

Early Detection of Pancreatic Cancer Using Jaundiced Eye Images

R. Reena Roy* and G. S. Anandha Mala

Department of Information Technology, Easwari Engineering College, Chennai, Tamilnadu, 600089, India

*Corresponding Author: R. Reena Roy. Email: rreenaroyphd@gmail.com

Received: 06 January 2021; Accepted: 15 June 2021

Abstract: Pancreatic cancer is one of the deadliest cancers, with less than 9% survival rates. Pancreatic Ductal Adeno Carcinoma (PDAC) is common with the general public affecting most people older than 45. Early detection of PDAC is often challenging because cancer symptoms will progress only at later stages (advanced stage). One of the earlier symptoms of PDAC is Jaundice. Patients with diabetes, obesity, and alcohol consumption are also at higher risk of having pancreatic cancer. A decision support system is developed to detect pancreatic cancer at an earlier stage to address this challenge. Features such as Mean Hue, Mean Saturation, Mean Value, and Mean Standard Deviation are computed after color space conversion from RGB to HSV. Fuzzy k-Nearest Neighbor (F-kNN) is designed for classification. The system proposed is trained and tested using features extracted from jaundiced eye images. The proposed system results indicate that this model can predict pancreatic cancer as earlier as possible, helping clinicians make better decisions for surgical planning.

Keywords: Pancreatic adeno carcinoma; fuzzy-based classifier; jaundiced eye; classification system

1 Introduction

Among all forms of cancer, pancreatic cancer is one of the deadly diseases [1,2], and it remains unnoticed by many people since cancer-specific symptoms are known only at later stages [3,4]. Due to this, many patients remain under-diagnosed, which does not lead clinicians to surgical planning. One of the earliest symptoms of pancreatic cancer is Jaundice [5,6]. The eyes become yellow discoloration of the skin [5] and whites of the eye as in Fig. 1. The sclera is the white part of the eye that will get affected during Jaundice. As carcinoma blocks the tube, the common bile duct that releases bile into the intestine. It helps develop bilirubin within the blood [7].

The blockage causes the urine to be dark yellow, light-colored stools, and itching sensation. The skin color and whites of the eyes will vary dynamically, counting on bilirubin levels in the blood. The yellow color appears when bilirubin exists in moderate levels, while very high levels will appear brown. The patient's severity will be known when there is an elevation in the blood's bilirubin levels. By recognizing sclera using computer vision and machine learning algorithms, it will detect the level of bilirubin in jaundice. The eye image of the patient with jaundice is used to detect whether the pancreas is affected by



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

cancer or not. The symptoms of pancreatic cancer [8] are yellowish of skin and eye, pain or pain in the abdomen, diabetes, nausea, and vomiting, gall bladder or liver enlargement, and blood clots

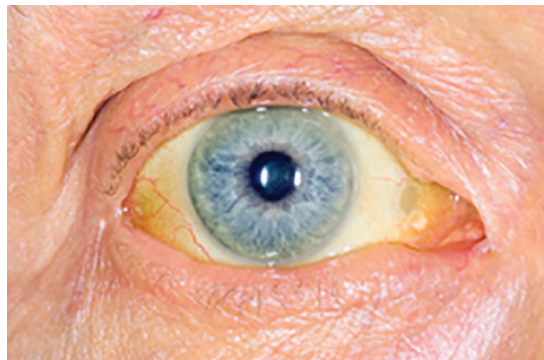


Figure 1: Jaundiced eye image

The research's main objective is to link jaundice and pancreatic cancer since jaundice is easily recognizable through the whites of the eye, i.e., sclera. Hence, this becomes the motive to focus on this research. The pancreas organ condition can be detected through the iris images of diabetes patients [9]. The beta cells of the pancreas organ using iris diagnosis can measure insulin deficiency [10].

The decision support system is built based on recognizing the sclera from eye images through machine learning algorithms and extracting features from the sclera through color descriptors. Hence Hue, Saturation, mean value, and standard deviation are calculated. The sclera is segmented through a modified min-max color space threshold algorithm. The extracted features are calculated and given as the input for the proposed F-kNN for classification and validation. The fuzzy-based initialization functions are modeled. The weights are created using the sigmoid function for improving the accuracy of classification in kNN. This enhanced F-kNN outperforms other classical kNN. The incorporation of deep learning in radiation theory predicts pancreatic cancer for effective decision-making [11]. Deep learning neural network architectures are very useful in processing large-size images [12–14].

2 Methodology

A computer decision support system is needed in the medical field for diagnostic purposes, which helps clinician's decisions. Normally, without the intervention of a Computer Aided Diagnosis (CAD) system, the computed tomography, magnetic resonance imaging, and position emission tomography images suffer from high observer variations as it needs well-trained radiologists. Therefore to reduce it, CAD systems are widely used. Likewise, an F-kNN approach to detect the severity of pancreatic cancer is designed. The F-kNN algorithm takes input from the features extracted from the Jaundiced Eye Image. The goal is to build an efficient CAD system framework for pancreatic cancer diagnosis from clinical data, assessing the patient's risk level.

The F-kNN approach is carried out in five stages: image preprocessing, segmentation, RGB to HSV representation, feature extraction, and classification. The block diagram of the proposed CAD system for pancreatic cancer is shown in Fig. 2. The proposed approach includes the following process:

- a) Image preprocessing: image acquisition and resizing
- b) Segmentation: min-max color space threshold algorithm
- c) RGB to HSV representation.
- d) Feature extraction
- e) Classification: proposed F-kNN Approach

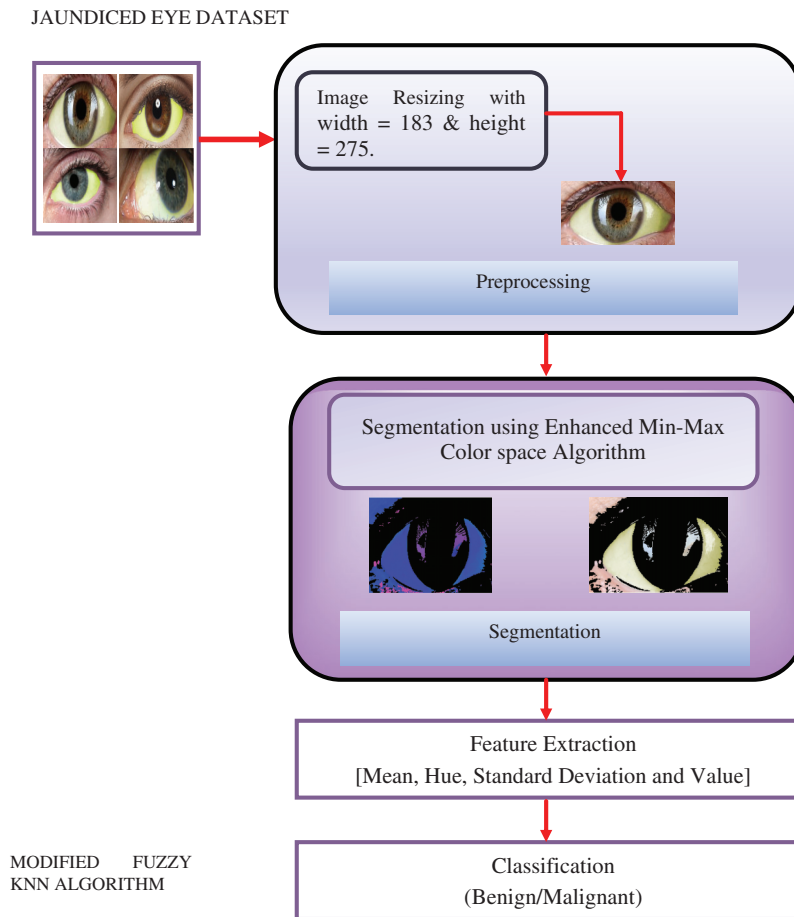


Figure 2: Block diagram of the proposed framework

2.1 Image Processing

The original eye image is loaded to detect whether the pancreas is affected by cancer through the ROI of the sclera. As carcinoma blocks the tube, which is named the common bile duct that releases bile into the intestine, it helps develop bilirubin within the blood. Jaundice turns the skin and the eyes (the sclera region) yellow, referred to as jaundice. The same blockage causes the urine to be dark yellow, light-colored stools, and itching sensation. In the eye image, we segment the sclera part to check whether the person is affected by cancer or not through the sclera's yellow part.

The eye image must be pre-processed before it is segmented. The input image is then resized to ensure all the images are in the same size for algorithm compatibility. The preprocessing leads to removing unwanted pixel information from the image. The image should be resized into the same size for further processing. The image's size is $W = 183$ and $L = 275$, as shown in Fig. 3.



Figure 3: (a) Original image and (b) Resized image

2.2 Segmentation

Segmentation is the process of removing a particular portion of Image. The jaundice is affected by the white part in the eye, so it should be segmented using the min-max sclera segmentation. Sclera segmentation is a very important process [15]. Min-max color space algorithm works well by converting the image into the color spaces as RGB split them into their respective channels and then find the min/max value using the minMaxLoc function by minimum and maximum value, the white part is sclera part is segmented from the eye image. Since the yellowish part buildup by bilirubin needs to be extracted from whites of the eye, this min-max algorithm is used to separate the color, which produces much accuracy. The min-max color space threshold algorithm's accuracy is validated using the canny edge algorithm, and k-means algorithm. Accurate segmentation of medical images leads to the exact location of the tumour [16]. After making validation, results found that the proposed technique produces much better accuracy than other algorithms. The validation of segmentation techniques with the k-means algorithm, canny edge algorithm and min-max color space algorithm is shown in Fig. 4. The canny edge operator is widely used in locating the region of interest [17]. The algorithm is as follows:

Algorithm 1: Min-max color space threshold algorithm for sclera segmentation

Input: Eye image

Output: Sclera ROI.

Step 1: Image smoothing.

Step 2: Resize the input image with the size of height 183, width 275.

Step 3: Segment the sclera part from the image using the min-max color space threshold algorithm.

Step 3.1: Define thresholds for each channels based on histogram. channel1Min = 180.000; channel1Max = 255.000; channel2Min = 172.000; channel2Max = 244.000;

Step 3.2: Create a segmented image based on chosen histogram thresholds.

Step 3.3: Initialize masked output image based on input.

Step 3.4: Remove the empty values present in the image by using any function.

2.3 RGB to HSV Conversion

After the affected sclera is segmented, the image's RGB color representation is converted into HSV color space. The HSV is the widely preferred model over RGB because the human eye tends to recognize color similar to the HSV model describes the color. RGB stands for Red, Green, Blue, and HSV is described in terms of Hue, Saturation, and Value. The Hue represents the type of color. It can be described in angles in a circle. The circle has 360-degree rotation with values normalized from 0 to 255, and 0 is the redcolor. The RGB representation of the image is converted into HSV color space to extract color features. The algorithm for the conversion is shown below.

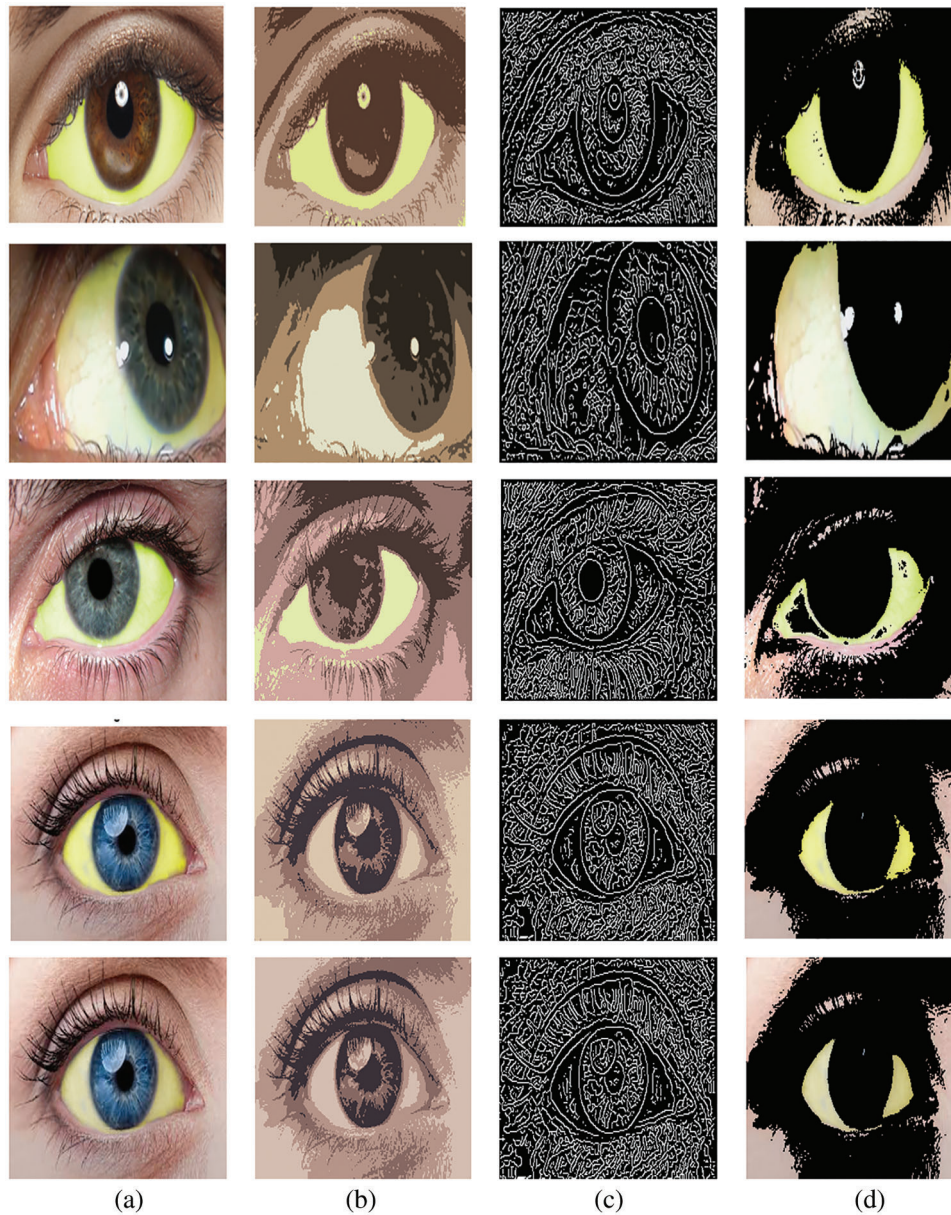


Figure 4: (a) Original image (b) Segmented image by k-means (c) Segmented image by canny edge operator, (d) Segmented image by min-max color space threshold

Algorithm 2: RGB to HSV conversion

Step 1: Divide the values of R,G, B by 255. $R' = R/255$, $G' = G/255$, $B' = B/255$

Step 2: $C_{\max} = \max (R', G', B')$

Step 3: $C_{\min} = \min (R',G',B')$

Step 4: Difference = $C_{\max} - C_{\min}$

Step 5: Hue calculation:

(Continued)

Algorithm 2: (continued)

$C_{\max} = C_{\min} = 0$, then $H = 0$

$h = (60 * ((G' - B') / \text{Difference}) + 360) \bmod 360$, if $C_{\max} = R'$

$h = (60 * ((B' - R') / \text{Difference}) + 120) \bmod 360$, if $C_{\max} = G'$

$h = (60 * ((R' - G') / \text{Difference}) + 240) \bmod 360$, if $C_{\max} = B'$

Step 6: **Saturation calculation:**

if $C_{\max} = 0$, then $s = 0$

if $C_{\max} \neq 0$ then $s = (\text{Difference} / C_{\max}) * 100$

Step 7: **Value Calculation:** $v = C_{\max} * 100$

The conversion from RGB to HSV of the affected eye image is shown in Fig. 5.

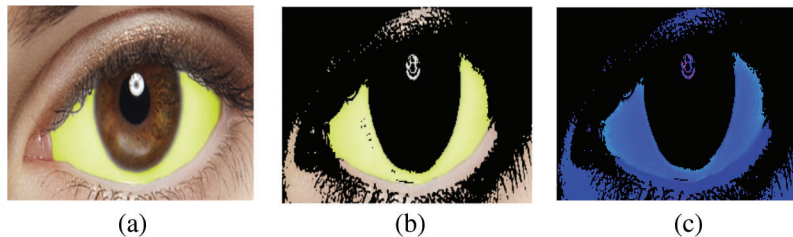


Figure 5: (a) Original image, (b) Segmented image by min-max color space threshold, (c) RGB image converted to the HSV color space

2.4 Feature Extraction

Feature extraction is a dimensionality reduction technique representing the interesting parts of the image in a compact form. After the segmented image is converted into HSV form, feature extraction is performed. 'Hue' indicates the original color, 'Saturation' denotes the quantity to which the corresponding color is mixed with white, and 'Value' denotes the quantity to which the corresponding color is mixed with black (Gray level). In RGB, the color information cannot be separated from luminance. HSV or Hue Saturation Value is used to separate image luminance from color information. Using the HSV, the mean, standard deviation, mean Value, and mean Saturation of each channel is calculated and used as features. This research mainly focuses on the classification of pancreatic cancer using the proposed kNN Approach. Tab. 1 depicts mean Hue, mean Saturation, mean value, and mean standard deviation extracted from the eye images and the graphical representation of features extracted in Fig. 6. The algorithm for feature extraction is shown below.

Algorithm 3: Feature extraction

Step 1: The input eye image is loaded.

Step 2: The image should be resized into the same size for further processing. The size of the image is $w = 183$ $L = 275$.

Step 3: The affected sclera part is extracted using the segmentation of the min-max color space algorithm.

Step 4: The RGB image is converted into HSV color space for feature extraction.

Step 5: The features like mean Hue, mean Saturation, mean value, and mean Standard Deviation are calculated.

Table 1: Comparison of proposed F-kNN with conventional kNN and fuzzy kNN (accuracy)

Image No	Mean hue	Mean saturation	Mean value	Mean standard deviation
I ₁	0.0382	0.0390	0.2212	0.0432
I ₂	0.0545	0.0596	0.3158	0.0673
I ₃	0.0332	0.0642	0.2407	0.0610
I ₄	0.0503	0.0438	0.1937	0.0608
I ₅	0.0297	0.0981	0.3143	0.0506
I ₆	0.0905	0.0283	0.2485	0.1025
I ₇	0.0322	0.0120	0.0685	0.0719
I ₈	0.0439	0.0382	0.3119	0.0551
I ₉	0.1138	0.0697	0.3760	0.0415
I ₁₀	0.0245	0.0606	0.1999	0.0570
I ₁₁	0.0792	0.0397	0.3440	0.0397
I ₁₂	0.0620	0.0585	0.2708	0.0494
I ₁₃	0.0628	0.0352	0.2286	0.0618
I ₁₄	0.0550	0.0508	0.2978	0.0481
I ₁₅	0.0186	0.1329	0.3775	0.0495
I ₁₆	0.1095	0.0503	0.2144	0.0632
I ₁₇	0.0459	0.0733	0.2713	0.0509
I ₁₈	0.0294	0.1040	0.3165	0.0515
I ₁₉	0.0477	0.0511	0.1865	0.0659
I ₂₀	0.0341	0.0796	0.2458	0.0629

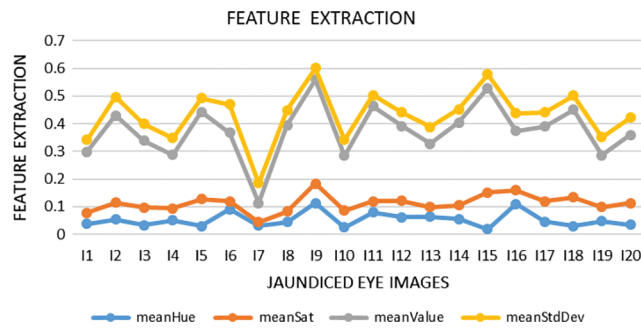


Figure 6: Extracted features such as mean Hue, mean Saturation, mean Value, and mean Standard Deviation

2.5 Classification

A brief description of F-kNN, as explained in [18] by is given below. In F-kNN, the membership values of samples are assigned to various categories are as follows [19]:

$$u_i(x) = \frac{\sum_{j=1}^k u_{ij} \left(1/\|x - x_j\|^{2/(m-1)} \right)}{\sum_{j=1}^k \left(1/\|x - x_j\|^{2/(m-1)} \right)} \quad (1)$$

where i takes the values from $1, 2 \dots c$ and j take the values from $1, 2 \dots c$ where c indicates the number of classes and k indicates the nearest neighbours.

Euclidean distance is the distance measure used to find the distance between x and its j^{th} neighbour, x_j . u_{ij} represents the degree of membership of the pattern x_j from the training set to the particular class i among the k nearest neighbours [20].

$$x_{ij}(x_k) = \begin{cases} 0.51 + \left(\frac{n_j}{K}\right) * 0.49, & \text{if } j = i \\ \left(\frac{n_j}{K}\right) * 0.49, & \text{if } j \neq i \end{cases} \quad (2)$$

The value of n_j represents the number of neighbours of class, the membership values calculated using Eq. (2) must satisfy the below equation.

$$\sum_{i=1}^C \mu_{ij} = 1 \quad (3)$$

where j takes the values from $1, 2 \dots n$, and C is the number of classes.

$$0 < \sum_{j=1}^n \mu_{ij} < n \quad \mu_{ij} \in [0, 1] \quad (4)$$

$$C(x) = \arg \max_{i=1}^C (u_i(x)) \quad (5)$$

In this methodology, sigmoid weighted F-kNN is proposed to optimize the feature weights and value of k . This novel work analyses the eye image to find the patient is affected by second-level pancreatic cancer using the proposed machine-learning algorithm called modified K-Nearest Neighbor. The existing kNN algorithm is modified using the weighted sigmoid function. Sigmoid Weighted kNN is one of the variations of the kNN method. It uses K nearest neighbours' concept, irrespective of their classes, and then uses weighted votes from each sample. The result shows the classification of pancreatic cancer level-2 intensity. The general framework of enhanced F-kNN is shown in Fig. 7.

The algorithm for classification is shown below.

Algorithm 4: Classification

Step 1: Determine the number of k , where K is the number of nearest neighbours.

Step 2: Compute K using the Euclidean distance formula.

$$(p, q) = d(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (6)$$

Step 3: Find out the membership of K -nearest neighbours.

(Continued)

Algorithm 4: (continued)

Step 4: Sigmoid Weight kNN uses weighted votes from each sample.

Step 5: Add the membership.

Step 6: Assign the class label with maximum membership.

Step 7: Using the label from the maximum membership result for the level of pancreatic cancer is shown

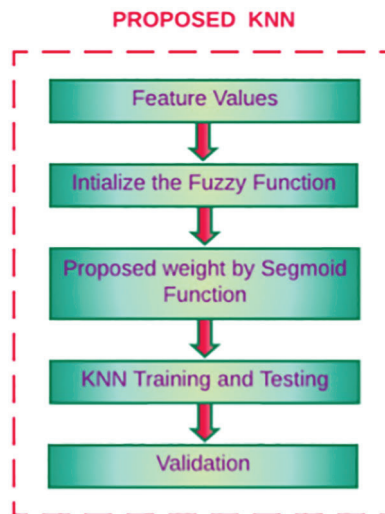


Figure 7: Detection of pancreatic cancer using proposed F-kNN

3 Experimental Analysis and Results

The dataset contains the four-feature attribute, in which each class label is assigned to have 3 values, 1, 2, and 3. The number of training instances is 42, and the testing instances are 18 instances. In percentage, 70 percent used as training and 30 percent used as testing. Tab. 2 shows the performance of the proposed F-kNN with conventional kNN and fuzzy kNN.

Table 2: Comparison of proposed F-kNN with conventional kNN and fuzzy kNN

Value of K	kNN		Fuzzy kNN		Proposed methodology	
	Accuracy	Error loss	Accuracy	Error loss	Accuracy	Error loss
k = 3	0.40	0.60	0.31	0.69	0.98	0.02
k = 4	0.42	0.58	0.38	0.62	0.97	0.03
k = 5	0.40	0.60	0.27	0.73	0.97	0.03
k = 7	0.36	0.64	0.36	0.64	0.96	0.04

All the results are performed in an Intel Core i3 Duo processor and MATLAB 2020. Here the comparison of the proposed F-kNN approach with Fuzzy kNN and Conventional kNN in terms of accuracy is done. k takes four different values, i.e., $k = 3, 4, 5, 7$ where k represents the number of nearest

neighbours used for the classification process [21]. The graphical representation of the analysis of accuracy is summarized in Fig. 8. The error rate calculated is summarized graphically in Fig. 9.

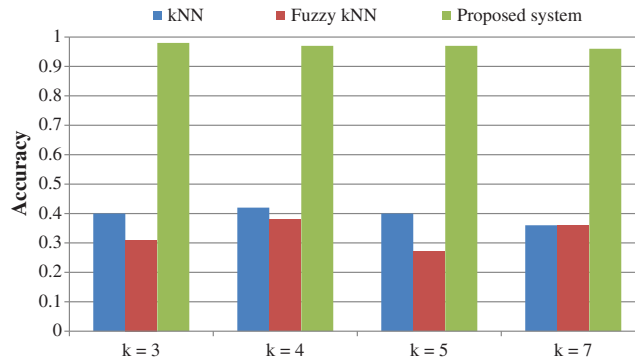


Figure 8: Graphical representation of analysis of accuracy for different values of k

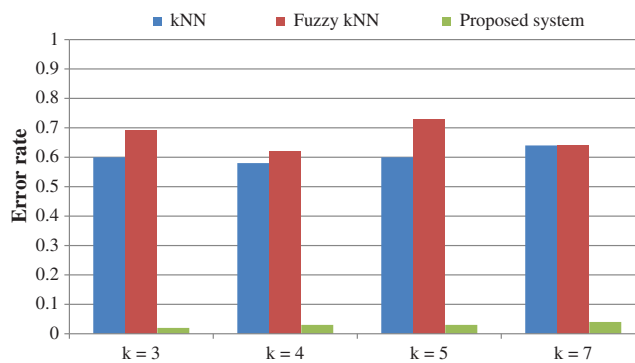


Figure 9: Error rate for different values of k

From the statistics and the graph obtained, it is concluded that the proposed F-kNN outperforms well compared to fuzzy kNN and conventional kNN algorithms. The extracted features are used in the proposed F-kNN for training and testing the samples to check whether the person is affected by pancreatic cancer by jaundice or not. The proposed F-kNN is a novel one that is different from the conventional kNN. The sigmoid-based fuzzy kNN is used in the proposed model. By changing the sigmoid weighted kNN, the accuracy level is improved. By trying the sample dataset, the accuracy improved by 98% percentage for k = 4.

4 Conclusions

An efficient decision support system is developed to detect pancreatic cancer using the highly related symptom with pancreatic cancer. i.e., jaundice. Experimental results show that the proposed Fuzzy kNN model efficiently classifies the various cancer stages and produces 97% accuracy. The proposed F-kNN model outperforms well compared to classical kNN and Fuzzy kNN, respectively. Here the significance of sclera segmentation with min-max color space, canny edge, and k-means algorithm with features

extracted is analyzed and used for risk assessment. Feature extraction is done when the RGB representation is converted into HSV form. This proposed decision support system is built based on recognizing sclera from eye images through machine learning algorithms and extracting features from sclera through color descriptors. Hence, mean Hue, mean Saturation, mean value, and mean Standard Deviation are calculated. This analysis helps clinicians in better decision-making which help in the future. Also, the error loss that occurred during segmentation and classification can be avoided, which still improves accuracy. In the future, this kind of classification model can be used for other kinds of symptoms specific to pancreatic cancer to check the correlations.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] C. L. Wolfgang, J. M. Herman, D. A. Laheru, A. P. Klein, M. A. Erdek *et al.*, “Recent progress in pancreatic cancer CA,” *A Cancer Journal for Clinicians*, vol. 63, no. 5, pp. 318–348, 2013.
- [2] American Cancer Society. *Cancer Facts & Figures 2016*. Technical report, 2016.
- [3] W. Muhammad, G. R. Hart, B. Nartowt, J. J. Farrell, K. Johung *et al.*, “Pancreatic cancer prediction through an artificial neural network,” *Frontiers in Artificial Intelligence*, vol. 2, pp. 2, 2019.
- [4] T. Sen and S. Das, “An approach to pancreatic cancer detection using artificial neural network,” in *Proc. of the Second Int. Conf. on Advances in Computer, Electronics and Electrical Engineering*, Mumbai, India, pp. 56–60, 2013.
- [5] A. Mariakakis, M. A. Banks, L. Phillipi, L. Yu, J. Taylor *et al.*, “Biliscreen: Smartphone-based scleral jaundice monitoring for liver and pancreatic disorders,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 2, pp. 1–26, 2017.
- [6] J. Shah, S. Surve and V. Turkar, “Pancreatic tumor detection using image processing,” *Procedia Computer Science*, vol. 49, pp. 11–16, 2015.
- [7] <https://www.pancan.org/news/the-acts-about-jaundice-and-pancreatic-cancer/>.
- [8] J. Shah, S. Surve and V. Turkar, “Pancreatic tumor detection using image processing,” in *4th Int. Conf. on Advances in Computing, Communication and Control*, Mumbai, India, pp. 11–16, 2015.
- [9] A. D. Wibawa and M. H. Purnomo, “Early detection on the condition of pancreas organ as the cause of diabetes mellitus by real time iris image processing,” *IEEE Asia Pacific Conference on Circuits and Systems*, pp. 1008–1010, 2006.
- [10] I. P. D. Lesmana, I. K. E. Purnama and M. H. Purnomo, “Abnormal condition detection of pancreatic beta-cells as the cause of diabetes mellitus based on iris image,” in *2nd Int. Conf. on Instrumentation, Communications, Information Technology, and Biomedical Engineering*, Bandung, Indonesia, pp. 150–155, 2011.
- [11] W. Zhao, L. Shen, B. Han, Y. Yang, K. Cheng, “Markerless pancreatic tumor target localization enabled by deep learning,” *International Journal of Radiation Oncology* Biology* Physics*, vol. 105, no. 2, pp. 432–439, 2019.
- [12] H. Mohsen, E. S. A. El-Dahshan, E. S. M. El-Horbaty and A. B. M. Salem, “Classification using deep learning neural networks for brain tumors,” *Future Computing and Informatics Journal*, vol. 3, no. 1, pp. 68–71, 2018.
- [13] S. Justin and M. Pattnaik, “Skin lesion segmentation by pixel by pixel approach using deep learning,” *International Journal of Advances in Signal and Image Sciences*, vol. 6, no. 1, pp. 12–20, 2020.
- [14] C. Srinivasan, S. Dubey and T. R. Ganeshbabu, “Complex texture features for glaucoma diagnosis using support vector machine,” *International Journal of MC Square Scientific Research*, vol. 7, no. 1, pp. 81–92, 2015.
- [15] P. Rot, M. Vitek, K. Grm, Ž. Emeršič, P. Peer *et al.*, “Deep sclera segmentation and recognition,” in *Handbook of Vascular Biometrics*, Cham: Springer, pp. 395–432, 2020.

- [16] D. J. Reddy, T. A. Prasath, M. P. Rajasekaran and G. Vishnuvarthanan, "Brain and pancreatic tumor classification based on GLCM k-NN approaches," in *Int. Conf. on Intelligent Computing and Applications*, Chennai, India, pp. 293–302, 2019.
- [17] A. Helwan, "ITDS: Iris tumor detection system using image processing techniques," *International Journal of Scientific and Engineering Research*, vol. 5, no. 11, pp. 76–80, 2014.
- [18] J. M. Keller, M. R. Gray and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 4, pp. 580–585, 1985.
- [19] W. Shang, H. Huang, H. Zhu, Y. Lin, Z. Wang *et al.*, "An improved kNN algorithm-fuzzy kNN," in *Int. Conf. on Computational and Information Science*, Xi'an, China, pp. 741–746, 2005.
- [20] Z. Cai, J. Gu, C. Wen, D. Zhao, C. Huang *et al.*, "An intelligent Parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach," *Computational and Mathematical Methods in Medicine*, vol. 2018, pp. 1–24, 2018.
- [21] S. Taneja, C. Gupta, S. Aggarwal and V. Jindal, "MFZ-kNN—a modified fuzzy based k nearest neighbor algorithm," in *Int. Conf. on Cognitive Computing and Information Processing*, Noida, India, pp. 1–5, 2015.