Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) with Harmony Search Optimization (HSO) for Early Breast Cancer Detection System

K. Sangeetha[1*], S. Prakash[2]

Abstract

For women, most common cause of death is Breast tumour and in worldwide, it is the second leading reason for cancer deaths. Due the requirement of breast cancer's early detection and false diagnosis impact on patients, made researchers to investigate Deep Learning (DL) techniques for mammograms. There are four stages in this proposed HIRResCNN framework, namely, Pre-processing, reduction of dimensionality, segmentation and classification. From images, noises are removed using two filtering algorithms called Median and mean filtering in pre-processing stage. Then canny edge detector is used for detecting edges. Gaussian filtering is used in canny edge detector to smoothen the images. In the next dimensionality reduction stage, attributes are correlated using Principal Component Analysis (PCA) inclusive of related features. So, this huge dataset is minimized and only few variables are used for expressing it. In order to detect the breast cancer accurately, foreground and background subtraction is done in the third stage called segmentation stage. At last, for detecting and classifying breast cancer, a Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) is introduced, which integrates Harmony Search Optimization (HSO) to tune bias and weight parameters and classification accuracy is enhanced using HIRResCNN-HSO model. Strength of Recurrent Convolutional Neural Network (RCNN), Residual Network (ResNet) and Inception Network (Inception-v4), are combined in a powerful Deep Convolutional Neural Network (DCNN) model called HIRResCNN. using Mammographic Image Analysis Society (MIAS) dataset, various experiments are conduced and results are compared with other available techniques. Around 92.6% accuracy rate is produced using this proposed HIRResCNN classifier in finding breast cancer.

Key Words: Principal Component Analysis (PCA), Computer Aided Detection (CAD), Deep Convolutional Neural Network (DCNN), Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN), Breast Cancer, Mammography, Deep Learning (DL).

DOI Number: 10.14704/nq.2021.19.7.NQ21079 **NeuroQuantology 2021; 19(7):01-24**

Introduction

A familiar form of tumour is breast tumour and around the world, between women, it is the next foremost reason for demise (Ferley et al, 2015). Abnormality or uncontrolled division of breast cell tissues results in breast cancer. A large tissues lump are formed by these abnormal cells, which leads to a tumour (Mehdy et al, 2017). It was accounted for that 1.7 million instances of breast malignant growth were recognized on the planet in 2012.

Corresponding author: K. Sangeetha
Address: [1*]CSE Department, SNS College of Technology, Coimbatore, India; [2]CSE Department, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India.
[1*]E-mail: sangithaprakash@gmail.com
[2]E-mail: prakashdharsan@gmail.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 03 May 2021 **Accepted:** 30 June 2021

Breast malignant growth is the second reason for disease demise with the normalized death pace of 12.9 per 100,000, and its frequency has expanded throughout the years (Ghoncheh et al, 2016).

Breast tumour is one among basic driver of demise between ladies and possibly recognized as non-skin illness in ladies (Ayer et al, 2013). Breast cancer growth happens when the breast's cell tissues changes to weird and wildly isolated. These are huge piece of tissues, that turns into a cancer (Tarique et al, 2015). Such disarranges might efficiently treated as possibility that recognized before time. In this way, it is important to have suitable methods for viewing the largely prompt sign of Breast cancer growth.

According to the view under microscope, breast cancers are classified into various types (Abdelwahed et al, 2015). Carcinomas is a most common type of breast cancer. It begins with cells that line organs as well as in breast like tissues. Breast cancer is a carcinoma type, which is termed as adenocarcinoma. In glandular tissue, carcinoma starts to grow.

Sarcoma is another breast cancer type, which begins from muscle cells, connective tissue or fat. In certain conditions, combination of various cancer types or invasive mixture produces the single breast cancer. However, in some rare cases, tumor may be formed by tumour cells. Breast tumour is a familiar form of cancer among ladies (Rizzi et al, 2012).

Treatment options are increased by early detection as well as efficient breast cancer treatment and this leads to reduction morality rate. For breast tumour identification in an before time periods various modalities like MRI, ultrasound and mammography are used as a effective tools. In radiologists' interpretation, breast cancer's misdiagnosis is increased due to technical reasons having relation with human error and image quality.

For overcoming these restrictions, breast cancer's are detected automatically by developing CAD systems and they are classified as malignant and benign lesions. Between abnormal and normal tissues, performance of radiologists in computation and discrimination are enhanced using CAD systems.

In relation to breast cancer, CAD system focuses on design of reliable as well as accurate system for decreasing observational oversights and discrimination of malignant and benign lesions performance. For breast cancer diagnosis and detection, different modalities like biopsy histopathological images, MRI, ultrasound and mammography are used.

various imaging strategies is produced for before time location and healing of breast malignant growth and to lessen the quantity of passing, and several helped breast disease analysis techniques are utilized to build the analytic precision.

There are diverse imaging policies available for showing and finding of breast malignancy, highly important among them are mammography, ultrasound, and thermography. (Hasanien et al, 2014) Mammography is one maong the significant before time analysis techniques for breast malignant growth.

Breast screening is done using a dedicated image modality called mammography which utilizes low-dose X-ray in breast examination. In recent days, breast cancer are detected in early stages using an effective tool called mammography. In breast cancer diagnosis, best and most commonly used technique is mammography. In this low- dose x-ray system is used for examining breast. In some cases, tumor or cancerous mass are cannot be detected by radiologist, which leads to the use of CAD techniques for breast masses identification (Kim et al, 2014).

Microcalcifications and masses are soonest signs of breast malignant growth that is recognized using current period process. Microcalcifications are group of calcium stores that are minimum sized and is within the delicate breast tissues. For the most part, identification of masses in breast tissues is additionally testing contrasted with the recognition of microcalcifications, not just because of the enormous variety fit as a fiddle yet in addition since masses regularly display poor picture differentiate when utilizing mammography.

The trouble in order of kind and dangerous microcalcifications additionally causes a huge issue in clinical picture handling. as mammography is not exceptionally efficient for solid breasts, ultrasound or analytic sonography process are recommended (Brorders et al, 2012).

Over the most recent couple of decades, a few information mining and AI procedures have been created for breast malignant growth identification and grouping (Brorders et al, 2012), that is partitioned as three fundamental levels: preprocessing, feature extortion, and characterization. Pre-preparing which has extraordinary criticalness in mammographic picture investigation because of low class of mammograms as they are found at little portion of radiation whereas high measure of radiation might endanger individual's wellbeing.

Noise restoration, image smoothing, image quality are enhanced using various techniques (Talha et al, 2016). Highlight extraction and arrangement techniques have just been utilized to distinguish and characterize Regions Of Interest (ROI) in clinical pictures when all is said in done and breast tumors in digitized mammograms specifically. After extraction, picture properties, for example, perfection, coarseness, profundity, and consistency are removed by characterization (Bicket et al, 2017). Absence of early identification prompts a large number of ladies experiences agonizing, lower endurance rate and scar actuating medical procedures.

In this way, information investigation systems have become helpful partners for doctors when settling on malignancy determination choices. The current tumor include informational indexes were characterized into dangerous and kind sets independently. To build the precision and handle the significantly expanding tumor highlight information and data, various scientists have gone to Machine Learning (ML) and Deep Learning (DL) methods for anticipating breast malignancy (Dhahri et al, 2019).

Sadly, ML is as yet a field with high hindrances and regularly requires master information. Structuring a successful AI model including the phases of pre-preparing, highlight choice and order forms requires a lot of aptitudes and ability (Zheng et al, 2014).

For solving these issues, this proposed work discussed about HIRResCNN with HSO technique based on Deep CNN (DCNN) to classify breast tumor disease and it is termed as HIRResCNN-HSO. MIAS images are classified as malignant/ benign using this technique. There are four stages in this proposed HIRResCNN framework, namely, Pre-processing, reduction of dimensionality, segmentation and classification. In the proposed HIRResCNN-HSO framework, brief explanation about following steps are given:

1. Primarily the MIAS image data is received as input and it follows HIRResCNN-HSO Model for the breast cancer detection.
2. HIRResCNN-HSO Model pre-processing stage removing noise from images by using mean, median filtering algorithms, canny edge detector use Gaussian filter algorithm for smooth the image, result in the blurring of fine-scaled image edges.
3. Dimensionality reduction stage, PCA technique is applied to correlate attributes it gives better result, so that the huge dataset can be reduced and can be expressed within few number of variables.
4. Segmentation stage is easily and/or modify the depiction of image into somewhat which is significant and simple to process. Foreground and background subtraction using for accurate breast image detection.
5. Classification stage, Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) for breast tumour categorization and detection by integrating Harmony Search Optimization (HSO) for tuning the parameters of weight and bias HIRResCNN model, resulted improve classification accuracy.
6. Further the results of the HIRResCNN-HSO models are evaluated with the parameters of precision, recall, f-measure and accuracy.

HIRResCNN is a powerful Deep Convolutional Neural Network (DCNN) form that integrates strength of Inception Network (Inception-v4), Residual Network (ResNet) and Recurrent Convolutional Neural Network (RCNN). Remaining article is ordered as: related studies on breast cancer detection existing schemes are briefly discussed in section 2. The proposed methodology on breast tumour identification method using deep learning in section 3. Discussion about outcome are briefly explained in section 4. Conclusion with upcoming plan is described in section 5.

Literature Review

In medical field, prior identification of breast tumor are done using a mammography, which is followed by some other screening techniques. Mammography based techniques are presented in this literature with outcomes, reliability and affordability analysis. Clear explanation about demerits and merits of different breast cancer detection techniques are presented also.

Wang et al (2018) proposed a system gestalt hypothesis to identify breast masses in digitized mammograms. It very well may be separated as three phases: sensation coordination, semantic reconciliation, and confirmation. In the wake of investigating the advancement of radiologist's mammography viewing, a progression of optical guidelines dependent on the morphological qualities of breast masses are introduced and measured by scientific techniques.

Design is viewed as a viable exchange among base up sensation and top down acknowledgment strategies. This is another exploratory strategy for the programmed discovery of sores. The trials are done on Mammographic Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM) informational collections. Proposed structure has accomplished a superior presentation contrasted and different calculations.

Lee et al (2019) proposed a Fast Non Local Means (FNLM) de-noising calculation for prior breast malignancy recognition dependent on clinical mammography. In order to correlate with ordinary denoising techniques, the Wiener channel and Total Variation (TV) denoising calculation are utilized. worldly goals, Coefficient Of Variation (COV), and Contrast to Noise Ratio (CNR) are assessed for three Introduction situations:

(a) different cylinder voltages at fixed 40 mAs, (b) different cylinder flows at fixed 28 kVp, (c) auto introduction control mode. outcomes indicated that introduced FNLM denoising calculation can accomplish a comparative worldly goals to the Wiener channel and might proficiently lessen picture commotion by utilizing COV and CNR data in mammography.

Huang et al (2019) proposed a novel Biclustering mining and AdaBoost for breast tumor grouping. It is novel PC helped analysis conspire with distinguish the considerate and harmful breast tumors in ultrasound. In this structure, include procurement is performed by a client took an interest highlight scoring plan that depends on Breast Imaging Reporting and Data System (BI-RADS) dictionary and experience of specialists.

Biclustering mining is then utilized as a helpful apparatus to find the segment consistency designs on the preparation information. The examples regularly showing up in the tumors with a similar name can be viewed as a likely demonstrative guideline.

Along these lines, the demonstrative principles are used to build segment classifiers of the Adaboost calculation by means of a novel guidelines mix methodology which settle the Problem of Classification in Different Feature Spaces (PC-DFS). At last, the AdaBoost learning is performed to find compelling mixes and coordinate them into a solid classifier.

The proposed approach has been approved utilizing an enormous ultrasonic dataset of 1,062 breast tumor occurrences (counting 418 kind hearted cases and 644 harmful cases) and its exhibition was contrasted and a few traditional methodologies. The test results show that the proposed strategy yielded the best forecast presentation, demonstrating a decent potential in clinical applications.

Valvano et al (2019) introduced Convolutional Neural Networks (CNN) for distinguish and section breast Micro calcifications inside mammographic pictures. This model is made out of two back to back squares dependent on CNN: the indicator and the segmentator. Proposed CNN the identification and division of micro calcification bunches. Because of the fundamental investigation completed by the first CNN, the computational weight is extensively decreased and the absolute division process doesn't become tedious. In this work, utilized 283 mammograms to prepare and approve CNN model, getting most noteworthy grouping exactness on micro calcification location and a lower bogus positive rate.

Liu et al (2018) introduced Fully-Connected Layer First Convolutional Neural Network (FCLF-CNN) for breast cancer order. In which the completely associated layers are installed prior to initial convolutional layer. the completely associated layer as an encoder or an approximate to move crude examples into portrayals with greater area. So as to show signs of improvement execution, prepared four sorts of FCLF-CNNs and does outfit FCLF-CNN by coordinating them.

At that point used with WDBC and WBCD databases and got the outcomes through fivefold cross approval. outcomes demonstrated that the FCLF-CNN can accomplish a superior order execution than unadulterated multi-layer perceptrons and unadulterated CNN for both databases. troupe FCLF-CNN can accomplish an exactness of 99.28%, an affectability of 98.65%, particularity of 99.57% for WDBC, and a precision of 98.71%, an affectability of 97.60%, and an explicitness of 99.43% for WBCD. outcomes for the databases are serious contrasted with the consequences of other examination.

Deniz et al (2018) proposed an exchange learning, profound component extraction strategies are AlexNet, Vgg16 and characterization utilizing Support Vector Machines (SVM) for breast malignancy identification dependent on the histopathological pictures. For profound element extraction, two well known profound Convolutional Neural Network (CNN) structures in particular

AlexNet and Vgg16 frameworks are used for training.

The BreaKHis database is preferred for trial tasks because of tremendous test pictures. Three diverse exploratory tasks are thought of. initially, the component vectors starting at fc6 layers of both AlexNet and Vgg16 systems are extricated and afterward are connected. then subsequently, the fc7 layers of AlexNet and Vgg16 systems are utilized to include extortion and got highlight vectors are linked.

A SVM classifer is utilized for initial and next trial task of grouping the pictures into kind hearted and dangerous modules. next analyses, the pretrained AlexNet system is additionally adjusted with breast disease pictures. Test outcome indicated adjusted AlexNet outflanked and the main examinations outcomes are superior to the subsequent. acquired places of interest are then grouped by SVM. wide tests on a freely available histopathologic breast malignant growth database are done and the precision values are verified for implementation evaluation. evaluation results show that exchange learning created best result over intense element extortion and SVM arrangement.

Wang et al (2019) introduced Dual-mode Deep Transfer Learning (D2TL) for breast malignant growth location utilizing CEDM. Deep Learning (DL)- engaged analytic framework utilizing CEDM. The proposed framework is creative in a few viewpoints, including (1) a double mode profound engineering structure; (2) utilization of move figuring out how to encourage vigorous method assessment in little sized example; (3) improvement of perception procedures to assist decipher the method outcomes and encourage between and intra-cancer threat evaluation; (4) reduced human predisposition. At long last, apply D2TL to characterize considerate versus dangerous tumors utilizing the CEDM information gathered as of the Mayo Clinic in Arizona. D2TL beats contending systems and techniques.

Badr et al (2019) cross breed techniques to be specific SVM and Gray Wolf Optimizer (GWO) are proposed for recognizing the sort of the breast cancer as dangerous or generous. Improved SVM for adequately recognizing the breast malignant growth to abuse the greatest capability of SVM, an improved GWO technique was built up to choose the best boundaries of SVM for order. Analyze prescient execution of this cross breed method with SVM for breast malignancy information accessible from Wisconsin, UCI AI with an absolute 569 lines and 32 segments.

The test outcomes have shown that the created approach has accomplished progressively better order execution over SVM in precision, in this manner it very well may be securely presumed that the created clever framework can fill in as a promising elective choice emotionally supportive network for breast cancer identification.

Khan et al (2019) anticipated a novel profound CNN models: GoogLeNet, Visual Geometry Group Network (VGGNet) and Residual Networks (ResNet) for discovery and characterization of breast cancer in breast cytology pictures using the thought of move learning. In proposed method, places of interest from pictures are extorting using pre-prepared CNN models, to be specific, GoogLeNet, VGGNet and ResNet, that are taken care of into a entirely connected layer for order of threatening and favorable cells using usual pooling characterization. Additionally, likewise proposed the idea of information expansion to expand the size of an informational collection to improve the productivity of CNN structure.

To evaluate the exhibition of the proposed method, tests are performed on normal level informational indexes. It is viewed that the proposed structure outflank diverse profound learning models to the extent that precision in position and order of breast cancer in cytology pictures. At long last, the presentation of the proposed system is contrasted and diverse CNN models autonomously and furthermore contrasted and other ongoing techniques. It has been seen that introduced structure provides fantastic outcomes in regards to exactness without preparing without any preparation which improves arrangement productivity.

Wang et al (2019) proposed a CNN and Extreme Learning Machine (ELM) for breast cancer detection. CAD framework dependent on mammograms empowers premature breast malignant growth identification, analysis, and treatment. Breast CAD method reliant on contains mixture with CNN thoughtful things to see. first, introduced a mass recognition system reliant on CNN thoughtful things to see and unaided ELM clustering.

Second, form record abilities melding deep things to see, morphological things to see, surface things to see, and thickness things to see. then, an ELM classification method is formed using the collective record of capacities to order benevolent and hazardous breast masses. wide investigations

exhibit the precision and productivity of introduced mass recognition and breast cancer categorization grouping strategy.
Jamal et al (2018) anticipated Principal Component Analysis (PCA) and K-Means bunching for breast tumour forecast. Bolster Vector Machine (SVM) and Extreme Gradient Boosting strategy is thought about for grouping reason. Before the grouping, the quantity of information quality will be decreased from the crude information by separating highlights utilizing PCA. A clustering technique, in particular K-Means is likewise utilized for dimensionality decrease other than the PCA.
article introduces a correlation between four forms dependent on two dimensionality decrease techniques joined with two classification models that put in Wisconsin Breast Cancer Dataset. correlation is estimated by utilizing exactness, affectability and explicitness measurements assessed from disarray lattices. test outcomes have shown that K-Means strategy, that isn't typically utilized for dimensionality decrease can perform all around contrasted with the famous PCA.
Abdel-Zaher and Eldeib (2019) proposed a Deep Belief Network (DBN) utilizing Computer-Aided Diagnosis (CAD) plot for recognition of breast tumour. Computer aided design conspire for discovery of breast malignant growth has been created utilizing Deep Belief Network (DBN) unaided way followed by back spread regulated way. The development is back-propagation neural system with Liebenberg Marquardt learning capacity whereas loads are instated from Deep Belief Network Path (DBN-NN).
This procedure tried on Wisconsin Breast Cancer Dataset (WBCD). Classification method complex provides a precision of 99.68% demonstrating talented outcomes above beforehand distributed investigations. The proposed framework gives a successful grouping model to breast cancer. Furthermore, analyzed the engineering at a few train-test segments.
From the above writing survey papers, it very well may be discovered that heterogeneous breast densities make masses all the more testing to identify and group contrasted and calcifications. The conventional ML strategies present kept methodologies restricted to either specific thickness type or datasets.
DL techniques show promising upgrades in breast cancer conclusion, there are still issues of information shortage and computational cost, which have been defeated to a huge degree by applying information expansion and improved computational intensity of DL calculations. To tackle these current methodologies issues proposed the deep learning model of HIRResCNN is used for breast cancer identification with swarm strategy of HSO.

Proposed Methodology

This proposed work discussed about HIRResCNN with HSO technique based on Deep Convolutional Neural Network (DCNN) to classify breast tumor disease and it is termed as HIRResCNN-HSO. MIAS images are classified as malignant/ benign using this technique. There are four stages in this proposed HIRResCNN framework, namely, Pre-processing, reduction of dimensionality, segmentation and classification.
From images, noises are removed using two filtering algorithms called Median and mean filtering in pre-processing stage. Then canny edge detector is used for detecting edges. Gaussian filtering is used in canny edge detector to smoothen the images. Quality enhanced image is obtained using this edge detection and filtering process, which can be used in further image analysis process.

In the next dimensionality reduction stage, attributes are correlated using Principal Component Analysis (PCA) inclusive of related features. So, this huge dataset is minimized and only few variables are used for expressing it. In order to detect the breast cancer accurately, foreground and background subtraction is done in the third stage called segmentation stage. This results in a meaningful image, which is easy for analysing. Image representation is changed using this stage.
At last, for detecting and classifying breast cancer, a Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) is proposed, which integrates Harmony Search Optimization (HSO) to tune bias and weight parameters and classification accuracy is enhanced using HIRResCNN-HSO model. Figure 2 shows the HIRResCNN with HSO's proposed architecture.
Figure 1 shows Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN)'s flowchart designed for MIAS medical images. Image data is taken as an input of a first layer in a network as a fixed-length low-level feature vector. For disease diagnosis, it does not requires any feature engineering like various

traditional techniques and in hidden layer, automatic learning of extractor is performed. Through a HIRResCNN classifier's hidden layer, input vector is passed and in hidden layers, parameters are optimized using Harmony Search Optimization (HSO). In HIRResCNN networks, performance quality and structure are highly influenced by this. In the following supervised training phase, final classification layer called output layer is included for learning final prediction model of breast cancer disease.

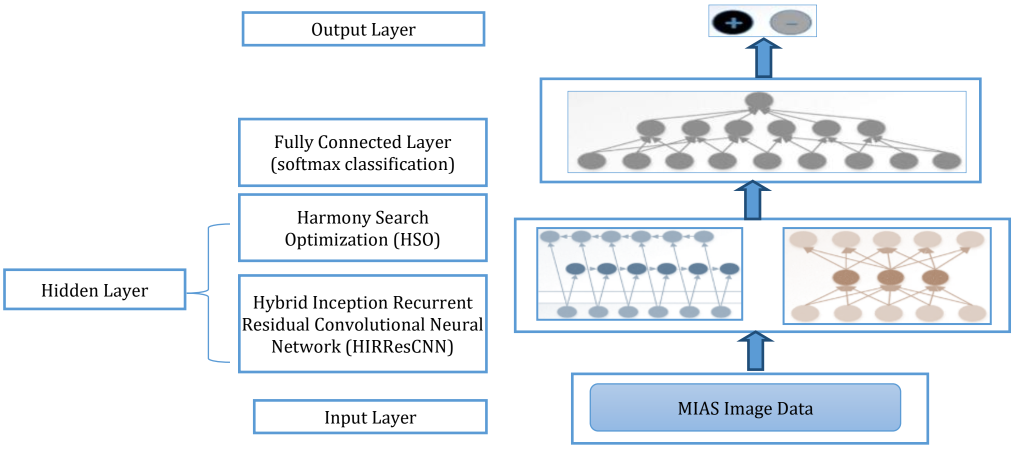

Figure 1. Flowchart of Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRresCNN)

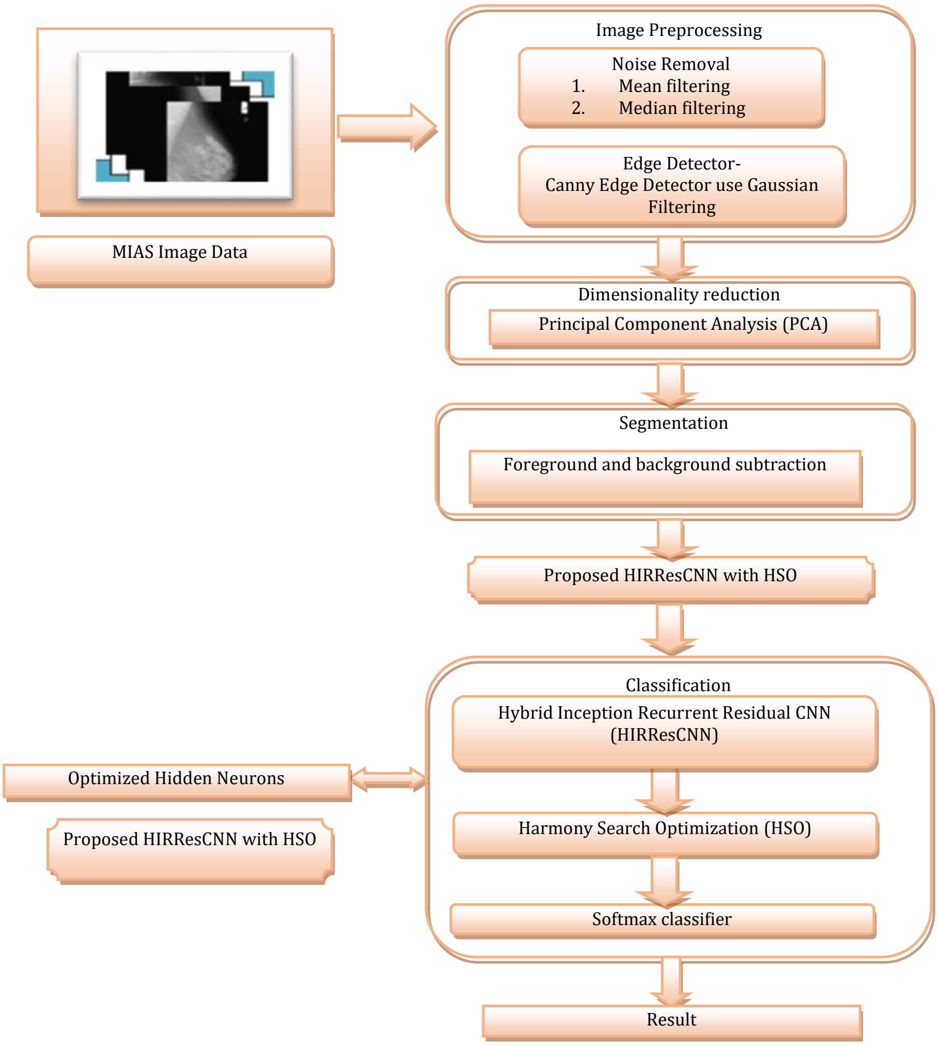

Figure 2. Proposed Architecture of HIRResCNN with HSO

A. Preprocessing

There are various unwanted and irrelevant parts in scanned medical image's actual format. An image composed of pixels or also termed as picture element is called as digital image. Every pixel is having a finite discrete quantity, which represents its gray level or intensity and it is an output of two-dimensional functions given with spatial coordinates x on x axis and y on y axis as input. It is classified as raster or vector type based on its fixed resolution. Spatial filter can be used for reducing Gaussian noise. But high frequencies are blocked by this filter. Because of that, in image smoothing, fine-scaled image details and edges may be blurred as a result of it. Gaussian smoothing, median filtering, mean or average filtering are the conventional spatial filtering methods used for removal of noise (Makandarand et al, 2016).

Image quality enhancement is a major objective of preprocessing techniques. In mammogram image background's surplus and unrelated parts are removed by this process and it makes the image ready for further processing. Mammograms is a medical image set, which complicates its interpretation. Noise removal using compute foreground markers, these are pixel connected blobs within every foreground objects. A "closing-by-reconstruction" and "opening-by-reconstruction" are the morphological techniques used for noise reduction (Li et al, 2019).

Mean Filtering

A simple sliding window spatial filter where, average value of all the window pixel is used for replacing center pixel value is termed as mean filter. Square shape of kernel or window is used in general case, but it may take any kind of window. Within a local image region, all pixels average is used for defining arithmetic mean filter (Charate et al, 2017).

In a rectangular sub image window with $m \times n$ size, assume coordinate set is represented as S_{xy}, which is centered as point (x, y). In an area defined by S_{xy}, corrupted image $g(x, y)$'s average value is computed using arithmetic mean filtering process. At any point (x, y), restored image value will equals mathematical mean of pixels in the area shown by S.

$$\hat{f}(x,y) = \frac{1}{mn} \sum_{(s,t)\in S_{xy}} g(s,t)$$

Median Filtering

A nonlinear signal processing method is median filtering which is statistic based. Median value of neighborhood is used for replacing digital sequence or images noisy value. Based on gray levels, mask pixels are ranked. Noisy value is replaced by storing group's median value. $g(x, y)\ \mathrm{med}\{ f(x - i, y - j), i, j \in W\}$ is the output of median filtering, where, novel image is symbolized as $f(x, y)$, output image is symbolized as $g(x, y)$, two-dimensional mask is represented as W, size of the mask is $n \times n$(with n as an odd value) like $3 \times 3, 5 \times 5$. Cross, circular, square, linear shaped mask can be used (Sha et al, 2020). For image having random noise, mathematical analysis of median filtering is a complex one because of its non-linear nature. Under normal distribution, image having zero mean, median filters noise variance can be approximated as,

$$\sigma_{med}^2 = \frac{1}{4nf^2(n)} \approx \frac{\sigma_i^2}{n + \frac{\pi}{2} - 1} \cdot \frac{\pi}{2}$$

Where, input noise power is represented as σ_i^2, median filtering mask size is represented as n, noise density function is represented as $f(\bar{n})$. 8
Average filtering noise variance is represented as,

$$\sigma_0^2 = \frac{1}{n}\sigma_i^2$$

In reduction of random noise, better performance is exhibited by median filtering when compared with average filtering performance. But less effective performance is exhibited on impulse noise of narrow pulse with less than n/2 pulse width. If average filtering algorithm is combined with median filtering algorithm, then its performance can be enhanced and based on noise density, mask size is varied adaptively. Figure 3 shows MIAS images benign, malignant noise removal.

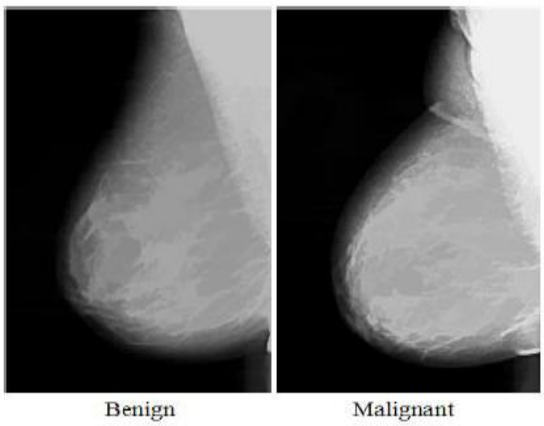

Figure 3. Noise Removal for Mias Images

Canny Edge Detection

In this work edge detection is done by using canny edge recognition method. Canny edge detector utilizes multiple stage procedure to identify a broad variety of edges in images. The advantage of canny edge detector is listed below:

1. existence of Gaussian filter permits eliminating of a few noise in an image.
2. signal is improved regarding the noise proportion by non-maxima containment system that outcomes in one pixel wide ridges as outcome.
3. identifies edges in a noisy condition by utilizing thresholding method.
4. efficiency is modified through factors.
5. provides superior localization, answer and is resistant to a noisy surroundings.

Various mathematical techniques are included in this edge detection for points identification in a digital image, where there exist a sharp brightness changes in a formal way. Points having sharp brightness changes are organized as a curved line segments set called edges. In feature extraction and detection, especially in computer vision, machine vision and image processing, edge detection plays a major role. From various vision objects, useful structural information can be derived using Canny edge detection, which decrease total data that needs to be practiced dramatically. On diverse vision systems, edge detection application's requirements are relatively similar as found by Canny (Soulami et al, 2019).

Process of Canny Edge Detection Algorithm

There are five steps in canny edge detection algorithm process.

1. For noise removal, image is smoothened by applying Gaussian filter.
2. Image's intensity gradient is computed.
3. To get better edge detection results, non-maximum suppression is applied.
4. Potential edges are computed using double threshold
5. Edge detection is finalized by suppressing all other weaken edges and not having any connection with strong edges.

1. Gaussian Filter

In Graphics software, it is a most commonly used effect for reducing image noise. Image blurring filter type corresponds to Gaussian blur, where transformation is computed using Gaussian function, which is applied to every image pixel. In statistics, normal distribution is expressed by it T. In an image, noise will easily effect the results of all edge detection. So, in order to prevent, false detection, noise has to be filtered out. Convolve the Gaussian filter with image for smoothening it (Kwon et al, 2016). Image is slightly smoothened using this step and its effects on edge detector is reduced. By convolving Gaussian filter G with input image $I(i,j)$, image with noise is smoothened. It is mathematically expressed as,

$$F(i,j) = G * I(i,j)$$

Finding Gradients

Various directions are pointed out by an edge in image. So, in blurred image, diagonal, vertical and horizontal edges are detected using four filters in Canny algorithm. In vertical direction (G_j), and horizontal direction (G_i), first derivative value is computed using an edge discovery operators like Sobel, Prewitt and Roberts. Using this direction and edge gradient are computed as,

$$G = \sqrt{G_i^2 + G_j^2}$$

Where, hypot function is used for computing G and arctangent function is represented as arctan and it has two arguments. Round the edge detection angle to $0°, 45°, 90°$ and $135°$ for representing two diagonals, horizontal end vertical directions. Gradient direction is represented as $\theta = \arctan\left(\frac{G_j}{G_i}\right)$, where, i-th direction gradient is represented as G_i and j-th direction gradient is represented as G_j.

2. Non-maximum Suppression

An edge thinning method corresponds to a non-maximum suppression method. There will be blurred nature edge which is extracted from gradient value, after gradient calculation application. Except, local maxima values, all other gradient values are suppressed using a non-maximum suppression. Intensity value's sharpest change locations are indicated using this method. For every gradient image pixel, algorithm is given by,

1. In negative and positive directions, pixels edge strength is compared with current pixels edge strength.

2. Preserve current pixels edge strength value, if it is high when compared with other pixels of mask which are in similar direction. Pixel which is pointed towards y direction is compared with pixels in its vertical axis. E.g. Above and below pixels. Else, suppress the edge strength.

In gradient image, all local maxima's are preserved using non maximum suppression and in thin edges, everything else is deleted. For a pixel $m(i, j)$:

- Rounded gradient angle is θ and it lies between 0 to 45°. In north west–south east direction, edges will be present. If pixels magnitudes in south west and north east directions are smaller than points gradient magnitude, then that point is considered on edge. In east gradient direction, pixels gradient in west $\mathrm{W(i, j)}$and east $\mathrm{E(i, j)}$directions are compared.
- Pixel $\mathrm{m(i, j)}$ is marked as edge pixel, if $\mathrm{E(i, j)}$ and $\mathrm{W(i, j)}$ are less than pixel $\mathrm{m(i, j)}$'s edge strength and gradient value is preserved, else that pixel is removed or suppressed.

3. Double Threshold

In an image, real edges highly accurate representation is provided by remaining pixels, after non-maximum suppression application. But, there exist few edge pixels, which are color variation and noise produced pixels. Edge pixels having weak gradient values must be filtered out while preserving high gradient value edge pixels for producing better responses. Minimum and maximum threshold values are chosen for performing the same. An edge pixel is noticeable as a strong edge pixel, if high threshold value is smaller than its gradient value.

An edge pixel is decided as a weak edge pixel, if a high threshold value is greater than its gradient value and low threshold value is lower than it. Still there exist a local maxima in non-maxima suppression output which is a noise created one. In order to avoid streaking problem, two thresholds T_{low} and T_{high} are selected instead of one single threshold.

4. Edge Tracking by Hysteresis

Pixels extracted from noise/ color variations and from true edge are termed as weak edge pixels. Edges produced due to latter reasons are removed for producing better results. With unconnected noise response, strong edge pixel can be connected with a weak edge pixel produced from true edge. Blob analysis is applied for tracking edge connection with respect to weak edge pixel and its neighbourhoods pixels in 8-connected manner. Following conditions are used for detecting a pixel $m(i, j)$ with gradient magnitude G as edge.

B. Edge is discarded, If G $< T_{low}$
C. Edge is maintained, If G $> T_{high}$
D. Edge is maintained, If T_{low} < G < andT_{high} and in a 3×3 region, any of its neighbors around is having a gradient magnitudes greater than T_{high}.

In the above sections Mean and median filters are used to remove the noises. Gaussian is filter is an part of canny edge detector which will smoothen the image when image edges are detected. After edge detection, dimensionality reduction is performed which is explained below:

E. Dimensionality Reduction

Number of random variables are reduced using the process called dimensionality reduction by computing principal variable's set. High storage space is required by high-resolution image rather than data transmission. Attributes are correlated by applying this and it gives better result. Related features are included in it for reducing huge dataset and only few variables are used for expressing it (Taghanaki et al, 2017).

1. Principal Component Analysis (PCA)

Most popular technique utilized for decreasing linear dimension is PCA. For a few cases, it is used separately and in some cases, it is used with other dimension reduction techniques. PCA is projected based technique, where data is projected as a set of orthogonal axes. A quantitatively rigorous technique used for data set simplification is Principal Component Analysis (PCA).

New variable set are generated using this technique and are termed as principal component. Original variable's linear combination are used for making every principal component. They are orthogonal to each other, there wont be any redundant information. For data space, an orthogonal basis is formed using principal component (Xie et al., 2016).

A simplified datasets are formed by linear transformation is PCA, where original dataset's

characteristics are retained. Data analysis are done using PCA. PCA is a dimension reduction method, where related features are included. So, huge dataset is reduced and few variables are used for expressing it.
Better results are produced if correlated attributes are applied with PCA. Breast cancer dataset's training and test attributes are applied with PCA. In dataset, patterns are detected using PCA and between every individual attributes, differences and similarity are computed as well (Wahdan et al., 2016).
Following steps are included in the process of PCA:

Step 1: Data is received

Step 2: Mean is subtracted

Step 3: Covariance matrix is computed

Step 4: Covariance matrix's Eigen values and eigenvectors are computed

Step 5: Components are selected and a feature vector is formed

Step 6: The 'new data' set is derived, and later getting old data back.

Following is used for stating low-dimensional feature representation problem: Assume $n \times N$ data matrix is represented as $X = (x_1, x_2, \ldots, x_i, \ldots, x_n)$, where, a feature vector with n dimension is given by x_i. In breast cancer input, total attributes count are represented as n and in training set, input cases count are represented as N.
Linear transformation of original input vector into a projection feature vector (Salleh et al., 2017) corresponds to PCA, i.e.

$$Y = W^T X$$

Where, $\mathrm{m} \times \mathrm{N}$ feature vector matrix is represented as Y, feature vector dimension is represented as m and an $\mathrm{n} \times \mathrm{m}$ transformation matrix is represented as W. Columns of this matrix are m largest Eigen value's eigenvectors and are computed as,

$$\lambda_{ei} = S_{ei}$$

Where, eigenvectors are represented as e_i, Eigen values matrix are represented as λ. Total scatter matrix is represented as S and all samples mean are computed as,

$$S = \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T$$

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Transformed feature vectors $\{y_1, y_2, \ldots, y_N\}$ scatter is represented as $\mathrm{W^T SW}$, after the application of linear transformation W^T. The projected sample's total scatter matrix's determinant is maximized by selection projection W_{opt}, i.e.,

$$W_{opt} = \arg\max_{W} |W^T S W| = (w_1 w_2 \ldots w_m)$$

Where, S's set of n −dimensional eigenvectors are represented as $\{\mathrm{w_i} | \mathrm{i} = 1, \ldots .. 2, \mathrm{m}\}$, which are corresponding to m highest Eigen values. In can also be stated as, attribute vector in an m–dimensional subspace can be formed by reducing the input vector in an n −dimensional space. Input faces vector n's dimension greater than the reduced feature vector m's dimension.

F. Segmentation

Partitioning process of digital image into various segments like pixel set called image objects is termed as segmentation in computer vision and digital image processing. Image representation is changed and/or simplified into something, which makes analyses process as a simple one. Segment set which covers entire image are produced using image segmentation or it may results in extraction of contours. In a region, there will be a similarity between every pixels based on some computed property or characteristics like texture, intensity, color. With respect to same characteristics, there will be difference in adjacent regions.

Foreground and Background Subtraction

In general, from background, image is segmented and for image analysis, only front region image pixels are used. Various techniques like gradient analysis, local gray-value range's global histogram analysis are used for mammogram segmentation into background and breast regions.
Without analysis background are segmented from foreground using these techniques, if image analysis process requires background pixels. Necessary information are contained in the mammogram's background pixel's and in body, cancer cells presence are detected using this information (Salleh et al., 2017). Regions having variance value less than threshold are assigned with a zero gray value for forming final segmented image.
The gray scale variance for a block of size R*R is defined

$$Z(k) = \frac{1}{R^2} \sum_{i=0}^{R-1} \sum_{j=0}^{R-1} ((x(i,j) - y(k))$$

A block is assigned as a background, if global threshold is greater than variance, else it will be assigned as a foreground. Where, block 's variance is represented as $Z(k)$, at pixel (i, j), gray level value is represented as $x(i, j)$ and for a block, mean gray level value is given by $y(k)$.

1. Binarization

Image thresholding is used for achieving Binarization, where, pixels having less gray level than specified threshold are grouped into foreground or background and balance pixels are grouped into another class. Selection of threshold (T) for extracting an object or various objects having similar value is a basic concept of thresholding. Every block's mean gray value is computed as threshold T. Below mentioned expression is used for the same. Performed binarization by making a comparison of mean value with every pixel in a block.

$$g(x, y) = \begin{cases} 1 \; if \; f(x, y) > T \\ 0 \; if \; f(x, y) <= T \end{cases}$$

Where, gray scale pixel values are represented as $f(x, y)$ and binarized image is represented as $g(x, y)$. Peaks of histograms must be split by deep valleys for selecting good threshold. Some pixel points may get damaged after binarization and image may have zero and one gray values. So, for enhancing quality of an image for extracting more information, thinning process is required by binarized image.

2. Thinning

From binary images, selected foreground pixels are removed using a morphological operation called thinning algorithm. Original region's topology called extent and connectivity are preserved and most of the original foreground pixels are thrown away. Minutiae extraction speed increases using this thinning process. A matrix is used for defining binary digitized picture, where every pixel $g(x, y)$ is either 1 or 0. Based on small set of neighbours of point (i, j) values, in a point by point manner, iterations are applied.

G. Classification

Process of predicting specified data points class is termed as classification. Classes are also termed as categories, labels or targets. categorization is a type of supervised learning, wherein, input data is specified with targets. To identify and recognize breast tumour Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) is formed by integrating Harmony Search Optimization (HSO) to tune HIRResCNN model parameters and HIRResCNN algorithm is used in softmax classifier.

1. Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN)

With the availability of sufficient labelled data, tremendous success is shown using Deep Learning (DL) techniques. During last few years, in various medical imaging and computer vision modalities, various advanced DL techniques are proposed, which shows better performance than modern techniques (Ball et al., 2017).

RCNN architecture, residual networks and inception based class of enhanced hybrid DCNN architecture is Inception Recurrent Residual Convolutional Neural Network (IRResCNN) (Alom et al., 2018). With same or lesser network parameters, better recognition performance are provided using this model when compared with alternative deep learning techniques like residual network, RCNN, inception. It is a major advantage of this technique. Based on Inception-v4 model, utilized the inception-residual units in this system. Equivalent inception-residual networks are compared with IRResCNN and it shows better performances. In the Recurrent Residual Convolutional Neural Network. Inception V4 model is integrated. So it became Hybrid Inception recurrent residual convolutional neural network. Here performance of classifier is enhanced by choosing the optimal values of parameters which is done by using Harmony Search Optimization.

There are stacks in IRResCNN model, which includes both transition units and Inception Recurrent Residual Units (IRResU) (Szegedy et al., 2017). Figure 4 shows the entire model. There are various convolution layers, transition blocks, IRResUs and softmax at output layer. Figure 5 shows the IRResU's pictorial view. Steps used to train the system are shown in upper part of the figure 4 and testing phase is displayed in lower part of this figure. Various performance evaluation metrics are used for evaluating the results of this technique.

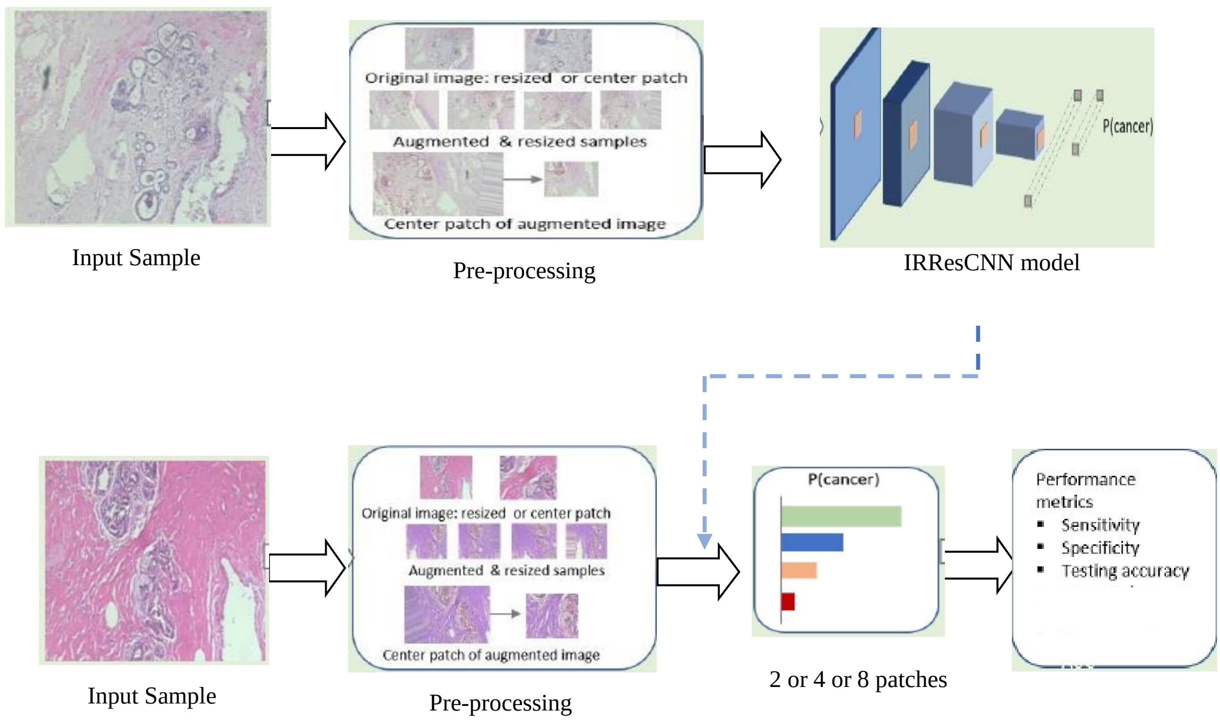

Figure 4. Illustration of Breast Tumour Identification through IRResCNN FRAMEWORK

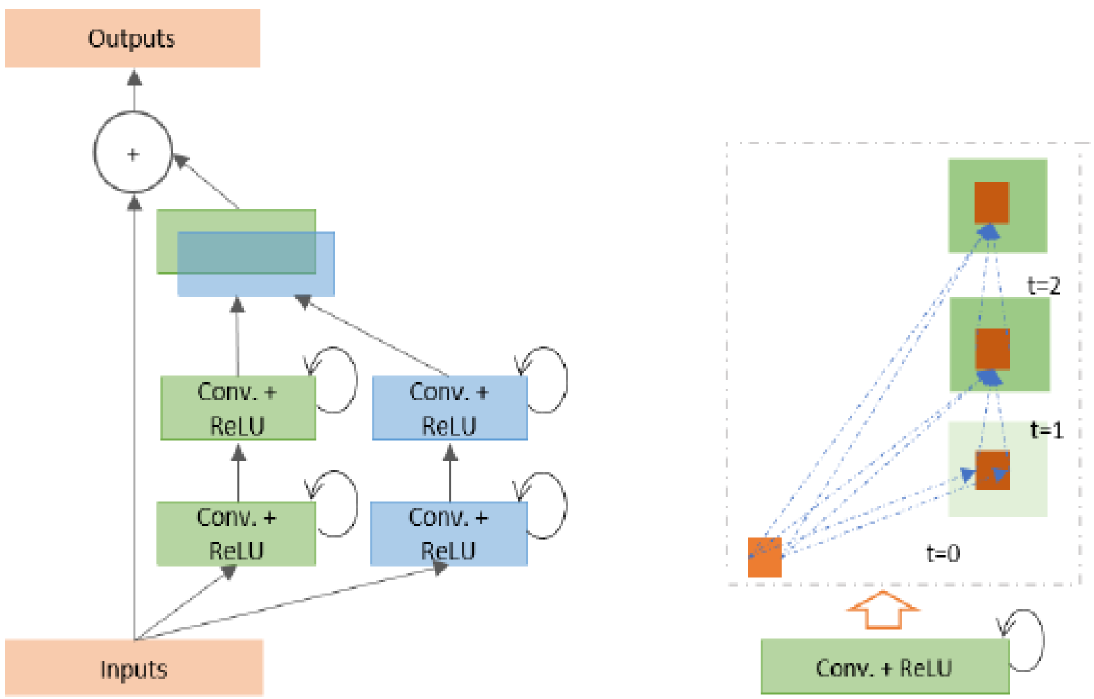

Figure 5. Illustration of Inception Recurrent Residual Unit (IRResU)

Figure 5 has recurrent convolutional layers and inception unit which are concatenated and residual units. Input features summation with inception element output is viewed prior to output block.

IRResU is a highly important unit of IRResCNN architecture, where residual layer, inception units, Recurrent Convolutional Layers (RCLs) are included (Wu et al., 2018). Input layers are fed with inputs and through inception units, they are passed and RCLs are applied. Then IRResU inputs are added with inception units outputs.

In inception unit, based on various kernel size, performed the recurrent convolution operations. Previous time step outputs are added with present time step outputs because of recurrent structure within convolution layer. Then, for next time step, current time step outputs are given as inputs. Based on assumed time steps, same operations are performed.

For instance, in IRResU, single feed forward convolution with 2 included RCLs are represented as $t = 2\ (0{\sim}2)$. Figure 5 shows the RCLs operation based on various time steps $\left(t = 2\ (0{\sim}2)\right)$and $(t = 3\ (0{\sim}3))$. Based on time steps, feature map accumulation is performed using IRResU. This ensures better representation of features and with similar amount of network parameters, superior performance can be achieved using this system.

Based on discrete time steps, performed the RCL operations and based on IRResCNN, they are expressed. Assume, $x_l$input sample in IRResCNN block's l^{th} layer and unit (i,j) from input example in k^{th} attribute map in RCL. In addition, at time step t, network output is assumed as $\mathrm{O}^{\mathrm{l}}_{\mathrm{ijk}}(\mathrm{t})$. With this information, output is expressed as,

$$O^l_{ijk} = \left(w^f_k\right)^T * x^{f(i,j)}_l(t) + (w^r_k)^T * x^{r(i,j)}_l(t-1) + b_k$$

Where, input for standard convolution layers is represented $x^{f(i,j)}_l(t)$, its weight value is represented as $w^f_k$and input for l^{th}RCL is represented as $x^{r(i,j)}_l(t-1)$, its weight value is represented as w^r_k and bias is represented as b_k. HSO optimization techniques are used to tune these weight and bias parameters and thus enhances HIRResCNN model's accuracy in classification.

$$\mathrm{y} = \mathrm{f}\left(\mathrm{O}^{\mathrm{l}}_{\mathrm{ijk}}(\mathrm{t})\right) = \max\left(0, \mathrm{O}^{\mathrm{l}}_{\mathrm{ijk}}(\mathrm{t})\right)$$

Where, activation function of standard Rectified Linear Unit (ReLU) is represented as f. With Exponential Linear Unit (ELU) activation function, this model's performance is evaluated as follows: for various sized kernels, inception units output y is given by $y_{1x1}(x), y_{3x3}(x)$, and average pooling layer output is represented as $y^p_{1x1}(x)$. The $\mathcal{F}(x_l, w_l)$ represents Inception Recurrent Convolutional Neural Network (IRCNN) unit's final output and is expressed as,

$$\mathcal{F}(x_l, w_l) = y_{1x1}(x) \odot y(x) \odot y^p_{1x1}(x)$$

Where, based on feature map axis or channel, concatenation operation is represented as $\odot$. Then, IRResCNN block's input is added with IRCNN unit's output. The IRResCNN block's residual operation is given by,

$$\mathrm{x}_{\mathrm{l+1}} = \mathrm{x}_{\mathrm{l}} + \mathcal{F}(\mathrm{x}_{\mathrm{l}}, \mathrm{w}_{\mathrm{l}})$$

Where, input of immediate next transition block is referred as $\mathrm{x}_{\mathrm{l+1}}$, IRResCNN block's input samples are represented as x_{l}, l^{th}IRResCNN block's kernel weights are represented as w_{l} and output of IRCNN unit's l^{th}layer is represented as $\mathcal{F}(\mathrm{x}_{\mathrm{l}}, \mathrm{w}_{\mathrm{l}})$. HSO optimization technique is used to tune parameters $\mathrm{w}_{\mathrm{l}}$and $bias$, which enhances HIRResCNN model's accuracy of classification.

Nonetheless, the quantity of highlight maps and the components of the element maps for the remaining elements are equivalent to in the IRResCNN unit appeared in Figure 5. Cluster standardization is applied to the yields of the IRResU. In the end, the yields of this IRResU are taken care of to the contributions of the prompt next progress unit (Liang et al., 2015). In the change unit, various activities containing convolution, pooling, and dropout are done relying on the situation of the progress part in the model. initiation element are remembered for the progress unit. 14

Down-inspecting activities are acted in the progress elements, where max-pooling tasks is done with a 3×3 fix and a 2×2 step. non-covering max-pooling activity negatively affects model regularization, in this manner we utilized covered max-pooling for standardizing the system that is significant while preparing a profound system engineering. The late utilization of a pooling layer assists with expanding the non-linearity of the highlights in the system, as this outcomes in higher dimensional component maps being gone during the convolution layers in the system.

This additionally assists with keeping the quantity of system boundaries at the very least. The advantage of including a 1×1 channel is that it assists with expanding the non-linearity of the choice capacity without having any effect on the convolution layer. as the information size and yield highlights doesn't change in the IRResCNN units. At long last, utilized a softmax, or standardized exponential capacity layer toward the finish of the engineering. To improve arrangement tune boundaries weight and inclination of HIRResCNN model use Harmony Search Optimization (HSO) calculation.

2. *Harmony Search Optimization (HSO)*

An improvisation process used for musicians is a Harmony Search Optimization (HSO). A kind of beautiful sound combination is a music harmony, which comes from aesthetic view. Music performance is computed for finding optimal state, which computed using aesthetic evaluation. After applying optimization problem, decision variable's objective function is represented using musicians and optimum state is computed using a heuristic procedure known as Harmony Search (HS), which is done via objective function value (Maleki et al., 2014).

Enhancing procedure of music players in searching a perfect harmony state is mimicked in developing a meta-heuristic procedure named Harmony Search (HS). Limited mathematical requirements are imposed by HS algorithm, when compared with previous meta-heuristic optimization algorithms and initial value settings are not affected by this. Harmony memory, notes harmonics and musicians are the key concepts in HS algorithm.

Functions decision variable's corresponds to musician. Notes played by all musicians are there in harmony and termed as solution vector, which has one value per variable. Harmonies played by musicians are stored in harmony's memory and it is used as a storage place of solution vectors (Elyasigomari et al., 2017).

The note that a performer plays is the estimation of each choice variable. Concordance memory comprises of the amicability which is played by the artist, or is an extra room for the arrangement vector. Specifically, the amicability memory is a two-dimensional lattice, in which the column vector portrayal concordance (the arrangement vector), and the quantity of lines speaks to the size of the congruity memory. Every section stores the agreement that is played by various artists, that is, the congruity edge of every artist.

Real performance function f(x) and generating solve vector, which is acting as a memory are used for forming Harmony memory in a random manner.

$$\min F(X) = |f_1(x), f_2(x), f_3(x)|$$

$$Minimize\ f(x) x_i \in X_i, i = 1,2, \ldots, N$$

$$x_i' = \begin{cases} x_i' \in \left(x_i^1, x_i^2, \ldots, x_i^{HMS}\right), if\ rand < HMCR \\ x_i' \in X_i, otherwise; i = 1,2, \ldots, N \end{cases}$$

The HIRResCNN model parameters are given by $w_k^f, w_k^r$and b_k, which are used for enhancing accuracy of classification.

$$x_i' = \begin{cases} \left(w_k^f\right)^T * x_l^{f(i,j)}(t) + b_k \in \left(x_i^1, x_i^2, \ldots, x_i^{HMS}\right), if\ rand < HMCR \\ x_i' \in X_i, otherwise; i, j = 1,2, \ldots, N \end{cases}$$

$$x_i' = \begin{cases} (w_k^r)^T * x_l^{r(i,j)}(t-1) + b_k \in \left(x_i^1, x_i^2, \ldots, x_i^{HMS}\right), if\ rand < HMCR \\ x_i' \in X_i, otherwise; i, j = 1,2, \ldots, N \end{cases}$$

For the i^{th}decision variable, random value is represented as $rand$ which lies between 0 to 1.

Number of single-objective sub-problems are formed by decomposing fitness function and objective values are formed by solving this in a collaborative manner.

Fitness function is expressed as

$$fitness(X) = \left|\frac{1}{T}\right| \sum_{i=1}^{|T|} z_i(X)$$

$$subject\ to\ z_i(X) = \min F(X)$$

A heuristic HSO solution is applied in three stages. Any famous music piece, which is a pitch series in harmony, is played from her or his memory. A known piece which is similar to previous one is played by adjusting the pitch slightly and random or new notes are composed. With selected solution, applied the random harmony heuristic for forming new solutions set.

Compare the new random harmony solution and analyse them based on its properties like configuration for deciding it to include in existing solutions set or for terminating from it for including new solutions obtained in next iterations. One or more solutions are considered and new solutions set are formed by modifying or combing them. Various search based operations are used for forming this solutions.

There are five steps in HSO:

Step 1: instead of $f(x)$ in below mentioned expression.

$$Minimize\ f(x) x_i \in X_i, i = 1,2, \ldots, N$$

Computation of parameters value: Number of Improvisation (NI), Pitch Adjustment Rate (PAR), Harmony Memory Consideration Rate (HMCR), Solution vector count in harmony memory, Harmony Memory Size (HMS).

Step 2: Harmony Memory is shaped and created according to below mentioned matrix.

$$HM = \left[\begin{matrix} x_1^1 & x_2^1 & \cdots & x_n^1 \\ x_1^2 & x_2^2 & & x_n^2 \\ \vdots & & \ddots & \vdots \\ x_1^{HMS} & & \cdots & x_n^{HMS} \end{matrix}\middle|\begin{matrix} f(x^1) \\ f(x^2) \\ \vdots \\ f(x^{HMS}) \end{matrix}\right]$$

Real performance function f(x) and generating solve vector, which is acting as a memory are used for forming Harmony memory in a random manner.

Step 3: In harmony search algorithm, it is a highly important step, where all the changes in present harmonics are done here. Using below mentioned expression, amount of inner harmony memory used are clarified using HMCR and new random harmony creation probability is shown using 1-HMCR.

$$\acute{x_l} \leftarrow \begin{cases} x_l \in \{x_l^1, x_l^2, \ldots, \acute{x_l^{HMS}}\}\ w.p\ HMCR \\ x_1 \acute{\in} X_l\ w.p(1 - HMCR) \end{cases}, HMRC \in [0,1]$$

It is not recommended to use the value of 1 for HMCR. For offering best solution, in harmony memory, saved values are used for enhancing the solution. From inside the memory, if one value is selected, according to PAR probability, this value can be changed as,

$$\acute{x_l} \leftarrow \begin{cases} Yes\ w.p\ PAR \\ No\ w.p\ (1 - PAR) \end{cases}, PAR \in [0,1]$$

If pitch modification choice for is $\acute{x_l}$ YES, $\acute{x_l}$is changed as,

$$\acute{x_l} \leftarrow x_l \pm \acute{rand} \times BW$$

Where, arbitrary distance bandwidth is represented as BW and a random number is represented as x and its value lies between 0 to 1.

Step 4: In this step, in memory, if worst number is worst than New Harmony, old one is replaced with New Harmony. Then, worst harmony will be eliminated. Also, at the top, according to best member, harmony memory are sorted. In this way, memory can be updated.

Step 5: Algorithm is terminated in this stage. Steps 3 and 4 will be repetitive, if closing is not fulfilled. However, termination condition can be adjusted to a certain optimum value, so, till the end, algorithm steps are repeated.

3. Softmax Classifier

For i^{th} class, with K distinct linear functions, weight vector W and input sample x, softmax operation is expressed as,

$$P(y = i|x) = \frac{e^{x^T w_i}}{\sum_{k=1}^{K} e^{x^T w_k}}$$

In convolution blocks, various convolutional layers count are used for evaluating this IRResCNN model and according to time step t, layers count are computed. A model with two convolutional layers are used in breast cancer recognition in the beginning state and it has four IRCCN blocks, transition block, fully connected layer, hidden layer and at the end, it has softmax layer.

Data are classified using this last hidden state. Following shows the expression used for classification,

$$y = softmax\left(w_k^f h_T, +b_k\right)$$
$$y = softmax(w_k^r h_T + b_k)$$

Where, predicted gait type is represented as y, output weight is represented as w_k^f, input weight is represented as w_k^r and output bias is represented as b_k. Image pattern are recognized using softmax classifier. Trained feature vector is obtained as a output from classifier. A fixed-length trained feature vector $\hat{x}$ is formed by transforming breast cancer image x_i's every patch. Classification results are obtained by feeding softmax classifier with this hidden layer. A class of multiclass classifier is softmax classifier and logistic regression is used in this classifier for data classification.

Every class probability to which the data is classified is estimated using this. Hence, this will leads to the sum of probability as one. 16
Normalization is done using softmax function and class probabilities are computed using exponentiation process. After training all network layers, the next training stage is termed as fine tuning. In classification process, final stage is a fine tuning stage, which is used for enhancing model performance.

Results and Discussion

Mini Mammographic (MIAS) breast cancer data set is used in this work. It is available on online and it has important risk factors. In labs, breast cancers are diagnosed using these factors in general and produces reliable results. Different cases are there in that dataset and they are classified as two classes of tumors called malignant and benign. With nine features in that dataset, tumor class is computed. Mammographic Image Analysis Society (MIAS) created this database and has 322 322 digitized breast images of resolution 1024x1024.

Available techniques such as RNN and DNN are compared with proposed HIRResCNN framework (Botchkarev et al., 2018). Promising outcomes are produced using this technique and nodules can be classified as malignant and benign. Metrics like F-measure, recall, precision, accuracy are used for evaluating proposed HIRResCNN-HSO's performance. The simulation outcomes are illustrated in the following figure 6.

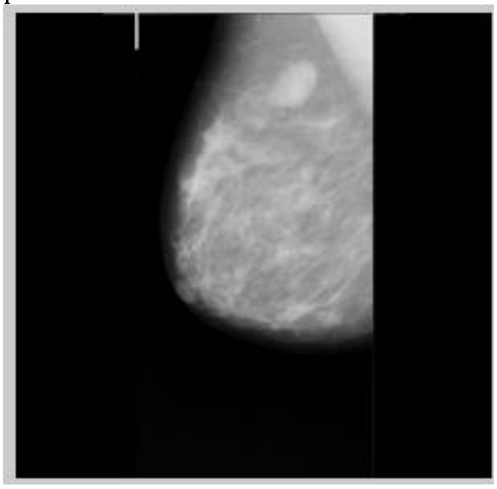

Figure 6. Preprocessed image

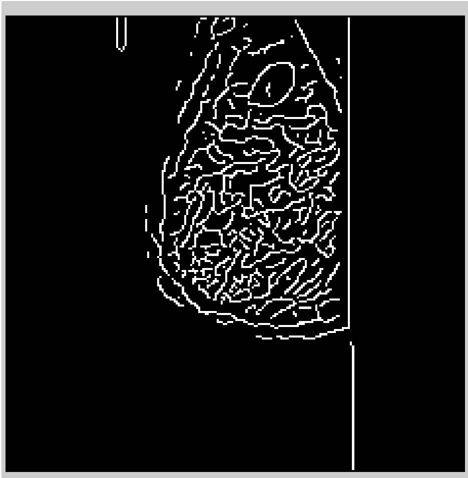

Figure 7. Canny edge detected image

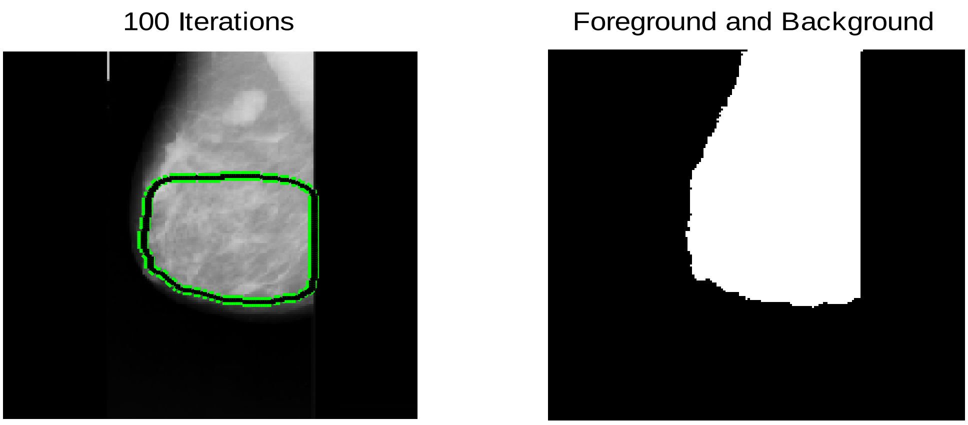

Figure 8. Segmented outcome

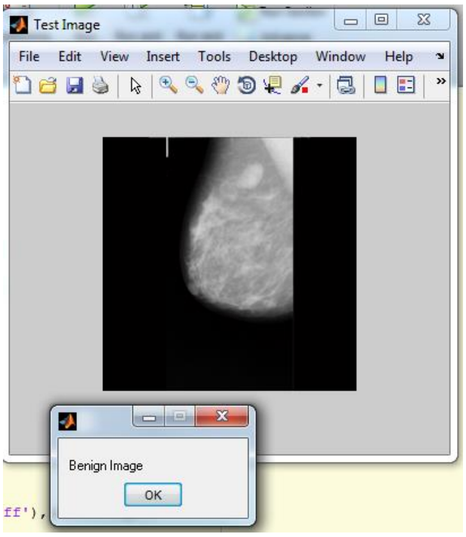

Figure 9. Classified outcome

The metric used for two class classification task is dependent on confusion matrix tabulated in table 1. On a test data set, classification model performance is described using table called confusion matrix where true values are well-known. permits performance visualization of procedure (Ohsaki et al., 2017).

In foretelling analytics, a confusion matrix (Table 1), involves a table with two rows and columns that corresponds to total false positive (fp), false negative (fn), true positive (tp), and true negative (tn). Classification performance is examined using a most popular and effective measure called accuracy, which applied in this study for prediction of recurrence. In addition, for analysing classifier's incorrect and correct decisions, measures like f-measure, recall and precision are used.

Truly detected breast cancer count is represented using a parameter True positive (tp), non-mitosis count that are misclassified as breast cancer is represented using a parameter false positive (fp), objects count that are undetected incorrectly is represented using a parameter false negative (fn) and truly detected count of non-breast cancer is represented using a parameter true negative (tn).

Table 1. Confusion Matrix

		Predicted	
		Positive	**Negative**
Actual	**Positive**	Tp	Fp
	Negative	Fn	Tn

- True Positive is correctly predicting a tag
- True Negative is correctly predicting the other tag
- False Positive is falsely predicting a tag
- False Negative is omitted and incoming tag

Precision (P) is called proportion of expected positive cases that were right expressed as

Precision (P)= tp/ (tp+fp)

True Positive Rate (TPR) or Recall is proportion of positive cases whicht are got correctly, expressed as

TPR= tp/ (tp+fn)

Precision and recall's harmonic mean corresponds to F1 score or balanced F-score or F-measure. This score is scaled to 1 by multiplying with a constant value of 2. It is given by,

F-score= 2tp/(2tp+fp+fn)

Accuracy is a proportion of total right predictions. It is given by

Accuracy= (tp+tn)/ (tp+tn+fp+fn) (13)

Overall results of the methods with performance evaluation metrics vs classification techniques are discussed in table 2.

Table 2. Varied Classifiers Performance

Techniques	**Measures**					
	Precision (%)	**Recall (%)**	**F-measure (%)**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
DNN	83.4	83.6	80.8	89.2	83.7	84.8
RNN	85.2	84.8	86.6	90	85.8	86.1
HIRResCNN	89.6	89.4	89	92.6	89.9	91.4

Precision Result Comparison

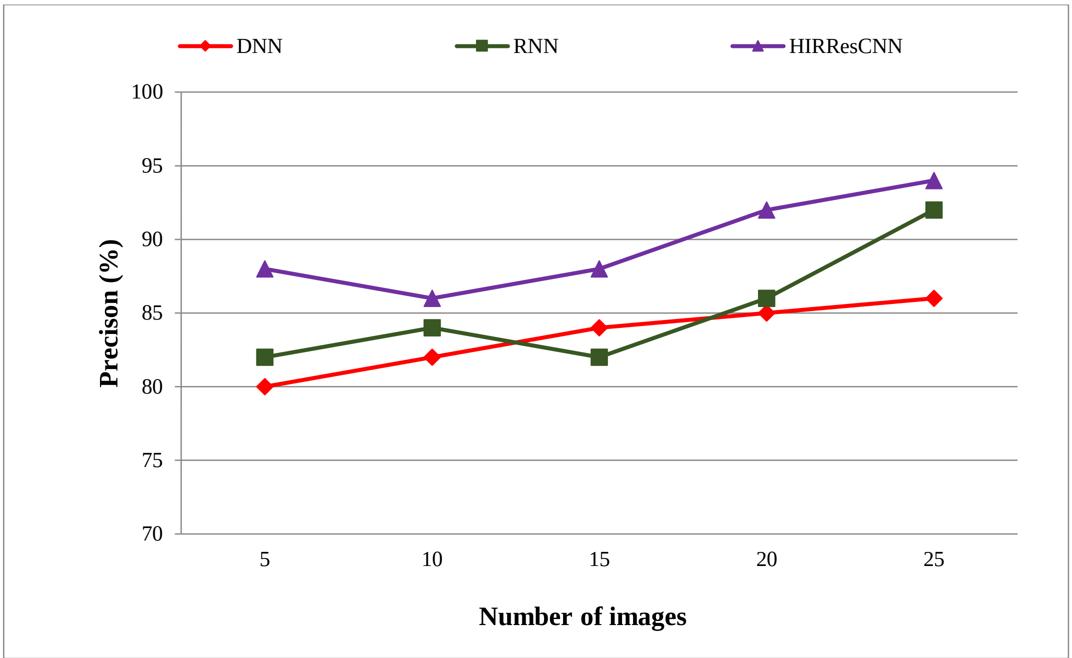

Figure 10. Precision Performance Comparison in Various Classification Methods

The correlation comparison in various classification methods between proposed AE-AMO-RNN and existing methods such as RNN and DNN are shown in figure 10. It indicates that proposed technique produces high precision rate when compared with existing techniques. Exact identification of breast cancer with high precision rate of 89.6%. is obtained. When compared with precision value of existing methods like RNN and DNN are providing good precision rates of 85.2% and 83.4%, however which is lower than the proposed method. Hence, the proposed model is effective and meaningful for short-term disease prediction.

Recall Result Comparison

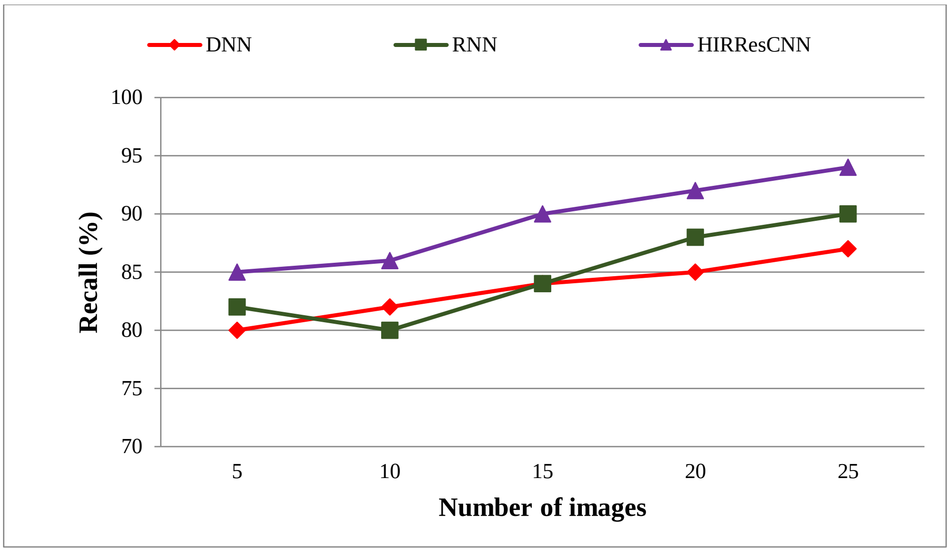

Figure 11. Recall Performance Comparison in Various Classification Methods

The recall comparison in various classification methods results between proposed HIRResCNN, and existing methods such as RNN and DNN were shown in figure 11. It shows that high recall results are produced by proposed technique when compared with existing techniques, which indicates better breast cancer detection. This is due to the fact that, proposed technique uses highly effective HIRResCNN-HSO based optimization for noise reduction and effective image enhancement stage. Around 89.4% of recall results are produced by proposed HIRResCNN in breast cancer classification, which is a greater one, while, existing RNN producing 84.8% and DNN producing 83.6% of recall results. Effectiveness of proposed system is shown by experimentation results.

F-measure Result Comparison

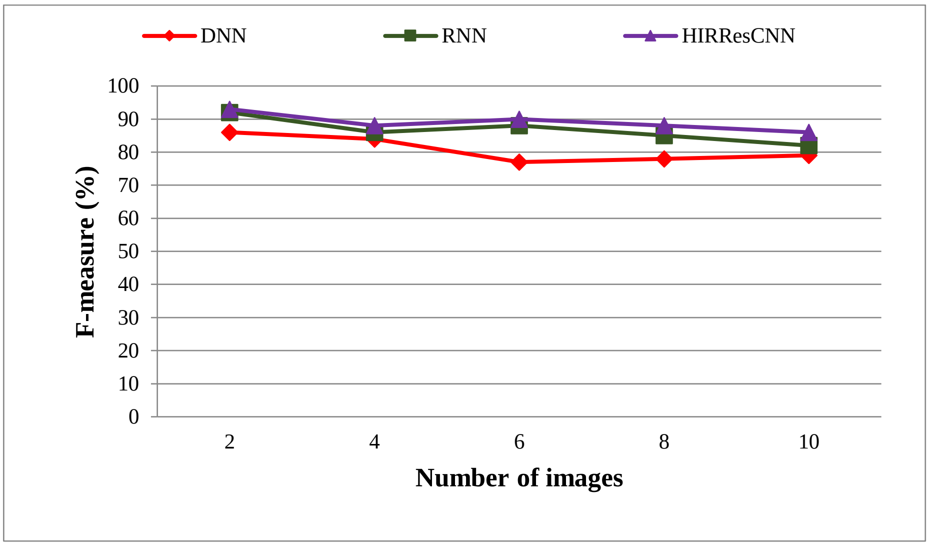

Figure 12. F-Measure Performance Comparison in Various Classification Methods

Figure 12 shows that the F-measure comparison in various classification methods results between proposed HIRResCNN, and other existing methods such as RNN and DNN. It shows that high F-measure results are produced by proposed technique when compared with existing techniques, which indicates better breast cancer detection. This is due to the fact that, proposed technique uses highly effective feature extraction detection and learning efficiency is enhanced by IRResCNN. Around 91% of F-measure results are produced by proposed IRResCNN in breast cancer classification, which is a greater one, while, existing RNN producing 86.6% and DNN producing 80.8% of F-measure results. Effectiveness of proposed system is shown by experimentation results.

Accuracy Result Comparison

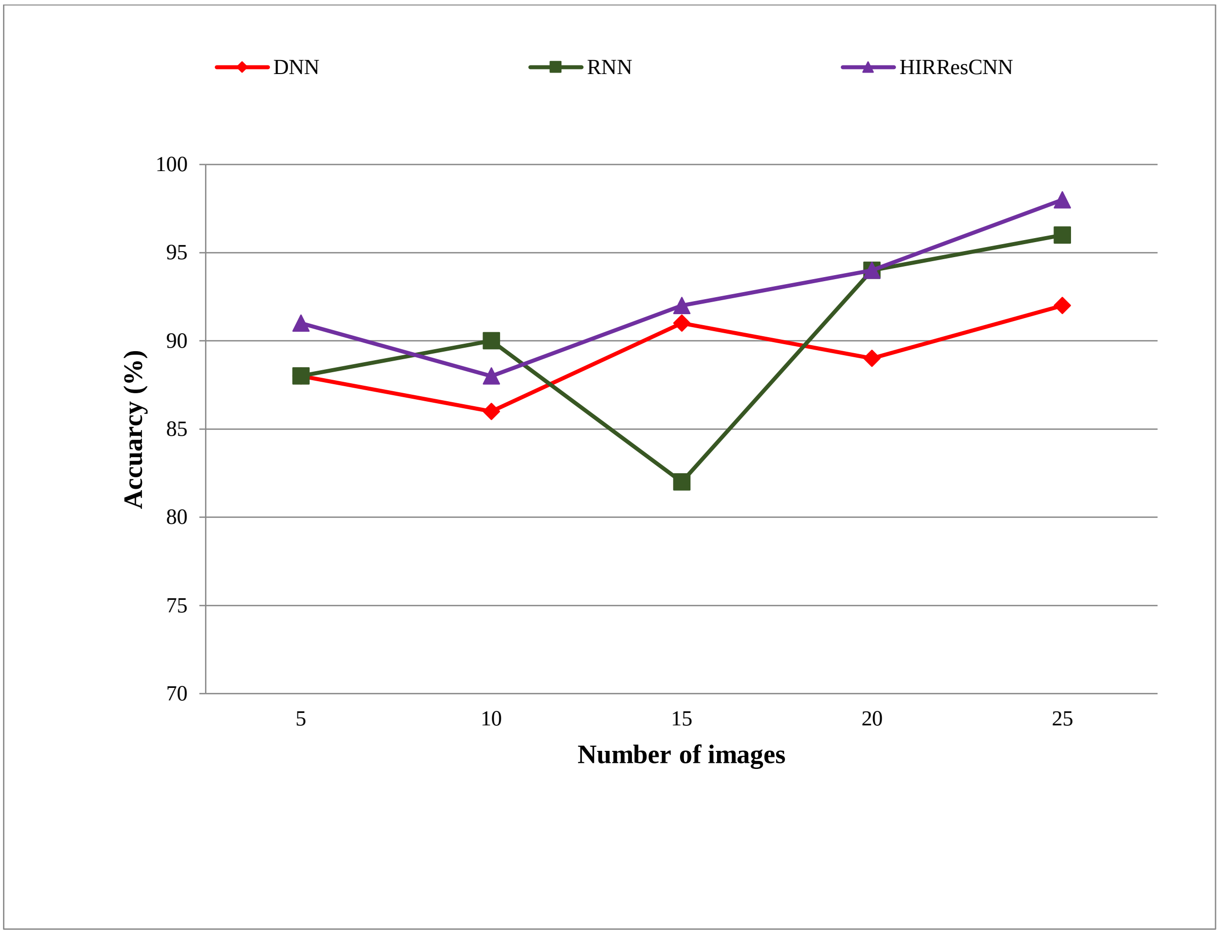

Figure 13. Accuracy Performance Comparison in Various Classification Methods

Figure 13 shows that the accuracy comparison in various classification methods results between proposed HIRResCNN and other existing methods such as RNN and DNN. It shows that high accuracy results are produced by proposed technique when compared with existing techniques. Around 92.6% of accurate results are produced by proposed AE-AMO-RNN in breast cancer classification, which is a greater one, while, existing CNN producing 90% and DNN producing 89.2% of accurate results. Effectiveness of proposed system is shown by experimentation results.

Sensitivity Result Comparison

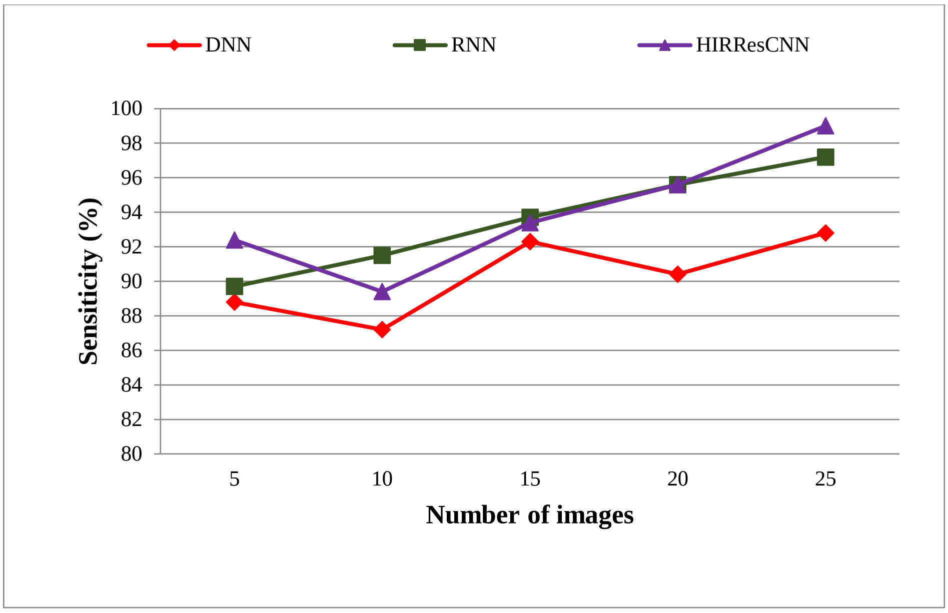

Figure 14. Sensitivity Performance Comparison in Various Classification Methods

Figure 14 shows that the sensitivity comparison in various classification methods results between proposed HIRResCNN and other existing methods such as RNN and DNN. It shows that high sensitivity results are produced by proposed technique when compared with existing techniques. Around 99% of sensivity results are produced by proposed AE-AMO-RNN in breast cancer classification, which is a greater one, while, existing CNN producing 97.2% and DNN producing 92.8% of sensitivity results.

Specificity Result Comparison

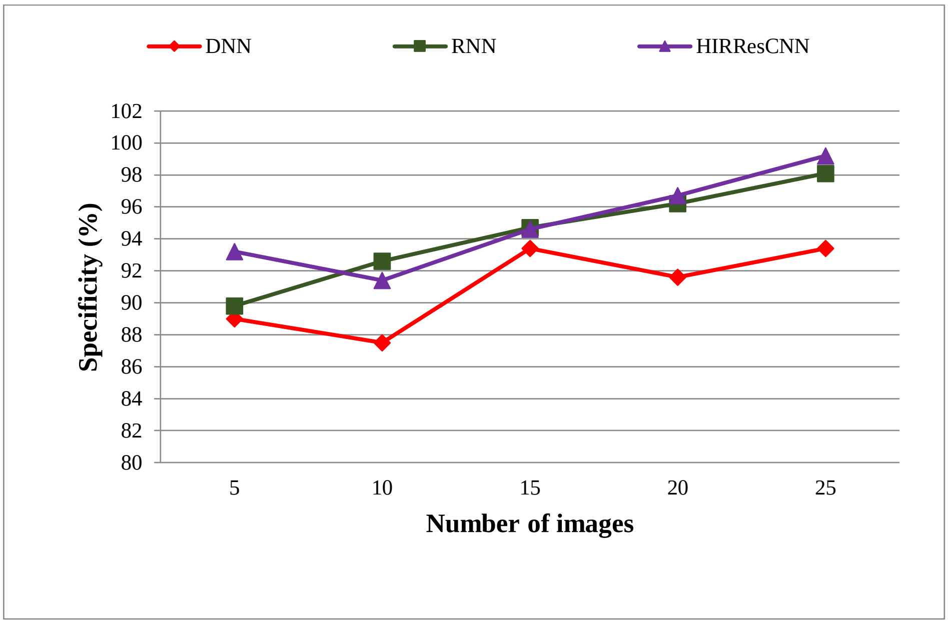

Figure 15. Specificity Performance Comparison in Various Classification Methods

Figure 15 shows that the specificity comparison in various classification methods results between proposed HIRResCNN and other existing methods such as RNN and DNN. It shows that high specificity results are produced by proposed technique when compared with existing techniques. Around 99.2% of specificity results are produced by proposed AE-AMO-RNN in breast cancer classification, which is a greater one, while, existing CNN producing 98.1% and DNN producing 93.4% of specificity results.

Conclusion with Future Work

This anticipated task discussed about HIRResCNN with HSO technique based on Deep CNN (DCNN) to classify breast tumor disease and it is termed as HIRResCNN-HSO. MIAS images are classified as malignant/ benign using this technique. There are four stages in this proposed HIRResCNN framework, namely, Pre-processing, reduction of dimensionality, segmentation and classification.

From images, noises are removed using two filtering algorithms called Median and mean filtering in pre-processing stage. Then canny edge detector is used for detecting edges. Gaussian filtering is used in canny edge detector to smoothen the images. Quality enhanced image is obtained using this edge detection and filtering process, which can be used in further image analysis process.

In the next dimensionality reduction stage, attributes are correlated using Principal Component Analysis (PCA) inclusive of related features. So, this huge dataset is minimized and only few variables are used for expressing it. In order to detect the breast cancer accurately, foreground and background subtraction is done in the third stage called segmentation stage. This results in a meaningful image, which is easy for analysing. Image representation is changed using this stage.

At last, for detecting and classifying breast cancer, a Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) is anticipated, which integrates Harmony Search Optimization (HSO) to tune bias and weight parameters and classification accuracy is enhanced using HIRResCNN-HSO model. On MIAS dataset, various experiments are conduced and results are compared with other available techniques. Promising outcomes are produced using this technique and nodules can be classified as malignant and benign. Metrics like F-measure, recall, precision, accuracy are used for evaluating proposed HIRResCNN-HSO's performance.

Around 92.6% accuracy rate is produced using this proposed HIRResCNN classifier in finding breast cancer, which is a greater value, when compared with available techniques similar to RNN and DNN. In future, autoencoders and Recurrent Convolutional Neural Networks (RCNN) can be combined and classification accuracy can be enhanced using a hybrid optimization technique using Cuckoo Search (CS) and Harmony Search Optimization (HSO) technique.

References

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer* 2015; 136(5): E359-E386.

Mehdy MM, Ng PY, Shair EF, Saleh NI, Gomes C. Artificial neural networks in image processing for early detection of breast cancer. *Computational and mathematical methods in medicine* 2017.

Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. *Asian Pacific Journal of Cancer Prevention* 2016; 17(3): 43-46.

Ayer T, Chen Q, Burnside ES. Artificial neural networks in mammography interpretation and diagnostic decision making. *Computational and mathematical methods in medicine* 2013.

Tarique M, El Zahra F, Hateem A, Mohammad M. Fourier Transform Based Early Detection of Breast Cancer by Mammogram Image Processing. *Journal of Biomedical Engineering and Medical Imaging* 2015; 2(4): 17-17.

Abdelwahed NM, Eltoukhy MM, Wahed ME. Computer aided system for breast cancer diagnosis in ultrasound images. *Journal of Ecology Health Environment* 2015; 3(3): 71-76.

Rizzi M, D'Aloia M, Castagnolo B. Health care CAD systems for breast microcalcification cluster detection. *Journal of Medical and Biological Engineering* 2012; 32(3): 147-156.

Hassanien AE, Moftah HM, Azar AT, Shoman M. MRI breast cancer diagnosis hybrid approach using adaptive ant-based segmentation and multilayer perceptron neural networks classifier. *Applied Soft Computing* 2014; 14: 62-71.

Kim JH, Cha JH, Kim N, Chang Y, Ko MS, Choi YW, Kim HH. Computer-aided detection system for masses in automated whole breast ultrasonography: development and evaluation of the effectiveness. *Ultrasonography* 2014; 33(2): 105-115.

Broeders M, Moss S, Nyström L, Njor S, Jonsson H, Paap E, Paci E. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *Journal of medical screening* 2012; 19(1): 14-25.

Talha M, Sulong GB, Jaffar A. Preprocessing digital breast mammograms using adaptive weighted frost filter. *Biomedical research* 2016; 27(4): 1407-1412.

Bickel H, Pinker K, Polanec S, Magometschnigg H, Wengert G, Spick C, Baltzer P. Diffusion-weighted imaging of breast lesions: Region-of-interest placement and different ADC parameters influence apparent diffusion coefficient values. *European radiology* 2017; 27(5): 1883-1892.

Dhahri H, Al Maghayreh E, Mahmood A, Elkilani W, Faisal Nagi M. Automated breast cancer diagnosis based on machine learning algorithms. *Journal of healthcare engineering* 2019.

Zheng B, Yoon SW, Lam SS. Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. *Expert Systems with Applications* 2014; 41(4): 1476-1482.

Wang H, Feng J, Bu Q, Liu F, Zhang M, Ren Y, Lv Y. Breast mass detection in digital mammogram based on gestalt psychology. *Journal of Healthcare Engineering* 2018.

Lee S, Park SJ, Jeon JM, Lee MH, Ryu DY, Lee E, Lee Y. Noise removal in medical mammography images using fast non-local means denoising algorithm for early breast cancer detection: a phantom study. *Optik* 2019; 180: 569-575.

Huang Q, Chen Y, Liu L, Tao D, Li X. On combining biclustering mining and AdaBoost for breast tumor classification. *IEEE Transactions on Knowledge and Data Engineering* 2019; 32(4): 728-738.

Valvano G, Santini G, Martini N, Ripoli A, Iacconi C, Chiappino D, Della Latta D. Convolutional neural networks for the segmentation of microcalcification in mammography imaging. *Journal of healthcare engineering* 2019.

Liu K, Kang G, Zhang N, Hou B. Breast cancer classification based on fully-connected layer first convolutional neural networks. *IEEE Access* 6: 23722-23732, 2018.

Deniz E, Şengür A, Kadiroğlu Z, Guo Y, Bajaj V, Budak Ü. Transfer learning based histopathologic image classification for breast cancer detection. *Health information science and systems* 2018; 6(1): 1-7.

Wang K, Patel BK, Wang L, Wu T, Zheng B, Li J. A dual-mode deep transfer learning (D2TL) system for breast cancer detection using contrast enhanced digital mammograms. *IISE Transactions on Healthcare Systems Engineering* 2019; 9(4): 357-370.

Badr EM, Salam MA, Ahmed H. Optimizing Support Vector Machine using the Gray Wolf Optimizer Algorithm for Breast Cancer Detection. *ResearchGate* 2019: 1-6.

Khan S, Islam N, Jan Z, Din IU, Rodrigues JJC. A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recognition Letters* 2019; 125: 1-6.

Wang Z, Li M, Wang H, Jiang H, Yao Y, Zhang H, Xin J. Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features. *IEEE Access* 2019; 7: 105146-105158.

Jamal A, Handayani A, Septiandri AA, Ripmiatin E, Effendi Y. Dimensionality reduction using pca and k-means clustering for breast cancer prediction. *Lontar Komput. J. Ilm. Teknol. Inf.*, 2018; 9(3): 192-201.

Abdel-Zaher AM, Eldeib AM. Breast cancer classification using deep belief networks. *Expert Systems with Applications* 2016; 46: 139-144.

Makandar A, Halalli B. Pre-processing of mammography image for early detection of breast cancer. *International Journal of Computer Applications* 2016; 144(3): 11-15.

Li H, Zhuang S, Li DA, Zhao J, Ma Y. Benign and malignant classification of mammogram images based on deep learning. *Biomedical Signal Processing and Control* 2019; 51: 347-354.

Charate AP, Jamge SB. The preprocessing methods of mammogram images for breast cancer detection. *International Journal on Recent and Innovation Trends in Computing and Communication* 2017; 5(1): 261-264.

Sha Z, Hu L, Rouyendegh BD. Deep learning and optimization algorithms for automatic breast cancer detection. *International Journal of Imaging Systems and Technology* 2020: 1-12.

Soulami KB, Saidi MN, Honnit B, Anibou C, Tamtaoui A. Detection of breast abnormalities in digital mammograms using the electromagnetism-like algorithm. *Multimedia Tools and Applications* 2019; 78(10): 12835-12863.

Kwon S, Lee H, Lee S. Image enhancement with Gaussian filtering in time-domain microwave imaging system for breast cancer detection. *Electronics Letters* 2016; 52(5): 342-344.

Taghanaki SA, Kawahara J, Miles B, Hamarneh G. Pareto-optimal multi-objective dimensionality reduction deep auto-encoder for mammography classification. *Computer methods and programs in biomedicine* 2017; 145: 85-93.

Xie H, Li J, Zhang Q, Wang Y. Comparison among dimensionality reduction techniques based on Random Projection for cancer classification. *Computational biology and chemistry* 2016; 65: 165-172.

Wahdan P, Saad A, Shoukry A. Automated breast tumour detection in ultrasound images using support vector machine and ensemble classification. *Journal of Biomedical Engineering and Biosciences (JBEB)* 2016: 1-13.

Salleh S, Mahmud R, Rahman H, Yasiran SS. Speed up Robust Features (SURF) with Principal Component Analysis-Support Vector Machine (PCA-SVM) for benign and malignant classifications. *Journal of Fundamental and Applied Sciences* 2017; 9(5S): 624-643.

Ball JE, Anderson DT, Chan Sr CS. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. *Journal of Applied Remote Sensing* 2017; 11(4): 042609.

Alom MZ, Hasan M, Yakopcic C, Tahaand TM, Asari VK. Improved inception-residual convolutional neural network for object recognition. *Neural Computing and Applications* 2018: 1-15.

Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence* 2017.

Wu S, Zhong S, Liu Y. Deep residual learning for image steganalysis. *Multimedia tools and applications* 2018; 77(9): 10437-10453.

Liang M, Hu X. Recurrent convolutional neural network for object recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* 2015: 3367-3375.

Maleki F, Nooshyar M, Fatin GZ. Breast cancer segmentation in digital mammograms based on harmony search optimization. *Tech. J. Eng. Appl. Sci.*, 2014; 4(4): 477-484.

Elyasigomari V, Lee DA, Screen HR, Shaheed MH. Development of a two-stage gene selection method that incorporates a novel hybrid approach using the cuckoo optimization algorithm and harmony search for cancer classification. *Journal of biomedical informatics* 2017; 67: 11-20.

Botchkarev A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv preprint arXiv:1809.03006*, 2018.

Ohsaki M, Wang P, Matsuda K, Katagiri S, Watanabe H, Ralescu A. Confusion-matrix-based kernel logistic regression for imbalanced data classification. *IEEE Transactions on Knowledge and Data Engineering* 2017; 29(9): 1806-1819.

Chen N, Zeng H. Spatial attention characteristics of table tennis players based on event-related potentials. *NeuroQuantology* 2018; 16(6): 811-815.

Authors

S. Prakash received his master's degree in Computer Science and Engineering and Ph.D. degree in Information and Communication Engineering from Anna University, Chennai. Currently he is working in the Department of Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology. His current research interests include the field of medical data analysis, data mining and data structures. He has published numerous papers in international journals and conferences.

K. Sangeetha received her master's degree in Computer Science and Engineering and Ph.D. degree in Information and Communication Engineering from Anna University, Chennai. Currently she is working in the Department of Computer Science and Engineering, SNS College of Technology. Her current research interests include the field of Artificial Intelligence, machine learning, and medical data analysis. She has published numerous papers in Journals and Conferences.