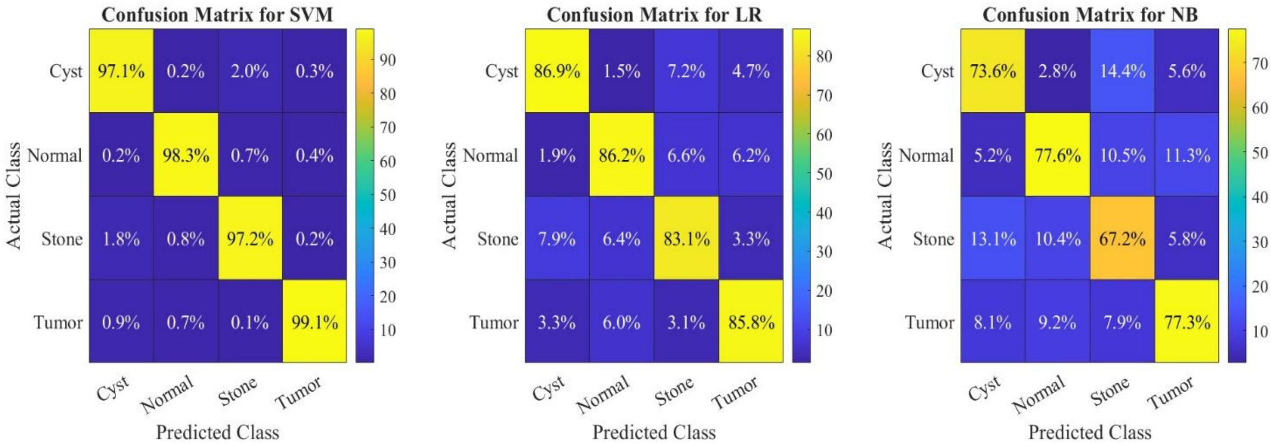

Figure 10. Confusion matrix of HIS-Net with three models pseudo-color images

4. Conclusions

The study proves the efficacy of our proposed algorithm (HIS-Net) to automatically classify kidney diseases. The comparative analysis between Pseudo-color and Grayscale imaging modes exposes a dependable performance lead for the grayscale-based models across all evaluated architectures. Specifically, the proposed model (HIS-Net) and SVM reached high performance using Grayscale images, achieving a perfect AUC of 1.000 and an Accuracy of 98.5%, surpassing its performance on Pseudo-color images (97.9%).

While other models such as Naive Bayes were unable to differentiate between classes owing to their complex overlaps, the SVM algorithm succeeded in separating them with hyperplanes. While pseudo-color representations improve visual contrast, grayscale images provide more accurate and stable results by preserving original intensity information. The fusion of deep features further enhances performance by integrating complementary information and reducing redundancy. Overall, the proposed framework provides a robust and reliable solution for medical image classification and computer-aided diagnosis.

References

[1] A. S. Reddi and K. Kuppasani, “Kidney Function in Health and Disease,” in *Nutrition in Kidney Disease*, Totowa, NJ: Humana Press, pp. 3–15. doi: 10.1007/978-1-59745-032-4_1.

[2] S. Sudharson and P. Kokil, “An ensemble of deep neural networks for kidney ultrasound image classification,” *Comput. Methods Programs Biomed.*, vol. 197, p. 105709, Dec. 2020, doi: 10.1016/j.cmpb.2020.105709.

[3] K. Yildirim, P. G. Bozdag, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, “Deep learning model for automated kidney stone detection using coronal CT images,” *Comput. Biol. Med.*, vol. 135, p. 104569, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104569.

[4] S. Qasim Hasan, “Shallow Model and Deep Learning Model for Features Extraction of Images,” *NTU Journal of Engineering and Technology*, vol. 2, no. 3, pp. 1–8, Nov. 2023, doi: 10.56286/ntujet.v2i3.449.

[5] N. Waisi, “Comparative Evaluation of Resnet-50 and Efficientnet-B1 for Pneumonia Detection in Chest X-Ray Images Using Transfer Learning,” *International Journal of Computational and Electronic Aspects in Engineering*, vol. 6, no. 2, pp. 89–97, Jun. 2025, doi: 10.26706/ijceae.6.2.20250405.

[6] N. Waisi, “Deep Feature Fusion Method for Images Classification,” *International Journal of Computational and Electronic Aspects in Engineering*, vol. 5, no. 4, pp. 148–153, Dec. 2024, doi: 10.26706/ijceae.5.4.20241103.

[7] M. K. Yousif, “Enhanced Diabetic Retinopathy Classification using Hybrid Efficient NetB0 with CBAM,” *International Journal of Computational and Electronic Aspects in Engineering*, vol. 7, no. 2, pp. 176–188, Jun. 2026, doi: 10.26706/ijceae.7.2.20260307.

[8] T. M. Kadhim, “Deep Convolutional Neural Network–Based Image Processing for Automated Mammography Analysis and Breast Cancer Detection,” *International Journal of Computational and Electronic Aspects in Engineering*, vol. 6, no. 4, pp. 243–251, Dec. 2025, doi: 10.26706/ijceae.6.4.20250904.

[9] A. M. Hadi, “Enhancing MRI Brain Tumor Classification with a Novel Hybrid PCA+RST Feature Selection Approach: Methodology and Comparative Analysis,” *International Journal of Computational and Electronic Aspects in Engineering*, vol. 5, no. 3, pp. 116–130, Sep. 2024, doi: 10.26706/ijceae.5.3.20240805.

[10] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[11] C. Szegedy *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2015, pp. 1–9. doi: 10.1109/CVPR.2015.7298594.

[12] H. A. Ahmed and E. A. Mohammed, “Using Artificial Intelligence to classify osteoarthritis in the knee joint: Review,” *NTU Journal of Engineering and Technology*, vol. 1, no. 3, Sep. 2022, doi: 10.56286/ntujet.v1i3.155.

[13] M. Wagih, F. E. Z. Abou-Chadi, H. El-Din, and N. Mekky, “Classification of Ultrasound Kidney Images using PCA and Neural Networks,” *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 4, 2015, doi: 10.14569/IJACSA.2015.060407.

[14] A. Martinez *et al.*, “Towards an automated classification method for ureteroscopic kidney stone images using ensemble learning,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, Jul. 2020, pp. 1936–1939. doi: 10.1109/EMBC44109.2020.9176121.

[15] M. Nadeem, G. Tan, M. Altaf, M. K. Mumtaz, A. Fatima, and Noshi, “Evaluation and Classification of Kidney Stone Detection Using Deep Learning Techniques,” in *2023 6th International Conference on Software Engineering and Computer Science (CSECS)*, IEEE, Dec. 2023, pp. 01–06. doi: 10.1109/CSECS60003.2023.10428612.

[16] K. S. Gill, V. Anand, R. Chauhan, P. Singh, and R. Gupta, “Detection of Kidney Disease in X-Ray Images Using Machine Learning on MobileNetV3 CNN Model,” in *2023 2nd International Conference on Futuristic Technologies (INCOFT)*, IEEE, Nov. 2023, pp. 1–4. doi: 10.1109/INCOFT60753.2023.10425269.

[17] Mohammad Faizan Maniyar, Yash Nagtode, and Bhalchandra Patil, “Kidney Disease Classification,” *International Journal of Advanced Research in Science, Communication and Technology*, pp. 542–546, Oct. 2024, doi: 10.48175/IJARSCT-19981.

[18] Y. A, M. K. Devi, M. Anitha, R. Praveena, M. K. Mol, and J. Lawrence, “Comparative Analysis of Deep Learning And Machine Learning Analysis For Kidney Tumor Analysis,” in *2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)*, IEEE, Aug. 2024, pp. 1554–1558. doi: 10.1109/ICCPCT61902.2024.10673250.

[19] R. Jain, V. Kukreja, S. Chattopadhyay, A. Verma, and R. Sharma, “Deep Learning Insights: Evaluating the Efficacy of CNN-SVM for Accurate Kidney Disease Identification,” in *2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)*, IEEE, Jun. 2024, pp. 1–5. doi: 10.1109/ICITEICS61368.2024.10625337.

[20] S. A. Shreif, O. F. Nassif, and M. K. Madi, “Kidney Disease Detection and Classification Based on CT Scan Images Using Deep Learning,” in *2025 Sixth International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)*, IEEE, Sep. 2025, pp. 1–6. doi: 10.1109/ACTEA66485.2025.11189946.

[21] T. Shirdhone and S. D. Chede, “Performance Evaluation of CNN Models for Kidney Abnormality Detection: A Comparative Study,” in *2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)*, IEEE, Jun. 2025, pp. 1–6. doi: 10.1109/ICICKE65317.2025.11136743.

[22] S. K. Basak, “Kidney CT Colorized (Normal, Cyst, Tumor, Stone),” 2025, *Kaggle*. doi: 10.34740/KAGGLE/DSV/11895398.

[23] T. Abedin, H. Xu, and S. Uddin, “The impact of K selection in K-fold cross-validation on bias and variance in supervised learning models,” *Sci. Rep.*, vol. 16, no. 1, p. 6084, Jan. 2026, doi: 10.1038/s41598-026-37247-x.

[24] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,” *arXiv preprint arXiv:1602.07360*, 2016.

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.

[26] S. H. Shetty, S. Shetty, C. Singh, and A. Rao, “Supervised Machine Learning: Algorithms and Applications,” in *Fundamentals and Methods of Machine and Deep Learning*, Wiley, 2022, pp. 1–16. doi: 10.1002/9781119821908.ch1.

[27] D. Saini, T. Chand, D. K. Chouhan, and M. Prakash, “A comparative analysis of automatic classification and grading methods for knee osteoarthritis focussing on X-ray images,” *Biocybern. Biomed. Eng.*, vol. 41, no. 2, pp. 419–444, Apr. 2021, doi: 10.1016/j.bbe.2021.03.002.

[28] K. J. D’Souza and Z. Ansari, “Big Data Science in Building Medical Data Classifier Using Naïve Bayes Model,” in *Proceedings of the 7th IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*, 2018, pp. 76–80. doi: 10.1109/CCEM.2018.00020.

[29] N. Nawful Maree, “A Classification Model to Detect Malicious Software in QR Code,” *NTU Journal of Engineering and Technology*, vol. 5, no. 1, Mar. 2026, doi: 10.56286/ngcs4e85.

[30] K. Pearson, “LIII. *On lines and planes of closest fit to systems of points in space*,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, Nov. 1901, doi: 10.1080/14786440109462720.