ISSN: 2093-5374 / E-ISSN: 2093-5382

Comparative Analysis for Feature Extraction and Prediction of CKD Using Machine Learning

K. Afnaan[1*], Dr. Peeta Basa Pati[2], Dr. Tripty Singh[3], and Dr.K.N. Bhanu Prakash[4]

[1*]PhD Research Scholar, Department of Computer Science and Engineering, Amrita School of Computing, Bengaluru, Amrita Vishwa Vidyapeetham, India. afnaankhadar06@gmail.com, https://orcid.org/0009-0009-5870-1444

[2]Professor, Department of Computer Science and Engineering, Amrita School of Computing, Bengaluru, Amrita Vishwa Vidyapeetham, India. bp_peeta@blr.amrita.edu, https://orcid.org/0000-0003-2376-4591

[3]Associate Professor, Department of Computer Science and Engineering, Amrita School of Computing, Bengaluru, Amrita Vishwa Vidyapeetham, India. tripty_singh@blr.amrita.edu, https://orcid.org/0000-0002-3688-4392

[4]Principal Investigator, Research Scientist, Bioinformatic Institute (BII), A*STAR Singapore. bhanu_prakash@bii.a-star.edu.sg, https://orcid.org/0000-0002-7555-7431

Received: February 17, 2024; Revised: April 04, 2024; Accepted: May 14, 2024; Published: June 29, 2024

Abstract

Purpose: Chronic Kidney Disease (CKD) is one of the world's top 20 causes of death. This novel study focuses on creating a prediction system for chronic kidney disease. It leverages T2 weighted MRI images and machine learning for efficient CKD classification, replacing labour-intensive manual processes. The adaptability of machine learning models accommodates changing disease patterns and diverse data sources. The purpose of this study is to investigate CKD, characterized by a sustained reduction in renal function lasting at least three months. CKD severity is gauged by kidney damage extent and glomerular filtration rate decline. The ultimate stage of CKD is end-stage renal disease.

Methods: The study focuses on various feature extraction from MRI data using (kNN), Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), and Gray level co-variance matrix (GLCM) along with a few morphological operations. Three Different sets of features are extracted, and Machine Learning Classification Models (Logistic Regression (LR), Support Vector Classifier (SVM), Decision Tree (DT), Random Forest (RF), k-Nearest-Neighbors (kNN), Naïve Bayes (NB)) are trained and tested on these set of features.

Results: Experiment results show that LR Classifier gives the highest Accuracy of 92% for GLCM features. SVM and RF Classifier provide the highest Accuracy of 91.5% for DCT features, and RF Classifier gives the highest Accuracy of 86.6%. Based on predictions made by each model, a soft voting classifier is trained and tested to achieve the best Classification for each set of features. This study helps analyse the influence of the voting classifiers obtaining an accuracy of 90% for GLCM features, followed by 89% for DCT features and 84% for DWT features.

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 15, number: 2 (June), pp. 202-225. DOI: *10.58346/JOWUA.2024.I2.014*

*Corresponding author: PhD Research Scholar, Department of Computer Science and Engineering, Amrita School of Computing, Bengaluru, Amrita Vishwa Vidyapeetham, India.

Conclusion: The suggested system offers a comparative analysis of classification models and techniques for feature extraction. The results of performance evaluation illustrate the efficacy of different feature extraction and classification approaches. This will contribute to the timely detection, routine examination, and efficient control of CKD.

Keywords: Discrete Wavelet Transform, Discrete Cosine Transform, GLCM, CKD, Random Forest, Machine Learning.

1 Introduction

CKD is a persistent decline in renal function for at least three months. Its severity is determined by the amount of kidney damage and the decrease in glomerular filtration rate. The end-stage renal disease represents the most critical stage of CKD. The National Kidney Foundation (NKF) uses the following definition of CKD: "if the kidney damage, such as proteinuria, haematuria, or pathological abnormality, or glomerular filtration rate is less than 60 mL/min/1.73 m2 and lasts for more than three months".

The kidneys aim to eliminate excess fluid and waste from the bloodstream to create urine. If an individual has Diabetes, high blood pressure, and heart disease are all potential causes of CKD as discussed in (Alnazer et al., 2021). CKD can be treated if detected early. Renal replacement therapy, such as transplantation or dialysis, is eventually required as in (Kim & Ye, 2021). CKD is diagnosed through physical examination. Blood Test measures the level of creatinine, which increases when kidneys are not functioning correctly. A higher level of creatinine is an indicator of CKD (Mishr et al., 2020). Urine Tests help to determine if the patient has increased levels of albumin protein, which indicates kidney damage. Glomerular filtration rate (GFR) test measures how well the kidneys filter waste from the blood. A low GFR suggests the presence of CKD is mentioned in (Vinayagam et al., 2022; Satukumati et al., 2019; Thamara et al., 2021).

Classification of medical images is important in understanding and categorizing abnormalities from various imaging modalities. MRI is one of the most high-resolution imaging modalities available (Faisal et al., 2018; Cenggoro et al., 2023; Abdullah., 2020; Kumar et al., 2023). T1 and T2 weighted MRI are the most common MRI sequences. T2 Weighted MRI is used in this study. T2 weighted image (T2WI) is a basic MRI pulse sequence. Longer TE (Time to Echo) and TR (Repetition time) times generate T2-weighted images (Hemasree et al., 2022). T2 images have a high contrast because, in imaging, tissues with a low T2 value will appear dark as they have lost most of their signal strength, whereas tissues with a high T2 value will appear bright as they have a stronger signal. For various reasons, machine learning outperforms the best approach for feature extraction and Classification of Kidneys as chronic and not chronic (Kutlu et al., 2021; Uyan, 2022). Machine learning techniques can automate the process of feature extraction and Classification, which takes a lot of time and effort to extract features and perform Classification manually (Bakti et al., 2021; Prakash et al., 2023; Srinivasa et al., 2023). Machine learning algorithms or models in the context of paper are adaptable because they can be re-trained on new data (data generated each time after applying different feature extraction methods), allowing them to adapt to changes in the underlying disease process (Juma et al., 2023; Salman et al., 2023; Park et al., 2019). This paper proposes a novel feature extraction method by converting the MRI images to grayscale images using various Grayscale as well as morphological operations along with the discrete mathematical wavelet transform and discrete cosine transform operations(Jonnerby et al., 2023; Nowakowski et al., 2021; Trivedi et al., 2023). Different Classification algorithms are given input with various feature extraction, and the accuracy and evaluation matrices of various classification algorithms are compared with the extracted features. Based on predictions made by each of the Classification models, a voting classifier is trained and tested for each of the feature sets to analyse the best

Classification this study can perform with the extracted features for Chronic and Healthy Kidney MRI images (Babu et al., 2021; Mosa et al., 2022; Karankar et al., 2017; Mathew et al., 2022).

SIFT (Scale Invariant Feature Transform) provides an exceptional example of features, such as edges, masses, and corners, of a picture at a specific intrigue point. It offers a collection of elements with no issues that some other methods have as discussed in (Pattanayak et al., 2022; Chen et al., 2019). The features utilized are Contrast, Angular second moment, Entropy, and Correlation. The KNN classifier classifies the images based on training data, resulting in an overall accuracy of around 91% as applied in (Kumar & Pati, 2023; Afza et al., 2021). Identifying patterns in kidney ultrasound images involves utilizing five-intensity histogram features and nineteen grey-level co-occurrence matrix features. This combination allows for the categorization of the images into four distinct groups: regular, bacterial infection, cystic disease, and kidney stones as in (Gayathri et al., 2022). The analysis of statistics and texture features, including using scale-invariant feature transform (SIFT), has revealed that features like Energy, Variance, and Kurtosis are elevated in typical kidney images compared to those with renal abnormalities. This distinction in feature values can be utilized to differentiate between normal and abnormal kidney conditions as presented in (Pati et al., 2022; Singh et al., 2019).

The process of creating a classification model for kidney ultrasound images involves preprocessing, feature extraction, and feature selection. Preprocessing steps such as cropping, interpolation, rotation, and background removal are performed to enhance image quality and diagnostic Accuracy. Texture features are extracted using grey level co-occurrence matrix (GLCM) and include statistics such as Energy, entropy, homogeneity, correlation, contrast, and dissimilarity. The obtained features are then reduced to a more manageable subset through principal component analysis (PCA) as in (Babu et al., 2022). This research introduces a modified multi-resolution DWT (MMDWT) an algorithm built on the Laplacian pyramid and 1D discrete wavelet transform (DWT) provides n-level decomposition with reduced computational complexity and compatibility with all types of mother wavelets as in (Vinayagam et al., 2022; Nair et al., 2019, Cantoni et al., 2020).

The remainder of the paper is as follows. Section II provides the data description and the image preprocessing. It elaborates on the different extraction methods used to collect multiple features along with the Machine Learning algorithms used in this study to perform Classification. Section III focuses on the experiments and results. Section IV provides brief discussion followed by conclusion.

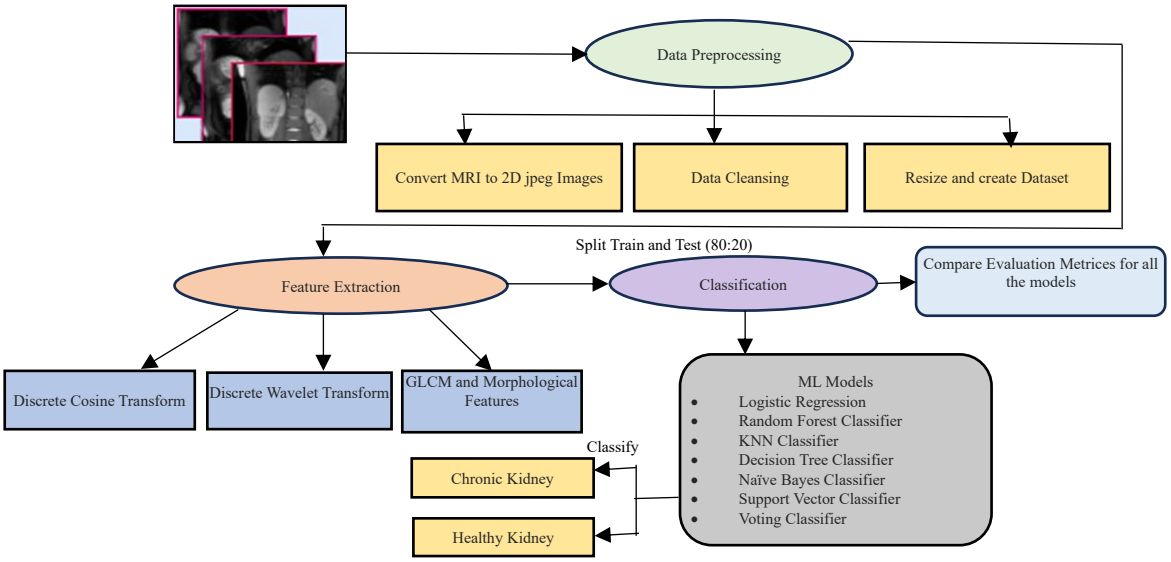


Figure 1: Block Diagram of the Proposed Methodology

2 Methods

Data Preprocessing and Description

In this paper, we have used Kidney T2 Weighted MRI Dataset patented in (Daniel et al., 2021)., which is a publicly available Dataset indexed in Open AIRE. The Dataset contains a collection of 100 T2-weighted abdominal MRI scans. The scans were obtained from 50 healthy control individuals and 50 patients suffering from CKD. The MRI scans are accompanied by kidney masks manually marked by experts. This MRI sequence aims to increase the Accuracy of segmentation by optimizing the contrast between the kidneys and the tissue surrounding them. Half of the scans were taken from individuals with no health issues, while the other half were from patients with CKD.

The acquired Dataset contained files in 3D.nii (Neuroimaging Informatics Technology Initiative) format. Each file was converted to jpeg format, resulting in about 8 to 14 jpeg slices per 3D MRI file. Jpeg images from 50 Healthy Kidney MRI and 45 Chronic Kidney MRI data were taken for classification purposes, respectively. The converted jpeg images for healthy Kidneys were 650, and for Chronic Kidneys were 585. Entirely, 1235 images were generated. All blank images generated during 3D to 2D file conversion were removed. Total of 1010 images were taken for further processing. All the images were converted to Grayscale before applying image processing operations and extracting features.

The Kidney MRI data used in this paper is used for performing kidney segmentations in one of the open-access journals. The following Dataset has yet to be used to train and test machine learning models or perform feature extraction techniques. MRI is one of the high contrasts and best imaging modalities, which helps perform good quality analysis. This is the only publicly available MRI CKD Dataset. MRI provides a better understanding of the textural patterns and features than other Modalities, such as ultrasound and Computed Tomography. The novelty lies in presenting various feature extraction and machine learning classification techniques to extract high-quality MRI features to perform better Classification.

Feature Extraction Techniques

The feature extraction consists of three variations in the paper, firstly using the Discrete Wavelet Transform operation. Secondly, use morphological and GLCM operations, and finally, use Discrete Cosine Transform operations.

Discrete Wavelet Transform

The discrete wavelet transform (DWT) is a method of wavelet transformation where wavelets are sampled at definite intervals. DWT simultaneously provides both spatial and frequency domain information for images. The Wavelet Transform employs a set of functions known as wavelets, each with a different scale.

The wavelet was applied to assess various frequencies present in an image at different scales. DWT is being utilized as a potent tool for feature extraction. DWT was used to obtain wavelet coefficients from magnetic resonance (MR) images of the kidneys.

$$\text{DWT } C(\text{s}) = \begin{cases} dk, l = \sum c(\text{s}) h * k(\text{s} - 2kl) \\ dk, l = \sum c(\text{s}) g * k(\text{s} - 2kl) \end{cases} \quad (1)$$

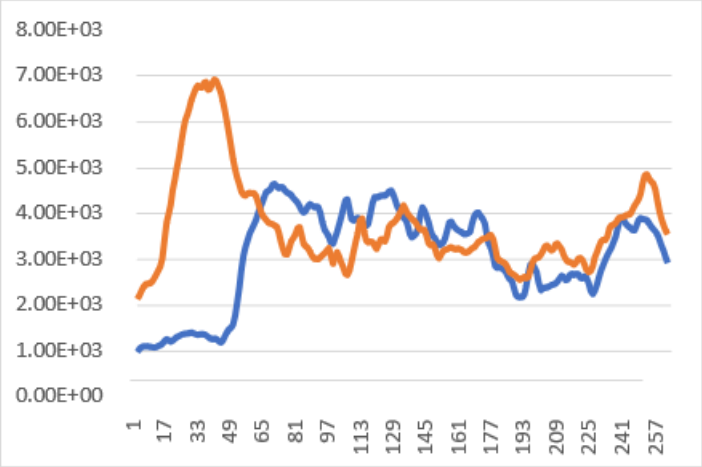


Figure 2: The Line Plot of 256 DWT Features for the Chronic Kidney MRI Sample

Equation 1 presents how DWT is calculated. The coefficients dk, l refers to the wavelet function's component attribute in signal c(s). In the mathematical equation, the coefficients of the high-pass and low-pass filters are represented by h(s) and g(s), respectively. The parameters I and j denote the wavelet scale and translation factors.

DWT is beneficial for extracting features from MRI images due to its ability to capture low-frequency and high-frequency components. It can also easily be used to analyse multi-dimensional signals such as the ones represented by MRI images. DWT can be applied to MRI images to extract a set of wavelet coefficients used as features for subsequent analysis of Kidney Image Classification tasks. DWT decomposes each image into a set of detailed coefficients at different scales. The number of coefficients extracted equals the number of pixels in the images at that scale. The total number of coefficients generated from a Gray level 256x256 image at a given level of decomposition is 256x256.

Features extracted are the individual measurable classify characteristics of Kidney MRI images that are used to them into different classes. The sum and average of features are statistical measures that can be used to summarize the distribution and behaviour of the features across a Feature set. The sum of features is the total value of a feature across all instances in a dataset; they indicate the overall magnitude or scale of features across a Feature set. The average of features, also known as the Mean, is the sum of a feature divided by the number of instances in a dataset. The average of features provides a measure of the central tendency value of a feature across a dataset. Fig. 2 plots all 256 features for one Sample Healthy Kidney (in orange) and one Sample Chronic Kidney (in blue) MRI Images.

All the coefficient values extracted from the sample Healthy Kidney Image range between 2115 and 6904. The sum and Average Mean of all the Healthy Kidney Sample Image features are 953000 and 3720, respectively. In Fig. 2, the blue line plot presents all 256 features for one Chronic Kidney MRI Image. All the coefficient values extracted from the Images range between 988 and 4630. The sum and Average Mean of all the Chronic Kidney Sample Image features are 788000 and 3080, respectively. The lowest and highest feature values, Sum, and Average mean feature values for the sample Healthy Kidney MRI Image are more significant than the Chronic Kidney MRI Image. For example, mean for the healthy image feature is 3720, and for Chronic, 3080. The image's spatial and frequency domain information, which means the coefficients generated during feature extraction. These coefficients are lesser in Chronic Kidney MRI Images when compared to Healthy images because Chronic Kidneys do not have a fuller appearance like healthy kidneys. Imaging findings show reduced kidney size in patients

with CKD, particularly at advanced stages, which has led to smaller coefficient values generation. DWT cannot accurately capture the differences in the high-frequency components of the abnormalities in Kidney MRI Images, generating smaller values for Chronic images when compared to DCT.

Discrete Cosine Transform

The Discrete Cosine Transform (DCT) represents an image as a combination of sinusoidal functions with varying intensity and frequency. The DCT function computes the 2D transformation of the image. DCT is applied to both the class of images and 1024 features are extracted for each image to provide a machine learning model as input to perform Classification. Equation 2 describes how DCT extracts feature from 2D MRI images.

$$B_{mn} = \alpha_m \alpha_n \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} A_{kl} \cos \frac{\pi(2k+1)m}{2K} \cos \frac{\pi(2l+1)n}{2L}, \quad 0 \le m \le K-1, \quad 0 \le n \le L-1$$
$$\alpha_m = \begin{cases} 1/\sqrt{K}, & m = 0 \\ \sqrt{2/K}, & 1 \le m \le K-1 \end{cases} \quad \alpha_n = \begin{cases} 1/\sqrt{L}, & n = 0 \\ \sqrt{2/\mathrm{L}}, & 1 \le n \le L-1 \end{cases} \tag{2}$$

DCT is a mathematical transformation used in image compression and feature extraction in Kidney MRI Images because it can reduce noise such as Gaussian noise, salt and pepper noise, and speckle noise from Kidney MRI images while preserving the most significant features. Since DCT has essential functions designed to be orthogonal and compact, it captures the fundamental characteristics while reducing redundancy. This transform is computationally efficient and can be applied to this large Dataset of Kidney MRI Images. Additionally, the DCT can detect subtle kidney changes that would otherwise go unnoticed. To transform each 256x256 gray image using DCT, the image is divided into 8x8 blocks. Each block is then altered using the DCT formula, which produces a set of 64 DCT coefficients. These coefficients represent the frequency content of the block in terms of low-frequency and high-frequency components. For each block, the DCT coefficient represents the average value of the block, and the other coefficients represent the variations from this average value. Since there are 64 coefficients for each 8x8 block, and the original image has 256x256 pixels, the total number of DCT coefficients produced is 256x256 / 8x8 x 64 = 1024.

Fig. 3 plots all 1024 features for one Chronic Kidney MRI Image. All the coefficient values extracted from the Images range between -1818 and 11302. For this sample image, the Sum and Average Mean of all the features are 1490000 and 1460, respectively. Fig. 4 plots all 1024 features for one Healthy Kidney MRI Image. All the coefficient values extracted from the Images range between -4564 and 54232. For this sample image, the Sum and Average Mean of all the features are 1440000 and 1410, respectively. We can see how the range of feature values extracted differs for healthy and chronic Kidney images, along with the Mean and the sum of the features. The plots in Figure. 3 and Figure. 4 present the line plots' differences and DCT is applied for Chronic and Healthy Kidney MRI Images. The DCT coefficients of CKD images tend to be higher than those of healthy kidney MRI images because of the presence of abnormal organs or tissues in CKD MRI images. For example, the mean for Chronic Images is 1460, and for Healthy Images, it is 1410. Thus, DCT coefficients can be used to accurately distinguish between healthy and CKD images by examining the differences in their respective pixel intensities. DCT is often used as a baseline or reference point for comparing and evaluating the feature values of individual instances or subsets of MRI image data.

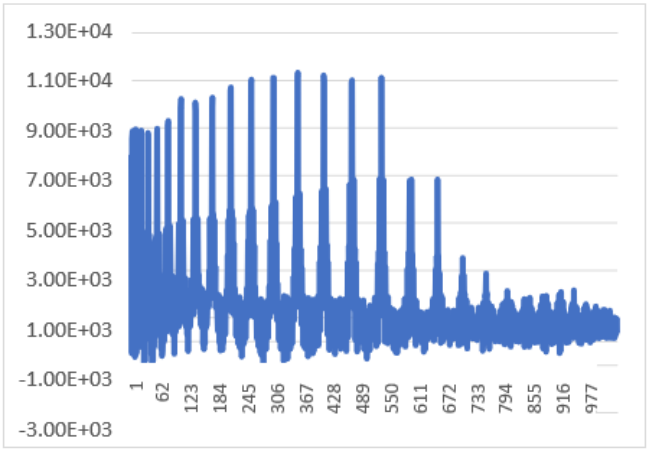


Figure 3: Line Plot of 1024 DCT Features for Chronic Kidney MRI Sample

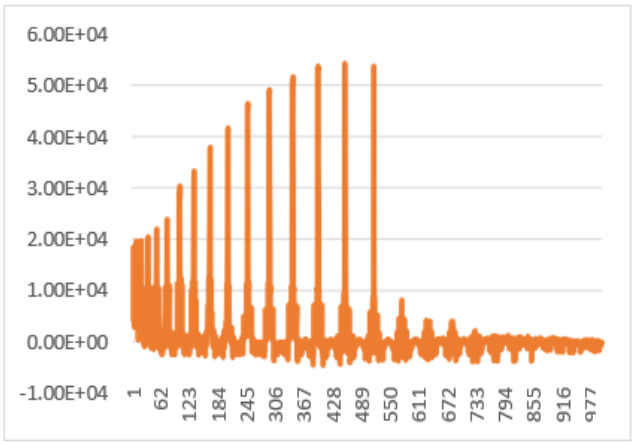


Figure 4: Line Plot of 1024 DCT Features for Healthy Kidney MRI Sample

GLCM and other Image Processing Operations

Firstly, the all-MRI images are converted to Grayscale, and various operations are applied to extract the Mean and standard deviation of each type of image, removing 20 crucial features. The different imaging operations include Histogram equalization, bilateral filters, clahe operation, and Otsu binarization: zero inversion, Sobel filter, erosion, and morphological dilation operations. GLCM features include Energy, correlation, dissimilarity, homogeneity, contrast, and entropy. Once all of these features are extracted, they are stored in a data frame and extracted into an Excel sheet containing 20 features. Table 1 presents Population mean values for all the features extracted from the combined Healthy and Chronic MRI Image dataset.

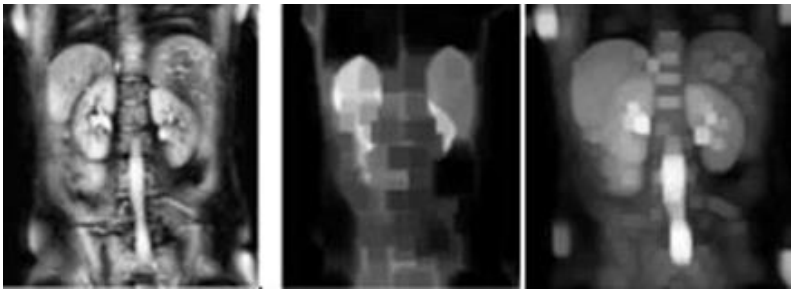
During preprocessing, all the 2D images generated are converted to grayscale images, and then various operations are applied onto them. GLCM features and other morphological operations. Features are generated from grayscale images. Histogram Equalization is used in Computer image processing to enhance the contrast in images by distributing the most frequent intensity values evenly. CLAHE (Contrast Limited Adaptive Histogram Equalization) is a variation of this technique that limits contrast amplification to prevent excessive noise amplification. In CLAHE, the amount of contrast amplification for each pixel is determined by the slope of the transformation function. Dilation is a technique used to

enlarge the pixels in an image by using a structuring element. Erosion, on the other hand, is used to remove pixels on the borders of an object. In other words, it shrinks the foreground objects.

Table 1: The Table Contains all the Features Extracted by Performing Various Operations, and their Calculated Population Mean Value Against Each Feature

Features	Mean value	Features	Mean value
Gray image Mean	55.4	Histogram image Mean	61.9
Gray Image STD_Dev	43.9	HistogramImageSTD_Dev	75.4
Adaptive image Mean	24.4	Dilation image Mean	79.1
AdaptiveImageSTD_Dev	46.7	Dilation Image STD_Dev	51.7
clahe image Mean	105.2	Energy	0.04
clahe Image STD_Dev	62.5	Entropy	176.2
Sobel image Mean	4.6	Correlation	0.9
Sobel Image STD_Dev	1110.4	Dissimilarity	7.5
Erosion image Mean	24.9	Homogeneity	0.2
Erosion Image STD_Dev	25.7	Contrast	176.2

a) Original b) Grayscale c) Histogram Equalization



d) Clahe Operation e) Erosion Operation f) Dilation Operation

Figure 5: Images After Performing and Applying Various Operations

In Fig. 5 sample images, Each Feature is extracted based on mathematical equations; the equations are presented below for a few features.

$$\textbf{Mean} \quad = \quad \mu_i = \sum_{k,l=0}^{N-1} k\left(P_{k,l}\right), \quad \mu_j = \sum_{k,l=0}^{N-1} l\left(P_{k,l}\right) \tag{3}$$

$$\textbf{Standard Deviation} = \sigma_k^2 = \sum_{k,l=0}^{N-1} P_{k,l}(k-\mu_k)^2, \sigma_l^2 = \sum_{k,l=0}^{N-1} P_{k,l}(l-\mu_l)^2 \tag{4}$$

$$\sigma_k = \sqrt{\sigma_k^2},\ \sigma_l = \sqrt{\sigma_l^2}$$

Mean and Standard deviation is calculated for all the images after applying various image processing operations, and the estimated mean and standard deviation for each type are stored in the extracted features data. Equations 3 and 4 are the commonly used equations.

Gray Image - (R + G + B / 3) (5)

RGB Image or color consists of 3 layers R, G, and B. It's a 3- dimensional matrix, and the grayscale image is of only two dimensions, and the values range between 0–255 (8-bit unsigned integers). All images used for processing are converted to Grayscale, and then various operations are applied in equation (5).

Histogram Equalization – $s_k = T(r_k) = (Q-1)\sum_{l=0}^{k} p_r(r_l) = \frac{(Q-1)}{N^2}\sum_{l=0}^{k} n_l$ (6)

In equation (6) shows Histogram Equalization enhances contrast in images by redistributing the frequency of intensity values by expanding the range of intensity values in the image resulting in a global contrast improvement. Adaptive Histogram Equalization is a variant of histogram equalization difference; instead of computing a single histogram for the entire image, it calculates multiple histograms, each for a distinct section. Contrast Limited Adaptive Histogram Equalization (CLAHE) is a modified version of adaptive histogram equalization that introduces contrast limiting with the additional steps of histogram clipping and re-equalization.

Sobel – Sobel is an edge detection filter in image processing. The Input MRI image convolves with the horizontal Sobel kernel to obtain a flat gradient image ***P***x, then convolve the image with the vertical Sobel kernel to obtain a vertical gradient image ***P***y. The gradient magnitude of each pixel in the image is then computed as the square root of the sum of the squares of the corresponding ***P***x and ***P***y values using in equation (7) and (8):

$$P = \sqrt{P_x^2 + P_y^2} \quad (7)$$

This gradient magnitude image P can be thresholder to detect edges in the image. The gradient direction of each pixel as the arctan of the ratio of the ***P***y and ***P***x values:

$$\Theta = arcta\,n\left(\frac{P_y}{P_x}\right) \quad (8)$$

Erosion and dilation are basic image-processing operations used to manipulate objects' shape and size within an image. These operations use a mathematical formula, usually represented as a matrix or kernel.

Erosion(P) = BP (9)

Dilation(P) = B $\oplus$ P (10)

In Equation (9) and (10), where P is the original MRI image, B is the kernel or structuring element used for erosion and $\ominus$ denotes the erosion operator. The erosion operation involves sliding the kernel over the MRI image and performing a logical AND operation between the kernel and the corresponding pixels. If all the pixels in the kernel are non-zero, the center pixel of the kernel is set to 1. Otherwise, it is set to 0. The resulting image after erosion is a smaller version of the original image where the boundaries of objects have been eroded. Dilation operation is similar to erosion difference being logical OR Operation, and the resulting image after dilation is a larger version of the original image where the boundaries of objects have been expanded.

GLCM Features

Energy - $\sum_{k,l=0}^{P_{k,l}(-ln\ P_{k,l})}$ (11)

In equation (11) shows Energy value describes the uniformity of the texture by measuring the sum of squared elements. Energy value ranges between 0 and 1.

$$\textit{Entropy} - \sum_{k,l=0}^{N-1} \frac{P_{k,l}}{1+(k-l)^2} \quad (12)$$

Entropy value is calculated depending upon the number of gray levels present in an image. It focuses on the degree of disorder in equation (12).

$$\textit{Correlation} - \sum_{k,l=0}^{N-1} P_{k,l} \left[\frac{(k-\mu_i)(l-\mu_j)}{\sqrt{(\sigma_k^2)(\sigma_l^2)}} \right] \sum_{k,l=0}^{N-1} P_{k,l} |k-l| \quad (13)$$

Correlation calculates how pixels are linearly dependent on each other or with the neighboring pixels within the matrix of image values in equation (13).

$$\textit{Dissimilarity} - \sum_{k,l=0}^{N-1} P_{k,l}^2 \quad (14)$$

$$\textit{Homogeneity} - \sum_{k,l=0}^{N-1} \frac{P_{k,l}}{1+(k-l)^2} \quad (15)$$

Homogeneity is inversely related to the contrast; it calculates the smoothness of the distribution of gray levels inside the image shows in equation (14) and (15).

$$\textit{Contrast} - \sum N-1 \; Pk,l(k-l)^2 \quad (16)$$

In equation (16) Contrast is calculated based on the presence of several edges and noises in the image. The higher the presence of these edges, the wrinkles in the image higher the contrast. Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Gray-Level Co-Occurrence Matrix (GLCM) (Tandel et al., 2021; Reddy et al., 2023) are popular methods used for feature extraction from medical images, including MRI images of the kidney. The various extracted feature vectors can experiment with classifiers to determine the most accurate classification method for each feature set. The DCT and DWT transform the Kidney MRI image into a set of Coefficients that represent the MRI image's frequency components. These frequency components can then extract various Kidney image features, such as texture and shape information. GLCM is used to quantify the spatial relationship between different Gray levels in a Kidney image. It provides information about the texture of a print, which can help differentiate between Healthy and Chronic Kidneys. Changes in the texture and shape of the kidney tissue identified with the help of features can be used to determine the presence of CKD.

A distplot, a short form for distribution plot, is a type of plot commonly used in data analysis and data visualization to display the distribution of a dataset. The density is calculated using kernel density estimation (KDE). KDE works by estimating the probability density function of the dataset at each point along the x-axis. It does this by creating a kernel function at each data and then summing up the contributions of all the kernels to estimate the density at each point. The density of the feature set refers to the degree of concentration of the data points in each Feature in a given range. For example, In Fig. 6, Gray Mean and Sobel Mean features have density values ranging from 0 to 5 and 0 to 20, respectively. A higher density indicates that the feature data points are concentrated in a smaller range. In comparison, a lower density suggests that the feature data points are spread out over a range. In Fig. 6, distplot is plotted for all 20 features containing GLCM and other components extracted by applying numerous image processing operations. This display describes the distribution of each of the features in the Feature set by presenting the spread and skewness. In the above plot, we can see different variations of plots like negative skewed, positive skewed and normally distributed, etc.; a narrow spread indicates that the values in the feature set are concentrated around a limited range or central tendency.

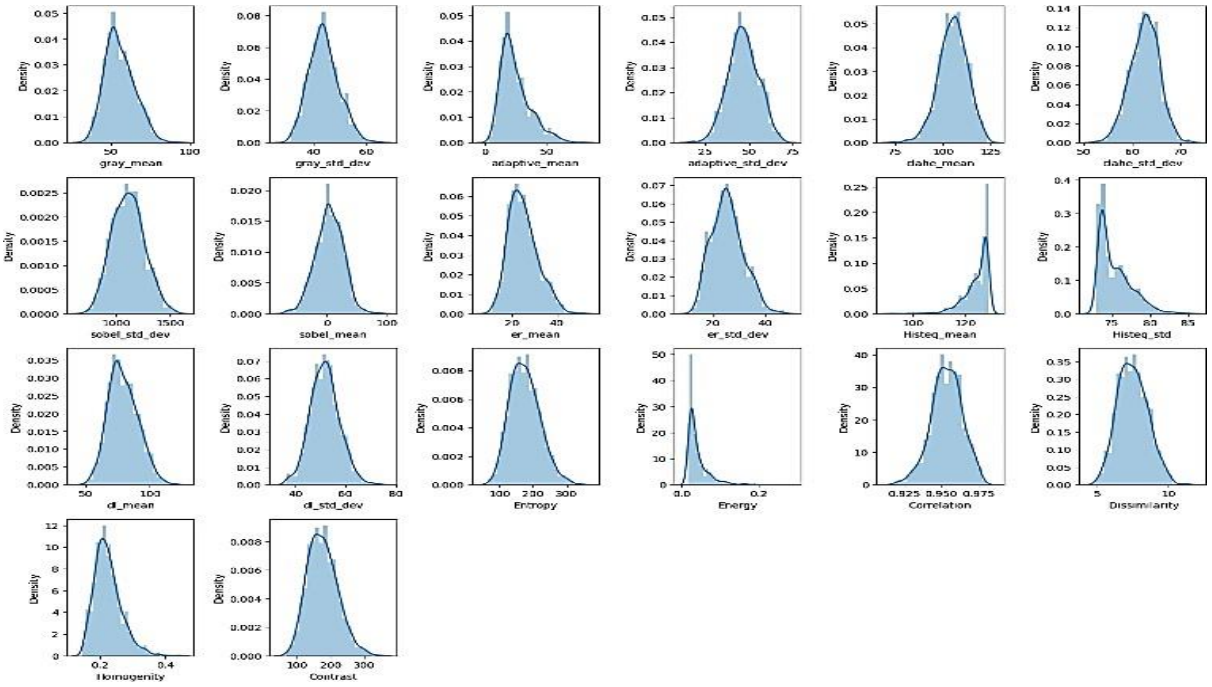


Figure 6: Plots of all 20 Features Extracted using GLCM and other Morphological Operations for a Complete Feature Set Comprising Image Features from Healthy and Chronic Kidney MRI Images

The distplot for Histogram equalization, mean, standard deviation and display for Energy GLCM features will have tall and narrow peaks, indicating that most of the observations fall within a small range of values. A wide spread means that the values in the Feature set are more spread out and less concentrated around a central tendency. The distplot for Sobel Standard Deviation (SD) and Contrast GLCM feature has flatter peaks and broader tails, indicating that the observations cover a more extensive range of values. The distplots for almost all the features extracted by applying SD have symmetric distribution has a similar shape on both sides of the central tendency, like Gray SD and Adaptive SD. The Mean, median, and mode are all equal, and the dirt plots for these features have a bell-shaped curve. Histogram Equalization Mean feature and Correlation feature have a positively skewed distribution with a longer tail on the right side of the central tendency, indicating higher values than low values. The Mean is greater than the median, and the distplot will have a longer tail on the right side and a shorter tail on the left side. Energy, Adaptive Mean, and Homogeneity features provide a negatively skewed distribution with a longer tail on the left side of the central tendency, indicating more low values than high values. The Mean is less than the median, and the distplot will have a longer tail on the left side and a shorter tail on the right side.

Classification Models Used

In the proposed method, various traditional machine learning methods are compared to identify the best-performing classifier for each of the ways of feature extraction. Here, five different models are used, they are as follows: Logistic Regression uses a sigmoid function in order to map the predicted probability values for binary Classification. In this paper, we have binomial image data, which can either be a Healthy Kidney or a Chronic Kidney. A few default parameters used in this work include a number of iterations, i.e., 100, Bias, or intercept in the function set to true (Alvee et al., 2021). A Decision Tree is used for classifying or predicting data based on a designated parameter. The model is trained and tested on data sets that include the desired categorization as Healthy and Chronic, making it a supervised

learning method (Gayathri et al., 2022). Random Forest Classifier uses multiple decision trees. Random Forest can be said as best suited for large feature datasets. A few default parameters of RF include 'Gini' to measure the quality of the split, no of Decision trees taken is 100, count of DCT, DWT, and GLCM features to consider for split 'sqrt.' The k-Nearest Neighbors algorithm employs supervised machine learning to label an unknown data point using previously labelled data. Distance algorithms, such as the Euclidean distance formula, are used to calculate how close two common features of healthy and chronic classes are from one another. A few default parameters are the number of neighbors given in numbers; weight is given uniformly by default, meaning each neighborhood data point is given equal weight. Naive Bayes, which is based on Bayes' Theorem. It uses the language of conditional probability. It assumes each category of features has its distribution. It calculates the probability of a data point in each category and gives the highest probability to a particular set of features. Support Vector Classifier SVC maps feature data points to a high-dimensional space and then locates the best hyperplane; SVC divides the data into Chronic Class and Healthy Class. (Alnazer et al., 2021) Parameters used for SVM are kernel as linear kernel and degree as 3, which means the polynomial degree of the kernel function. Soft Voting Classifier is one of the best methods for Accurate Classification. This approach to image classification uses an ensemble of all the classifiers. Each classifier is trained on the same Dataset and then voted on to decide the final output. The Soft Voting Classifier is especially effective for medical imaging tasks such as kidney MRI classification because it can combine the strengths of each classifier while mitigating weaknesses.

3 Results

Experiments of the proposed methodology have been categorized into three parts: Image preprocessing, feature extraction, and Classification. Jupyter Notebook is used for experimenting with the particular model with Python Version 3.11.1 on a Windows 11 Operating System with AMD Ryzen Processor with 16GB RAM. In Fig. 7, the distribution of classes before performing Classification is presented. Various machine learning models have been applied to three different feature sets, the first being the GLCM and morphologically combined 20 features extracted for both healthy and chronic kidney classes. The second is the Discrete cosine transform; 1024 features create a DCT feature set for Kidney MRI images. The DCT feature set contains all the compressed co-efficient generated after converting the MRI images to grayscale images and resizing them. DCT gives 1024 features because the DCT array gives average sample values forming vectors ranging from -1024 and +1023 for each of the input images. All the features in the feature sets are provided as a data frame input to perform machine learning classification. A similar procedure or feature extraction technique called Discrete Wavelet Transform is applied to MRI images using similar steps, except that the data extracted contains 256 features which extract numerical features from Raw image data. DWT performs feature extraction at the decomposition level for grayscale images. The intensity of the grayscale image is stored as an 8- bit integer, offering 256 possible variations of Gray from black to white. Therefore, DWT extracts 256 features from each of the 2D images. All three features are sent to different lists. We concatenate two classes list to form Feature sets, and then they are split with training and testing ratios of 80 and 20 percent.

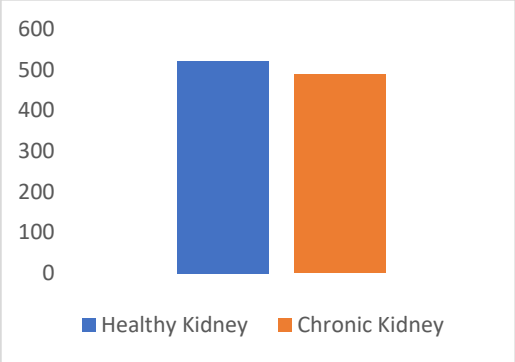


Figure 7: The above Bar Plot Describes the Class Distribution with Healthy Kidney and Chronic Kidney Containing 522 Healthy Kidney Images and 488 Chronic Kidney Images

The various parameters used for each classifier in this work are: Logistic Regression (solver='lbfgs', max_iter=1000), Random Forest Classifier (n_estimators =100), KNN Classifier (n_neighbors=9, metric= 'Minkowski,' p = 4), Decision Tree Classifier (criterian:'gini', splitter='best,' min_samples='2'), Naïve Bayes Classifier (GaussianNB, var_smoothing=1e-09) and Support Vector Machine Classifier (C=1.0, kernel=linear).

Predicted Labels	**Ground Truth (Actual Labels)**	
	Chronic Kidney	**Healthy Kidney**
Chronic Kidney	True Chronic (TC)	False Chronic (FC)
Healthy Kidney	False Healthy (FH)	True Healthy (TH)

Figure 8: Confusion Matrix for our CKD Classification

$$\text{Precision} = \frac{TC}{TC + FC} \quad (17)$$

$$\text{Recall} = \frac{TC}{TC + FH} \quad (18)$$

$$\text{F1 Score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (19)$$

$$\text{Accuracy} = \frac{TC + TH}{TC + FH + TH + FC} \quad (20)$$

Equation (17)-(20) shows all the tables containing performance evaluation metrics values are the Average Mean values of both Healthy and Chronic class performances for each feature set.

$$\text{True Chronic Rate} = \frac{TC}{TC + FH} \quad (21)$$

$$\text{False Chronic Rate} = \frac{FC}{FC + TH} \quad (22)$$

The ROC curve displays the relationship between TCR (True Chronic Rate) and FCR (False Chronic Rate) at various classification thresholds. Equations 21 and 22 present the calculation for TCR and FCR. By lowering the point, more items will be classified as positive, which will increase both False Positives and True Positives. The (Area Under the Curve-AUC) is determined by calculating the area under the ROC curve. ROC curves are plotted in this experiment, with the lowest value being 0 and the highest value being 1.

The confusion matrix for all the classification models for each feature set (GLCM, DCT, DWT) is presented in Table 8. Classifiers presented are as follows: LR (Logistic Regression), SVM (Support vector Machines), kNN (k nearest neighbors), DT (Decision Tree Classifier), RF (Random Forest), and NB (Naïve Bayes).

All metrics are presented to compare how various metrics perform for various machine learning models. This paper compares multiple classification models on 3 different feature sets. Every model works differently for each of the feature sets depending upon the training data size and quality, distribution of data, and complexity of models.

Chronic Kidney Disease Prediction using 6 Different Classification Models for GLCM, DWT, and DCT Feature Sets

First, we present how machine learning classification models work for GLCM and morphological sets of features.

Table 2: The Table Contains Performance Evaluation Metrics for 6 Models Arranged in the Decreasing Order of Accuracy, the First Accuracy being the Highest for GLCM Features Set

Model	Accuracy	Precision	Recall	F1 score
Logistic Regression	92%	92%	92%	92%
Support Vector Classifier	89.6%	90%	90%	90%
KNNeighbor Classifier	88.6%	89%	89%	89%
Decision Tree Classifier	88.1%	88%	88%	88%
Random Forest Classifier	86.6%	89%	89%	89%
Naïve Bayes Classifier	77.7%	79%	78%	78%

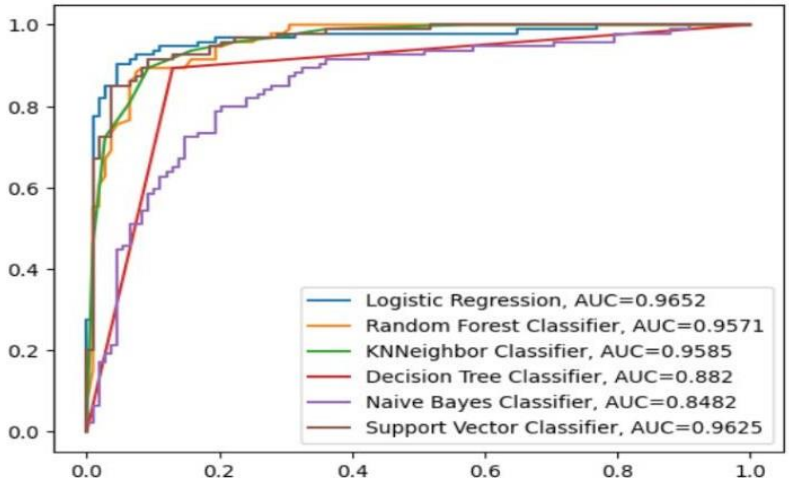


Figure 9: AUC (Area Under the Curve) Plot Comparison for all the Classification Models used is Described for the GLCM Feature Set

Table 3: Table Presents the Distribution of Data for all the Features GLCM, DCT, and DWT Features

GLCM Features	Images	Features (GLCM)	Features (DCT)	Features (DWT)	Class Labels
Original Set	1010	20	1024	256	1
Training Set	808	20	1024	256	1
Test Set	202	20	1024	256	1

Table 3 presents all the Features set to split before applying classification models for Training and Testing. Logistic Regression performs better, providing the highest Metrices being Accuracy, Precision, Recall, and F1-Score, as presented in Table 2. In Table 6, for the GLCM feature set, Logistic Regression (LR) has a higher number of Truly Chronic and fewer numbers of falsely Classified as Chronic; fewer are Falsely classified as Healthy and a more significant number of Truly Classified as Healthy. As presented in Table 2, all the metrics, when calculated, come to around 92%, making it the best performing classifier for the data. Accuracy comes to 92%. SVC proves to be the second-best classifier for the following data giving similar values in the confusion matrix like logistic regression, except that it has fewer FC and lesser TC, as presented in Table 6. In Fig. 9, AUC curves are plotted for all the models giving the highest AUC score of 0.96 for Logistic Regression and Support Vector Machines. TCR and FCR are almost similar for LR and SVM models providing similar plotting of ROC curves.

Table 4: The Table Contains Performance Evaluation Metrics for 6 Models Arranged in the Decreasing Order of Accuracy, the First Accuracy being the Highest for the DCT Feature Set

Model	Accuracy	Precision	Recall	F1 Score
Random Forest Classifier	91.5%	90%	90%	90%
Support Vector Classifier	91.5%	92%	92%	92%
Logistic Regression	76.2%	76%	76%	76%
KNNeighbor Classifier	76.2%	80%	77%	76%
Decision Tree Classifier	73.2%	74%	74%	74%
Naïve Bayes Classifier	69.3%	71%	69%	69%

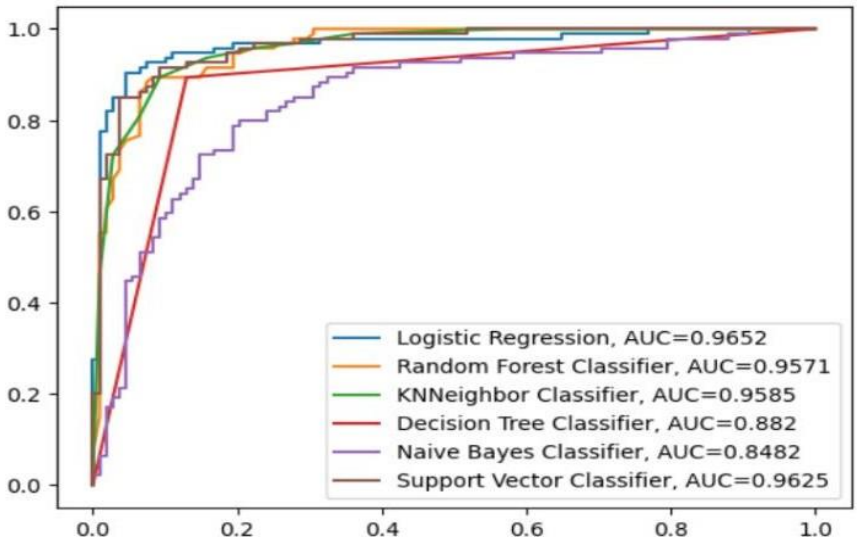


Figure 10: AUC (Area Under the Curve) Plot Comparison for all the Classification Models Used is Described for the DCT Feature Set

The confusion matrix for all classification models for the DCT feature set is presented in Table 6. Random Forest Classifier and Support Vector Machines provide a Higher number of TC and TH than all the other models, proving RF and SVM to be the best classifiers for the DCT feature set. RF and SVM perform better, providing the highest Metrics. In Table 4, SVM and RF give the highest accuracy score of 91.5%, followed by the LR Classifier with an Accuracy score of 76.2%. Fig. 10 is the plot of AUC curves for all the models for comparison. SVM holds 0.97, and RF holds 0.967, the highest AUC score, followed by kNN with an AUC of 0.889.

Table 5: The above Table Contains Performance Evaluation Metrics for Six Models Arranged in the Decreasing Order of Accuracy, the First Accuracy being the Highest for DWT Feature Set

Model	Accuracy	Precision	Recall	F1 Score
Random Forest Classifier	91.5%	90%	90%	90%
Support Vector Classifier	91.5%	92%	92%	92%
Logistic Regression	76.2%	76%	76%	76%
KNNeighbor Classifier	76.2%	80%	77%	76%
Decision Tree Classifier	73.2%	74%	74%	74%
Naïve Bayes Classifier	69.3%	71%	69%	69%

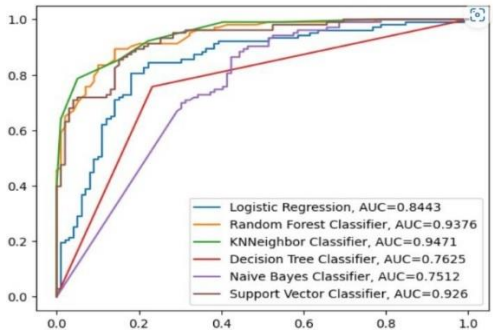


Figure 11: AUC (Area Under the Curve) Plot Comparison for all the Classification Models Used is Described for the DWT Feature Set

As Table 5 presents, Random Forest Classifier performs better, providing the highest Accuracy of 86.6% along with Precision, Recall, and F-Score, followed by the KNN Classifier with an accuracy of 85.1%. The confusion matrix for all the Classification models for DWT features is presented in Table 6. RF performs the best Classification compared to other classifiers, with more Truly Chronic and Truly Healthy. Fig. 11 shows the plotting of AUC curves for all the Classification models trained and tested for the DWT feature set. In Fig. 11, the AUC score of 0.947 for the kNN Classifier and the AUC score of 0.937 for the RF Classifier are presented.

Table 6: Summarizes the Performance of all Classification Models with the Predicted and Ground Truth Labels of the Confusion Matrix Under each Feature set for Chronic and Healthy Kidney MRI Image Classification Problem

Features	Classifiers	True Chronic (TC)	False Chronic(FC)	False Healthy(FH)	True Healthy(TH)
GLCM	LR	100	8	8	86
	SVM	94	14	7	87
	kNN	91	17	6	88
	DT	94	14	10	84
	RF	90	18	7	87
	NB	75	33	12	82
DCT	RF	93	6	11	92
	SVM	93	6	11	92
	LR	80	19	29	74
	kNN	94	5	43	60
	DT	67	32	22	81
	NB	54	45	17	86
DWT	RF	89	10	17	86
	kNN	84	15	15	88
	SVM	84	15	16	87
	LR	78	21	19	84
	DT	76	23	25	78
	NB	57	42	16	87

Chronic Kidney Disease Prediction using Voting Classifier

The Ensemble Vote Classifier is a meta-classifier for combining similar or conceptually different machine learning classifiers for Classification via majority or plurality voting. The Ensemble Vote Classifier implements "hard" and "soft" voting. In hard voting, we predict the final class label as the class label that has been predicted most frequently by the classification models. In soft voting, we predict the class labels by averaging the probabilities. In this experiment, we use a Voting Classifier provided by the sci-kit learn machine learning library.

As presented in Fig. 12, a soft voting classifier is used here to combine the predictions of all the machine learning models used in this study to make a final prediction. In soft voting, we predict the class labels based on the predicted probabilities Pred 1, Pred 2, …. Pred 6 for all classification models used. Our ensemble model makes predictions with the following:

$$\hat{y} = \arg\max_{c} \sum_{j=1}^{m} weight_j pred_{cj} \quad (23)$$

Where $weightj$ is the weight that can be assigned to the jth classifier, $pred$ are the predictions made by each classifier shows in equation (23) and (24).

$$\hat{y} = \arg\max_{c} [pred(c_0 \mid \mathbf{x}), pred(c_1 \mid \mathbf{x})] = 1 \quad (24)$$

The binary classification task with class labels c belongs to {0,1}, which is Chronic and Healthy. Using uniform weights, the average probabilities are computed. x belongs to C(x), which is a prediction made by classifiers C.

In soft voting, the final prediction is made based on the probability scores of each model's predictions. By combining the strengths of each classifier, Soft Voting Classifier can produce accurate results. Overall, Soft Voting Classifier is an excellent choice for accurately classifying kidney MRI images.

In Fig. 13, various performance metrics such as Accuracy, Precision, Recall, and F1 Score are plotted for each feature set. GLCM feature set gives an accuracy of 90% along with other metrics. All the predictions made for the GLCM feature set by Different classification models are given as input to the voting classifier to decide the best and most accurate result. Voting Classifier is trained and tested, and Fig. 13 shows results for the performance of the voting classifier for Test data. In Table 7, the Confusion matrix for each of the feature sets is presented.

Table 7: The Table Summarizes the Performance of the Soft Voting Classifier with the Predicted and Ground Truth Labels of the Confusion Matrix Under Each Feature set for Chronic and Healthy Kidney MRI Image Classification Problem

Classifier	Features	True Chronic (TC)	False Chronic (FC)	False Healthy (FH)	True Healthy (TH)
Soft Voting Classifier	GLCM	93	15	6	88
	DCT	88	11	12	91
	DWT	82	17	15	88

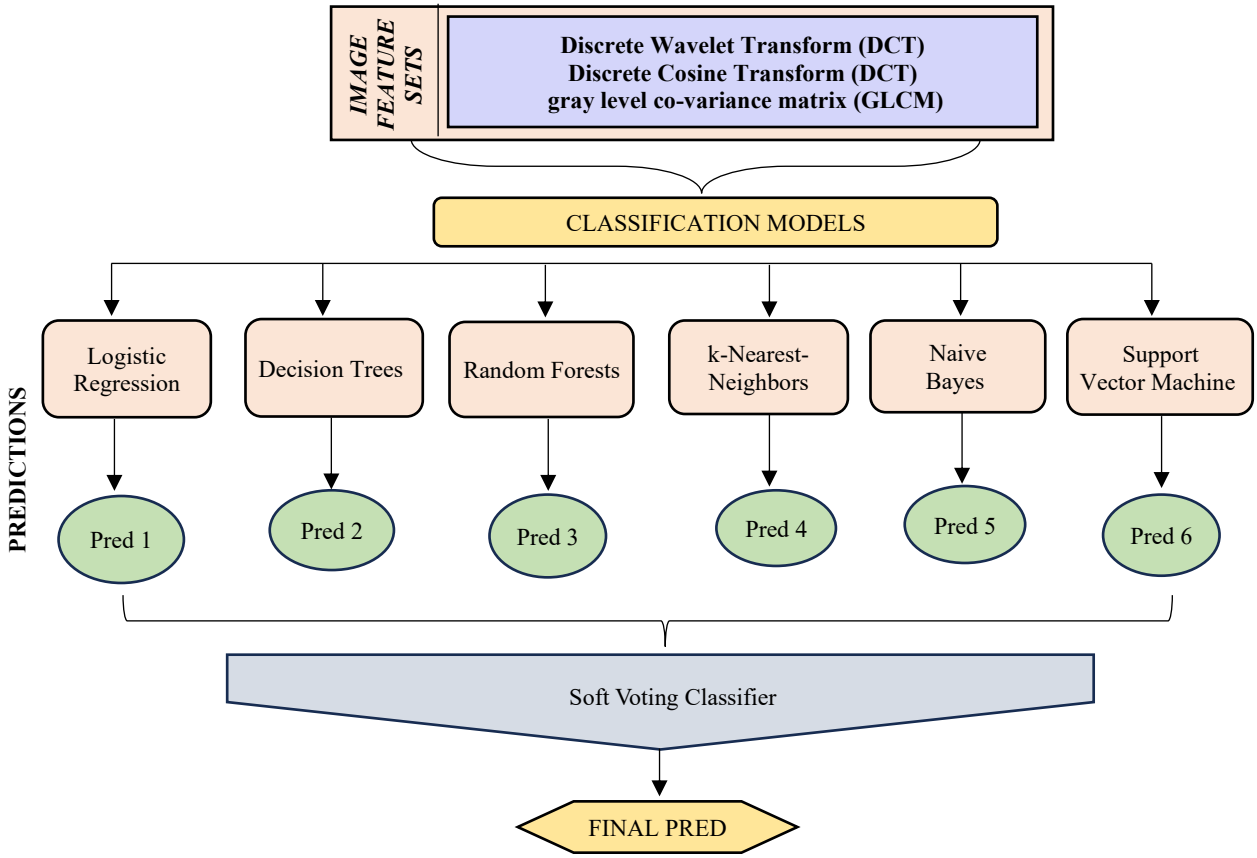


Figure 12: Voting Classifier for CKD Classification

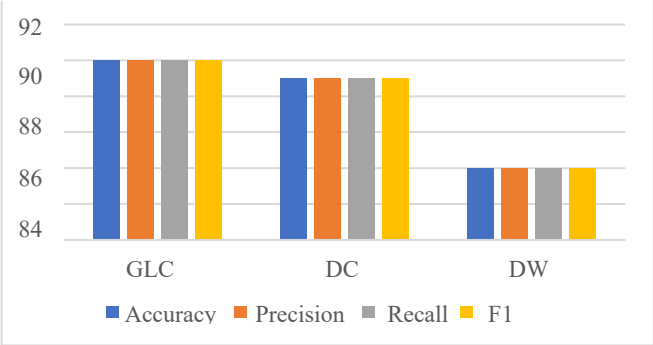


Figure 13: Accuracy, Precision, Recall, and F1 Score plot for Different Feature Sets (GLCM, DCT, DWT) for Soft Voting Classifier

4 Discussion

Overall, when comparing all the models for various Feature sets, we can say that Classification models such as Decision Trees and Naïve Bayes result in the least Accuracy for all the Feature sets. Naïve Bayes assumes no correlation between the features and considers all features are conditionally independent. Naïve Bayes may not generalize well for highly correlated data extracted from MRI images. There might be outliers on the level of individual features. MRI image data can have complex structures, such as non-linear relationships, that are difficult to capture with a Naive Bayes classifier. Due to the assumption of independence among features, the model cannot capture relationships between features. Decision Trees

have obtained less Accuracy because the tree structure may not be able to capture the linear relationships in the data. The Naïve Bayes Algorithm requires many training examples to estimate the parameters accurately. If the training set is small, being 202 images data, hence maybe the model is unable to generalize accurately. The question arises as to why none of the models can predict with 100 percent accuracy, the reasons being the features used to represent the Images in the Feature sets might not provide enough information for the classifiers to make higher accurate predictions and also machine learning algorithms are limited in their ability to capture complex relationships between the attributes of the Feature sets and the target or class.

This paper makes an important contribution to medical image classification by tackling the problem of categorizing 2D Kidney MRI images using various feature extraction approaches and machine learning models. For the GLCM feature set, the Logistic Regression Model gets an amazing 92% accuracy, while the Support Vector Classifier and Random Forest Classifier obtain 91.5% and 86.6% accuracy, respectively. Despite slight changes in accuracy, the aggregate predictions of the Voting Classifiers are regarded more dependable, emphasizing the study's emphasis on robustness.

Recognizing its limitations, the study primarily centres on classification accuracy metrics, with visual assessments conducted using confusion matrices and AUC plots. The dataset comprises 1010 2D MRI images, but future endeavours are aimed at refining performance metrics, enhancing dataset quality, and exploring publicly accessible datasets to validate medical image classification.

To pave the way for future research, the study suggests feature selection techniques to enhance diagnosis precision. By leveraging relevant image portions instead of exhaustive features, this approach aligns with the ongoing quest for streamlined and accurate medical image analysis. Overall, this study's fusion of feature extraction and machine learning advancements contributes to the field and propels future advancements in medical image classification.

5 Conclusions

The problem undertaken in this paper is to extract different features to classify 2D Kidney MRI images. Feature sets are created using feature extraction techniques such as GLCM, DCT, and DWT, and Machine Learning Models are applied to the different Feature sets. The logistic Regression Model gives the highest Accuracy of 92% for the GLCM Feature set, Support Vector Classifier gives the highest Accuracy of 91.5% for the DCT Feature set, and Random Forest Classifier gives the highest Accuracy of 86.6% for DWT set. DCT feature set contains the highest features of 1024. SVM Classifier gives the highest Accuracy, which depicts that SVM outperforms other models for large Datasets. Voting Classifiers provide an accuracy of 90% for GLCM, 89% for DCT feature sets, and an Accuracy of 84% for DWT Feature sets. Although Accuracy is less by a few numbers, we consider the voting classifier’s performance as best for each feature set because this Accuracy is not based on individual models but on a collection of predictions from different models. Hence, the predictions and classifications made by the voting classifier are considered the most accurate predictions as presented in (Tandel et al., 2021).

The experiments were conducted regarding classification accuracy score, precision, recall, score, and AUC Score. Also, we propose to use the visual evaluations for all the performance metrics along with the Confusion matrix and AUC Plots. This study is applied to 1010 2D MRI image Datasets. This study summarizes the effectiveness of various machine-learning techniques for MRI image classification tasks. Future work would be to improve Performance matrices and achieve higher Accuracy of machine learning models, often involving a combination of collecting more and higher-quality data by performing feature engineering and hyperparameter tuning. In the future, testing other publicly available image

datasets is suggested to determine the most appropriate models for medical image classification. As a future research direction, instead of using all extracted features, it is suggested to employ a feature selection method to select the most relevant parts of the medical Images for Diagnosis.

List of Abbreviations

CKD – chronic kidney disease
MRI – Magnetic Resonance Imaging
GLCM – Gray Level Co-variance Matrix
DCT – Discrete Cosine Transform
DWT – Discrete Wavelet Transform
AUC – Area Under Curve
RF – Random Forest
kNN – k- Nearest Neighbor
SVM – Support Vector Classifier
LR – Logistic Regression
DT – Decision Tree
NB – Naïve Bayes

6 Declarations

Availability of Data and Material

The datasets generated during the current study can be accessed at [Google Drive link - https://drive.google.com/drive/folders/1tTtwTRV_rQCuFkn0MJLks39uHjvn45r8?usp=sharing]. Access permissions have been set to public. For any inquiries related to the dataset, please contact Afnaan. K at afnaankhadar06@gmail.com.

Competing Interests

The authors declare that they have no conflicts of interest or Competing Interests regarding this study.

Funding

This project was conducted without external funding. The authors acknowledge that no financial support or grants were received for this research.

Authors' Contributions

AK preprocessed and generated the Kidney MRI data for further analysis. AK applied various machine learning predictions. PB Pati was a major contributor in performing the examination of results and writing the manuscript. T Singh and B P performed the analysis of the results acquired. All authors read and approved the final manuscript.

Acknowledgments

We want to express my gratitude to our university for providing the right environment to work on this paper. We want to acknowledge all those who have directly or indirectly contributed to completing this

work. We thank all the professors, colleagues, and friends from the Department of Computer Science and Engineering, Amrita School of Computing, for their immense support. We would like to express their sincere gratitude to doctors for their valuable contributions to this research project. Their expertise, guidance, and support have greatly enriched the quality of this study.

References

[1] Abdullah, D. (2020). A Linear Antenna Array for Wireless Communications. *National Journal of Antennas and Propagation (NJAP)*, *2*(1), 19-24

[2] Afza, F., Khan, M.A., Sharif, M., Kadry, S., Manogaran, G., Saba, T., & Damaševičius, R. (2021). A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection. *Image and Vision Computing*, *106*, 104090. https://doi.org/10.1016/j.imavis.2020.104090

[3] Alnazer, I., Bourdon, P., Urruty, T., Falou, O., Khalil, M., Shahin, A., & Fernandez-Maloigne, C. (2021). Recent advances in medical image processing for the evaluation of chronic kidney disease. *Medical Image Analysis*, *69*, 101960. https://doi.org/10.1016/j.media.2021.101960

[4] Alvee, B. I., Tisha, S. N., & Chakrabarty, A. (2021). Application of machine learning classifiers for predicting human activity. *In IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, 39-44.

[5] Babu, T., Singh, T., Gupta, D., & Hameed, S. (2021). Colon cancer prediction on histological images using deep learning features and Bayesian optimized SVM. *Journal of Intelligent & Fuzzy Systems*, *41*(5), 5275-5286.

[6] Babu, T., Singh, T., Gupta, D., & Hameed, S. (2022). Optimized cancer detection on various magnified histopathological colon imagesbased on dwt features and fcm clustering. *Turkish Journal of Electrical Engineering and Computer Sciences*, *30*(1), 1-17.

[7] Bakti, L. D., Imran, B., Wahyudi, E., Arwidiyarti, D., Suryadi, E., & Multazam, M. (2021). Data extraction of the gray level Co-occurrence matrix (GLCM) Feature on the fingerprints of parents and children in Lombok Island, Indonesia. *Data in Brief*, *36*.

[8] Cantoni, V., Cascone, L., Nappi, M., & Porta, M. (2020). Demographic classification through pupil analysis. *Image and Vision Computing*, *102*, 103980. https://doi.org/10.1016/j.imavis.2020.103980

[9] Cenggoro, T. W., & Pardamean, B. (2023). A systematic literature review of machine learning application in COVID-19 medical image classification. *Procedia computer science*, *216*, 749-756.

[10] Chen, Z., Edwards, A., Gao, Y., & Zhang, K. (2019). Learning discriminative subregions and pattern orders for facial gender classification. *Image and Vision Computing*, *89*, 144-157.

[11] Daniel, A. J., Buchanan, C. E., Allcock, T., Scerri, D., Cox, E. F., Prestwich, B. L., & Francis, S. T. (2021). T2-weighted Kidney MRI Segmentation (v1.0.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.5153568.

[12] Faisal, M. I., Bashir, S., Khan, Z. S., & Khan, F. H. (2018). An evaluation of machine learning classifiers and ensembles for early-stage prediction of lung cancer. *In IEEE 3rd international conference on emerging trends in engineering, sciences and technology (ICEEST)*, 1-4.

[13] Gayathri, R., Pati, P. B., Singh, T., & Nair, R. R. (2022). A framework for the prediction of diabtetes mellitus using hyper-parameter tuned xgboost classifier. *In IEEE 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1-5.

[14] Hemasree, V., & Kumar, K. S. (2022). Facial Skin Texture and Distributed Dynamic Kernel Support Vector Machine (DDKSVM) Classifier for Age Estimation in Facial Wrinkles. *Journal of Internet Services and Information Security, 12*(4), 84-101.

[15] Jonnerby, J., Brezger, A., & Wang, H. (2023). Machine learning based novel architecture implementation for image processing mechanism. *International Journal of Communication and Computer Technologies (IJCCTS), 11*(1), 1-9.

[16] Juma, J., Mdodo, R. M., & Gichoya, D. (2023). Multiplier Design using Machine Learning Alogorithms for Energy Efficiency. *Journal of VLSI Circuits and Systems, 5*(1), 28-34.

[17] Karankar, N., Shukla, P., & Agrawal, N. (2017). Comparative study of various machine learning classifiers on medical data. *In IEEE 7th International Conference on Communication Systems and Network Technologies (CSNT)*, 267-271.

[18] Kim, D. H., & Ye, S. Y. (2021). Classification of chronic kidney disease in sonography using the GLCM and artificial neural network. *Diagnostics*, *11*(5), 864. https://doi.org/10.3390/diagnostics11050864

[19] Kumar, A., & Pati, P. B. (2023). Offline HWR Accuracy Enhancement with Image Enhancement and Deep Learning Techniques. *Procedia Computer Science*, *218*, 35-44.

[20] Kumar, A., Joshi, P., Bala, A., Sudhakar Patil, P., Jang Bahadur Saini, D. K., & Joshi, K. (2023). Smart Transaction through an ATM Machine using Face Recognition. *Indian Journal of Information Sources and Services, 13*(2), 7–13.

[21] Kutlu, Y., & Camgözlü, Y. (2021). Detection of coronavirus disease (COVID-19) from X-ray images using deep convolutional neural networks. *Natural and Engineering Sciences, 6*(1), 60-74.

[22] Mathew, A., Antony, A., Mahadeshwar, Y., Khan, T., & Kulkarni, A. (2022). Plant disease detection using GLCM feature extractor and voting classification approach. *Materials Today: Proceedings*, *58*, 407-415.

[23] Mishr, R., Bhattacharjee, A., Gayathri, M., & Malathy, C. (2020). Kidney stone detection with CT images using neural network. *International Journal of Psychosocial Rehabilitation*, *24*(8), 2490-2497.

[24] Mosa, Q. O. (2022). Improved Kidney Stone Detection from Ultrasound Images Using GVF Active Contour. *Journal of Al-Qadisiyah for computer science and mathematics*, *14*(1), 24-31.

[25] Nair, R. R., & Singh, T. (2019). Multi-sensor medical image fusion using pyramid-based DWT: a multi-resolution approach. *IET Image Processing*, *13*(9), 1447-1459.

[26] Nowakowski, P., Zórawski, P., Cabaj, K., & Mazurczyk, W. (2021). Detecting Network Covert Channels using Machine Learning, Data Mining and Hierarchical Organisation of Frequent Sets. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 12*(1), 20-43.

[27] Park, M., You, G., Cho, S. J., Park, M., & Han, S. (2019). A Framework for Identifying Obfuscation Techniques applied to Android Apps using Machine Learning. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 10*(4), 22-30.

[28] Pati, P. B., Chandana, G. V., Reddy, C. R., Subash, G. B., & SriHarsha, J. (2022). Handwriting Quality Assessment using Structural Features and Support Vector Machines. *In IEEE 7th International conference for Convergence in Technology (I2CT)*, 1-5.

[29] Pattanayak, S., Pati, P. B., & Singh, T. (2022). Performance Analysis of Machine Learning Algorithms on Multi-Touch Attribution Model. *In IEEE 3rd International Conference for Emerging Technology (INCET)*, 1-7.

[30] Prakash, K., & Saradha, S. (2023). Efficient prediction and classification for cirrhosis disease using LBP, GLCM and SVM from MRI images. *Materials Today: Proceedings*, *81*, 383-388.

[31] Reddy, D. A., Roy, S., Kumar, S., & Tripathi, R. (2023). A Scheme for Effective Skin Disease Detection using Optimized Region Growing Segmentation and Autoencoder based Classification. *Procedia Computer Science*, *218*, 274-282.

[32] Salman, R., & Banu, A. A. (2023). DeepQ Residue Analysis of Computer Vision Dataset using Support Vector Machine. *Journal of Internet Services and Information Security, 13*(1), 78-84.

[33] Satukumati, S. B., & Shivaprasad Satla, R. K. (2019). Feature extraction techniques for chronic kidney disease identification. *Kidney*, *24*(1), 95-99.

[34] Singh, T., & Babu, T. (2019). Fractal image processing and analysis for compression of hyperspectral images. *In IEEE 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1-7.
[35] Srinivasa Rao, M., Praveen Kumar, S., & Srinivasa Rao, K. (2023). Classification of Medical Plants Based on Hybridization of Machine Learning Algorithms. *Indian Journal of Information Sources and Services*, *13*(2), 14–21.
[36] Tandel, G. S., Tiwari, A., & Kakde, O. G. (2021). Performance optimisation of deep learning models using majority voting algorithm for brain tumour classification. *Computers in Biology and Medicine*, *135*, 104564. https://doi.org/10.1016/j.compbiomed.2021.104564
[37] Thamara, A., Elsersy, M., Sherif, A., Hassan, H., Abdelsalam, O., & Almotairi, K. H. (2021). A novel classification of machine learning applications in healthcare. *In 3rd IEEE Middle East and North Africa COMMunications Conference (MENACOMM)*, 80-85.
[38] Trivedi, J., Devi, M. S., & Solanki, B. (2023). Step Towards Intelligent Transportation System with Vehicle Classification and Recognition Using Speeded-up Robust Features. *Archives for Technical Sciences, 1*(28), 39-56.
[39] Uyan, A. (2022). A Review on the Potential Usage of Lionfishes (Pterois spp.) in Biomedical and Bioinspired Applications. *Natural and Engineering Sciences, 7*(2), 214-227.
[40] Vinayagam, P., Sreemathi, M., Jeevitha, K., & Sandhya, S. (2019). Kidney stone detection using neural network. *International Journal of Applied Engineering Research*, *14*(6), 0973-4562.

Authors Biography

K. Afnaan is currently pursuing research within the field of medical image analysis at the Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Bengaluru. She earned her Master's degree in Computer Applications (MCA) from Jyoti Nivas College, Bengaluru, in 2020, following her Bachelor's degree in Computer Applications from Mount Carmel College in 2018.Her academic journey is marked by a deep-seated passion for the intricacies of medical image analysis, with a keen focus on areas such as classification, segmentation, optimization, and the translation of medical images. Afnaan's commitment to advancing the understanding and application of computer science in healthcare underscores her dedication to making meaningful contributions in the realm of medical imaging.

Dr. Peeta Basa Pati is a distinguished professor in the Department of Computer Science and Engineering at Amrita Vishwa Vidyapeetham, Bangalore. With a wealth of experience spanning over two decades, Dr. Pati holds a Ph.D. from the Indian Institute of Science, Bangalore (awarded in 2007), an M.Sc. (Engg.) from the same institute (2001), and a B.E. in Electrical Engineering from NIT (REC) Rourkela (1998). Currently serving as a Professor in the School of Computing, Bengaluru, Dr. Pati's qualifications include a BE, MSc, and Ph.D. His research interests encompass Camera Captured Document Analysis, Handwritten Document Analysis, Intelligent Document Processing, and Machine Learning Applications to Document Processing. Before joining Amrita, Dr. Pati worked as a Chief Architect at Cognizant Technology Solutions, contributing significantly to building and managing Intelligent Document Processing (IDP) systems. His responsibilities involved successfully implementing and productionizing IDP systems across various business domains, dealing with a variety of document types, including structured and unstructured, type-written, handwritten, and those with graphical content. Beyond his industry engagements, Dr. Pati has showcased his commitment to advancing knowledge through over 50 publications in Scopus-indexed conferences and journals. His profile reflects substantial contributions to both academia and industrial expertise in the domains of document digitization and information capture.

Dr. Tripty Singh is an Associate Professor in the Department of Computer Science at Amrita School of Engineering, Bengaluru, since 2012, accumulating over 18 years of expertise. She holds an Engineering degree in Electronics and Instrumentation (1999) from SGSITS Indore DAVV University, a Masters in Information Technology (2002-2004) from IIITM-Gwalior, and a Doctorate in Medical Image Processing (2008-2013) from MITS, RJTU (State Technical University), M.P. Her diverse research interests span Biomedical Image Processing, Artificial Intelligence, Computer and Information Sciences, Parallel and Distributed Systems, Ad-Hoc Networks, Network Security, and Digital Systems. Dr. Singh has earned international recognition, including the Best Idea Award in 3Bs, Young International Investigator Award, and accolades for the Best Paper Presentation and Best Contents in the track (ICEEE). With memberships in professional bodies like IEEE, MIAENG, and LMISOI. Dr. Singh's achievements extend to over 90 publications in Scopus-indexed journals and conferences, showcasing her significant contributions to the field. Her research impact is further exemplified by four Book Chapter Publications in Medical Imaging published by IGI Global. With a career spanning 17 years, she has supervised two Ph.D. scholars successfully and currently guides four research scholars in the realm of Medical Imaging, collaborating with renowned hospitals and research centres.

Dr.K.N. Bhanu Prakash is a Research Leader and Principal Investigator with extensive international experience in teaching, research, development, clinical product design, grant applications, and patent filing. He leads the Clinical Data Analytics & Radiomics group, focusing on developing advanced clinical research tools for accurate diagnoses, early interventions, and informed decision-making. Prakash has contributed to "Health & Medtech" solutions by automating preclinical and clinical research data analytics and creating computer-aided diagnoses for conditions like cancer, obesity, stroke, traumatic brain injury, and psychiatric disorders. He also mentor’s start-up companies, students, and staff, helping them realize their full potential and advance their skills. Prakash's key strengths lie in Biomedical Signal/Image Analysis, AI/ML/Data Analytics, Medical Imaging, Clinical Decision Support Systems, and Population health. He has published over 40 articles in Scopus-indexed journals and conferences, and is currently a Principal Investigator at the Bioinformatics Institute (BII) leading the way in advancing clinical data analytics and radiomics.