Pancreatic Cancer Detection Using Convolutional Neural Network

Ferdin Meriya Benitta, Dr. A. Lourdes Mary
Department of Computer Science and Engineering, Rohini College of Engineering and technology, Palkulam, Kanyakumari district, Anna University, Chennai 600025
Associate Professor, Department of Computer Science and Engineering, Rohini College of Engineering and technology, Palkulam, Kanyakumari district, Anna University, Chennai 600025

Abstract-One of the worst cancer diagnoses that a patient endures is pancreatic cancer, which has a very low 5 year chance of surviving. The primary cause of most cases of this syndrome is pancreatic cancer. Many cancer sufferers are now able to spot aberrations at the beginning of the disease because of medical image scans. It is challenging to spread the technology because of the high cost of the required infrastructure and equipment, which keeps it out of many people's price ranges. In this paper, image processing and PSO-CNN are used to identify pancreatic cancer in images. The Wiener reduction filter is used to remove noise from pictures within the image preparation stage. The Fuzzy C- means (FCM) algorithm divides the image into its individual parts using an organizing method. Image segmentation facilitates the process of detecting objects in an image while recognizing the areas of concern. To extract essential data from digital photos, the GLCM approach is used. The classification is carried out using the techniques PSO- CNN, naive Bayes, and AdaBoost. The PSO-CNN algorithm has higher precision, sensitivity, and specificity.

Keywords: PSO-CNN, wiener filter, FCM clustering, GLCM.

I. INTRODUCTION

The 5-year rate of survival for pancreatic cancer is only 9.4%, making it one of the most serious cancer diagnoses that is possible across the globe. Most instances of this illness are linked to pancreatic cancer. The pancreas, that is the digestion system's second-to-last organ after the liver, constitutes one of such organs. Certain kinds of fish have heads, bodies, and tails that resemble others in appearance only slightly. Even when fully developed to adult size, it is only approximately 5 centimeters (2 inches) wide [1, 2]. Once exocrine cells in the pancreas expand uncontrollably, a disease called pancreatic adenocarcinoma is appear. The most prevalent kind of pancreatic cancer is this specific one. The ducts and glands which are referred to as exocrine organs are made of exocrine cells. The pancreas contains these organs, including SNR, PSNR and RMSE [3-5]. The outcomes of this analysis were contrasted with the typical audio pattern. It is being demonstrated that by utilizing this method, the fundamental components of a medical image may be preserved while at the same time, the image noise may be greatly reduced. The MRI pictures are going to be more precise when utilizing this experimental technique. According to [6, 7], applying the median and mean filters proved necessary for denoising medical images. Comparable to the mean filter, the median filter is a kind of image filter that lowers the quantity of distortion in a photograph. The median filter accomplishes this without damaging the fine features of the image, in contrast to the mean filter. An intensity level's output is modified to be the median of the other nearby levels of brightness whenever the median filter is fitted to a particular pixel. The wiener filter clearly outperforms the competition with regard to of filtering and general functionality. Images with the wiener filter applied have better pixel quality than those with the other filter. In this specific research's [8, 9], which looked into the segmentation of 4D CT scans for nodules (collected at various times), nodules were the object of research. In an attempt to shrink the size of the power function, it has been suggested that this standard, which considers the similarity the images of every stages are to one another, be incorporated into the chart's cut method. The disadvantage of this approach is that it necessitates manual segmentation at the very beginning of the procedure. Deep learning and domain-dependent expertise are combined in the resulting combination cost function [10, 11]. The second group of researchers suggests an approach that starts with CNN and ends with graph cut segmentation. The core of this technology is the application of CNN as a filter to remove inaccurate results. A novel technique for segmenting colour images using local histogram equalization (LHE) and K-means clustering is presented in [12]. LHE is a technique for improving colour photos that alters pixels while utilizing the data present in the image. In the end, the K-means clustering approach is used to segment a colour image. The method is then compared to some other proven methods, including subtractive clustering, K-means, and fuzzy C-means techniques. The method of data collection and the qualities of the data play a role in how accurately a machine learning system predicts cancer [13, 14]. Support vector machines, random forests, naive Bayes, decision trees, K-nearest neighbors, artificial neural networks, fuzzy neural networks, RBFN, shuffled frog leaping with levy flights, particle swarm optimization The method of data collection and the qualities of the data play a role in how accurately a machine learning system predicts cancer [15]. Support vector machines, random forests, , K-nearest neighbors, naive Bayes, RBFN, decision trees Artificial Neural Networks (ANN), fuzzy neural networks, shuffled frog leaping with levy flights, back propagation neural networks, multilayered perceptron, and CNN are just a few of the techniques used to classify data. The results of this research show that the PSO-CNN method is the most successful way for predicting cancer illness from a given dataset., back propagation neural networks, multilayered perceptron, and CNN are just a few of the techniques used to classify data. The results of this research show that the (PSO-CNN) machine learning method is the most successful way for predicting cancer illness from a given dataset.

The main contribution of this work is to propose and PSO-CNN classifier for detecting the pancreatic cancer in early stage. With the implementation wiener filter the raw input image is preprocessed. For the purpose of reduce the processing time, the segmentation stage is used. Here, the Fuzzy C-means algorithm is utilized to segment the image. From the segmented image, the high ranked features are extracted with the aid of GLCM approach. Finally, the proposed PSO-CNN effectively predict the cancer from pancreatic organ.

II. PROPOSED SYSTEM

A Wiener filter is used during the process of the pre-processing to remove noise from the input image. Following pre-processing, the picture with the noise eliminated is supplied into a segmentation block to be divided into numerous segments for subsequent processing. This system employs the fuzzy-c-means clustering technique and for edge identification, a fuzzy filter is employed. To simplify categorization, the segmented images have been processed using the feature extraction technique to extract several features and pick the necessary characteristics. For feature extraction and choosing features in the present work, the GLCM approach is employed. In order to effectively categorize the image, the chosen characteristics are finally fed into the CNN classifier. The categorized image has a higher level of statistical accuracy. Figure 1 illustrates the block diagram for proposed system.

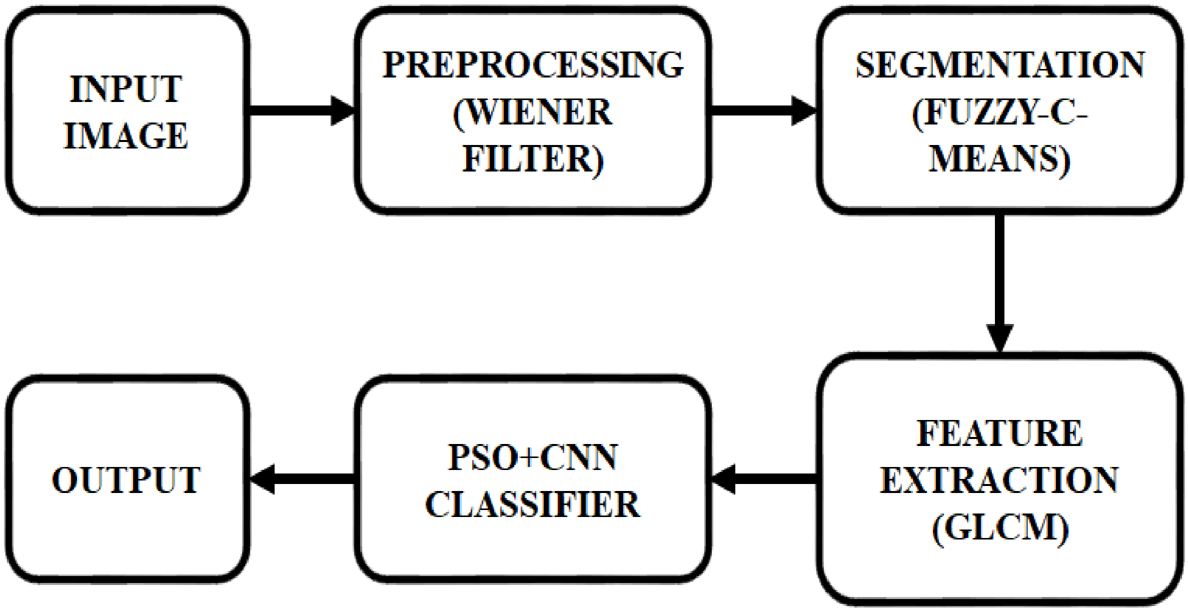

Figure 1 Proposed system

II. PROPOSED SYSTEM DESCRIPTION

A. WEINER FILTER

To remove unwanted interference from the corrupted signals, Wiener filter have been employed. The filters are set up to produce the desired frequency response, although establishing a Wiener filter isn't particularly simple. The outputs from the two primary Wiener filter components one knowing the spectral characteristics of the original signal and the other looking for a linear time invariant filter—would likely be identical to the original signal. The primary purpose of the Wiener filter is to reduce the effectiveness of the frequency of the weak chirp signal and to specify the frequency of the dominant chirp signal over the noise signal. If the chirp signal contains noise, then the obtained equation is,

$$R(u) = \frac{H(u)^*}{|H(u)|^2+K} \quad (1)$$

Where the signal's Fourier transform is $H(u)^*$ and R (u) is the output. A high pass filter is the counterpart to a blurred image. The low frequency aspect of the Wiener filter is connected to the variable K of the Wiener filter. The inverse filter and variable K are responsible for the low pass filter and high pass filter, accordingly, of the Wiener filter's behavior as a band pass filter. The key characteristic of a Wiener filter is that it allows the dependent on time noise factor to vary from image to image thus rendering it negatively correlated with SNR. This is guarantee that noise-only images are suppressed more severely and the chirp signal will be suppressed less severely during segmentation. Therefore, the noisy areas in the provided input image are removed using this filter.

B. FUZZY C MEANS ALGORITHM

The division of a picture into multiple subgroups (pixels) known as image objects is an operation known as segmentation. The segmentation process in this system uses the fuzzy-c-means technique that lessens the complexities of the image. A common data clustering approach is the fuzzy c-means algorithm (FCM), which determines the degree to which each data point belongs to a cluster through a membership grade. When FCM divides a set of n vectors, $X j, j = 1, \dots, n,$ into c fuzzy groups, it identifies a cluster centre in each group that reduces the expense function determined by the distance**.** The fuzzy clustering enables partial affiliation of the thing in question to all clusters and take into consideration cluster duplication. In simple terms, any collection with a degree of membership between ° and 1 can have a given data point as a member. The approach, which uses a repeated clustering approach, decreases the calculated within group sum of squared error objective function J_{FCM} to obtain an ideal c partitioning. Fuzzy c-partition, the outcome of a fuzzy clustering, is expressed by the formula that follows equation.

$$J_{FCM} = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^q \, d^2(x_k, v_i) \quad (2)$$

Where X is the data set in the p-dimensional vector space and $X = \{x1, x2, \cdots, xn\} \subseteq R\,p$.The quantity of data pieces is n. With 2 ≤ c < n is the total number of clusters.The degree of membership of $x_k$in the $i\,th$ cluster is represented by u_{ik}, q is a fuzzy membership's weighting exponent. The distance between an object and the cluster centre is measured as $d^2(x_k, v_i)$ is the prototype of cluster i centre. With help of proposed FCM technique, the input images are segmented.

C. GLCM METHOD

For texture analysis, the Grey Level Co-occurrence Matrix (GLCM) is used. The reference and neighbour pixels are the two pixels that we take into consideration at once. Prior to computing the GLCM, the specified certain spatial connection among the reference and neighbouring pixels. For instance, we may say that the neighbour is any pixel that is 1 pixels to the right of the current pixel, 3 pixels above, or 2 pixels diagonally (towards one of NE, NW, SE, or SW) from the reference. We build a GLCM of size (Range of Intensities x Range of Intensities) with all of its starting values set to zero after a spatial relationship has been established. For instance, a single channel image with 8 bits will have a GLCM of 256x256.

By combining it to its inversion and normalizing it, the matrix may be adjusted symmetrical so that each cell represents the likelihood of that pair of intensities transpiring in the image. After calculating the GLCM, use of matrix's texture attributes for representing the textures in the image.

GLCM ALGORITHM

$Input: Image$

$Output: Vector\ of\ textual\ features$

$Begin$
$Step\ 1$: $Call\ the\ Algorithm\ of\ computing\ GLCM\ matrix\ in\ four\ direction\ with\ distance\ d = 1$
$Step\ 2$: $Call\ the\ Algorithm\ of\ normalizing\ each\ GLCM\ matrix.$
$Step\ 3$: $For\ each\ GLCM\ matrix\ in\ certain\ angle$
$Step\ 4$: $Calculate\ textural\ features\ according\ to\ their\ equations.$
$Step\ 5$: $store\ computed\ features\ in\ a\ vector.$
End

D. CNN BASED PSO CLASSIFIER

The CNN predetermined weights are established by the CNN with the process of back propagation. Given that the MNIST dataset consists of images with a size of 28x28 pixels, a CNN with two hidden layers and 784 input nodes is used. Additionally, the CNN has two 2D max pool layers and two 2D convolutional layers. One hot vector is encoded as the output. The following stages are main steps in the algorithm. Pre-PSO training involves collecting weights and converting them into particles. PSO training includes updating CNN with weight values from phase. Prediction accuracy and results are calculated. Pre-PSO Training: Weight Capture and Weight Conversion into Particles": Since the weights must be optimized, train the CNN via back propagation. They are taken in as tensors and transformed into numpy arrays so that PSO can further optimize them. The acquired weights are transformed into particles for the following phase since the PSO algorithm optimizes particles.

After the determination of the memory, social, cognitive, particle number, stopping condition, and maximum epochs values. To achieve consistency, PSO optimization uses the CNN loss function as the aim for PSO training to explore the hyper plane for optimized solutions. The final calculation of the outcomes requires the development of an additional CNN with weight values specified with the numbers of weights received from phase. The CNN is not be trained using the output in this phase.

The finished outcome is the result of the CNN as developed, and its precision as well as the expense value are computed from the identical value. The system gathers weights after training with CNN and uses PSO to optimize these because the goal is to boost reliability. The suggested CNN algorithm's weight vectors can be extracted as tensors using Tensor flow, an inexpensive open-source tool designed exclusively for machine learning. The PSO training module receives the values after that and trains and updates them. The revised weight values matrix is transformed back to a tensor to update the CNN weights when the PSO training model converges.

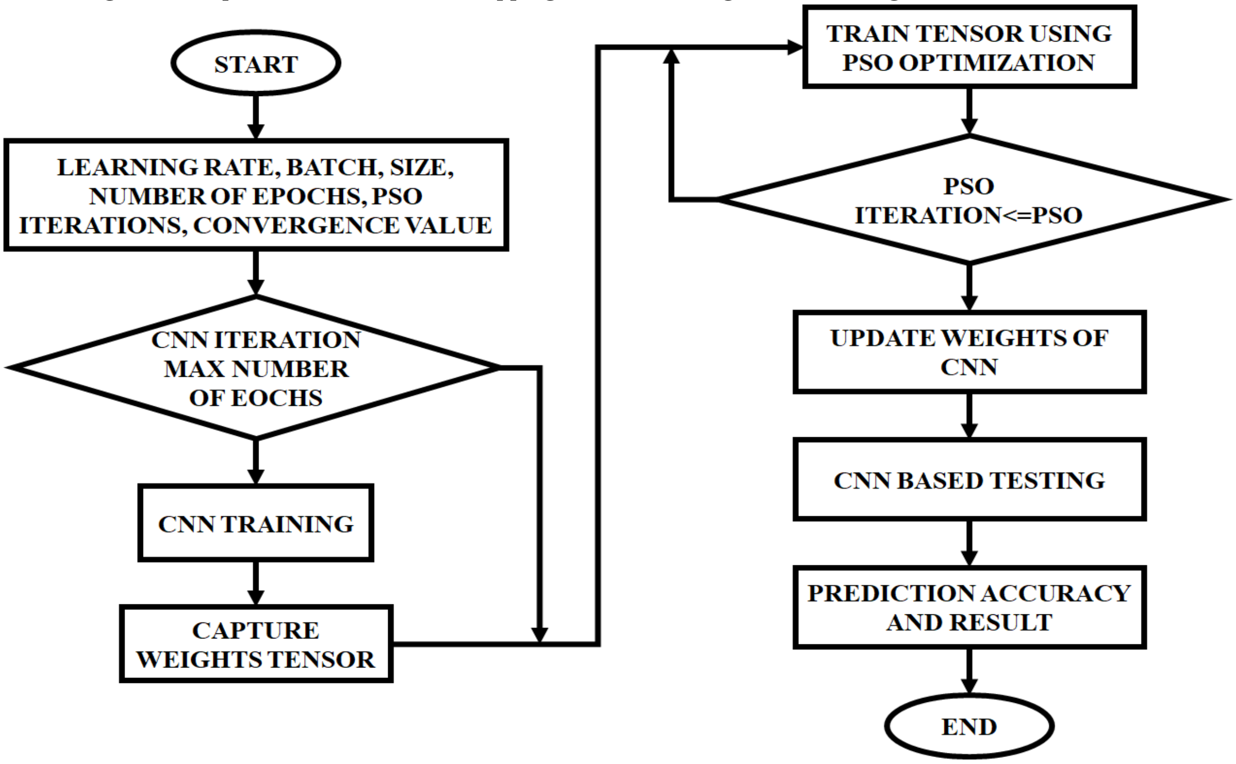

Figure 2 Suggested PSO-CNN flowchart

IV. RESULTS AND DISCUSSION

The proposed system implemented using MATLAB platform to verify its performance and the corresponding plots are presented below. Figure 3 indicates the input image of pancreas. With the implementation wiener filter in preprocessing, the noises from input image is removed effectively.

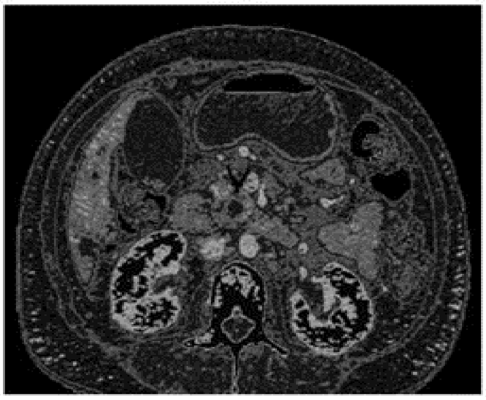

Figure 3 Pancreas input image

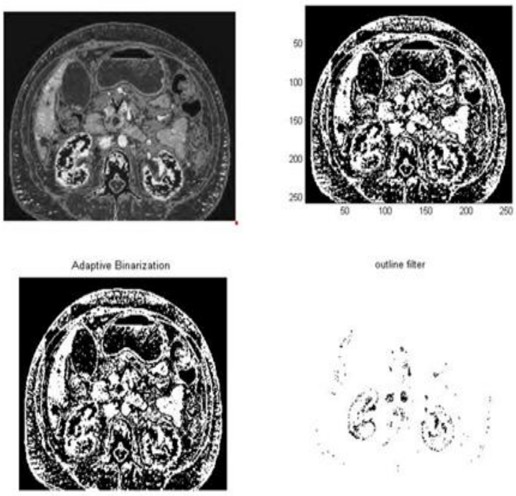

Figure 4 Preprocessed by using wiener filter

Following filtering, the image is segmented using the Fuzzy-C-Means approach. Thereby, Figure 5 displays the segmented image. Similarly, the Figure 6 represents the image for edge detection.

Figure 5 Segmented image using FCM

Figure 6 Edge detection

Figure 7 (a) Predicted result (b) Accuracy

The proposed PSO-CNN classifier effectively detect the pancreatic cancer with high accuracy, which is illustrated in Figure 7 (a) & (b).

V. CONCLUSION

Pancreatic cancer is considered to be one of the most severe forms of cancer a patient gets in their lifetime. Pancreatic cancer is the primary risk factor in the majority of cases of this condition. To avoid this early, the present article suggests using PSO-CNN and image processing to detect pancreatic cancer in photographs. During the image preparation process, noise is removed from photographs using the Wiener reduction filter. The Fuzzy C- Means (FCM) algorithm uses an organizing strategy to separate the image into its component pieces. Image segmentation makes it easier to find things in an image while also identifying problem regions. The GLCM technique is utilized to extract crucial information from digital pictures. Greater specificity, sensitivity, and precision are features of the PSO-CNN algorithm.

REFERENCES

[1] Chu, Linda C., Michael G. Goggins, and Elliot K. Fishman. "Diagnosis and detection of pancreatic cancer." *The Cancer Journal* 23.6 (2017): 333-342.

[2] Pereira, Stephen P., et al. "Early detection of pancreatic cancer." *The lancet Gastroenterology & hepatology* 5.7 (2020): 698-710.

[3] Hussein, Sarfaraz, et al. "Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches." *IEEE transactions on medical imaging* 38.8 (2019): 1777-1787.

[4] Yang, Annan, et al. "Autophagy Sustains Pancreatic Cancer Growth through Both Cell-Autonomous and Nonautonomous MechanismsAutophagy and Pancreatic Cancer." *Cancer discovery* 8.3 (2018): 276-287.

[5] Chan-Seng-Yue, Michelle, et al. "Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution." *Nature genetics* 52.2 (2020): 231-240.

[6] Liu, Kao-Lang, et al. "Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: a retrospective study with cross-racial external validation." *The Lancet Digital Health* 2.6 (2020): e303-e313.

[7] Yokoyama, Seiya, et al. "Predicted Prognosis of Patients with Pancreatic Cancer by Machine LearningPrognosis of Pancreatic Cancer by Machine Learning." *Clinical cancer research* 26.10 (2020): 2411-2421.

[8] Ma, Han, et al. "Construction of a convolutional neural network classifier developed by computed tomography images for pancreatic cancer diagnosis." *World Journal of Gastroenterology* 26.34 (2020): 5156.

[9] Ko, Jina, et al. "Combining machine learning and nanofluidic technology to diagnose pancreatic cancer using exosomes." *ACS nano* 11.11 (2017): 11182-11193.

[10] Sinkala, Musalula, Nicola Mulder, and Darren Martin. "Machine learning and network analyses reveal disease subtypes of pancreatic cancer and their molecular characteristics." *Scientific reports* 10.1 (2020): 1-14.

[11] Nasief, Haidy, et al. "A machine learning based delta-radiomics process for early prediction of treatment response of pancreatic cancer." *NPJ precision oncology* 3.1 (2019): 25.

[12] Li, Zhongqiang, et al. "Detection of pancreatic cancer by convolutional-neural-network-assisted spontaneous Raman spectroscopy with critical feature visualization." *Neural Networks* 144 (2021): 455-464.

[13] Kenner, Barbara J., et al. "Early detection of pancreatic cancer: the role of industry in the development of biomarkers." *Pancreas* 46.10 (2017): 1238.

[14] Muhammad, Wazir, et al. "Pancreatic cancer prediction through an artificial neural network." *Frontiers in Artificial Intelligence* 2 (2019): 2.

[15] Atsawarungruangkit, Amporn, et al. "Artificial intelligence for pancreatic cancer detection: Recent development and future direction." *Artificial Intelligence in Gastroenterology* 2.2 (2021): 56-68.