Received: 27 May 2023 | Revised: 7 June 2023 | Accepted: 8 June 2023

DOI: 10.1002/jbio.202300194

RESEARCH ARTICLE

JOURNAL OF BIOPHOTONICS

Fourier ptychographic and deep learning using breast cancer histopathological image classification

Leena Thomas[1,2,3] | **M. K. Sheeja**[1,2]

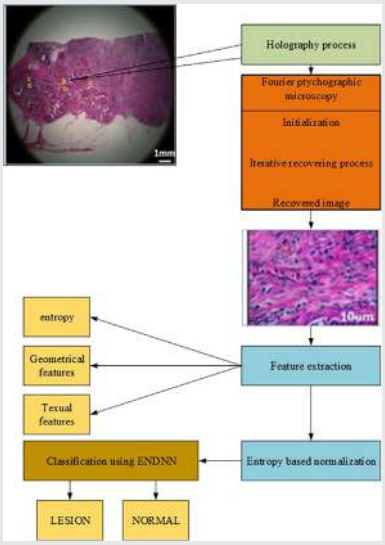
[1]Department of Electronics & Communication Engineering, Sree Chitra Thirunal College of Engineering, Thiruvananthapuram, Kerala, India

[2]APJ Abdul Kalam Technological University, Kerala, India

[3]College of Engineering Kallooppara, Pathanamthitta, Kerala, India

Correspondence
Leena Thomas, Department of Electronics & Communication Engineering, Sree Chitra Thirunal College of Engineering, Thiruvananthapuram, Kerala 695018, India.
Email: leenathomasy@gmail.com

Abstract



Automated, as well as accurate classification with breast cancer histological images, was crucial for medical applications because of detecting malignant tumors via histopathological images. In this work create a Fourier ptychographic (FP) and deep learning using breast cancer histopathological image classification. Here the FP method used in the process begins with such a random guess that builds a high-resolution complex hologram, subsequently uses iterative retrieval using FP constraints to stitch around each other low-resolution multiview means of production owned from either the hologram's high-resolution hologram's elemental images captured via integral imaging. Next, the feature extraction process includes entropy, geometrical features, and textural features. The entropy-based normalization is used to optimize the features. Finally, it attains the classification process of the proposed ENDNN classifies the breast cancer images into normal or abnormal. The experimental outcomes demonstrate that our presented technique overtakes the traditional techniques.

KEYWORDS

deep neural network, entropy, fourier ptychographic, geometrical features, normalization, textural features

1 | INTRODUCTION

Breast cancer is the most common cancer among women and is also linked to a high rate of morbidity and mortality in women [1]. The assessment of histological images underneath the microscope has been one of the main pillars in clinical applications. Histopathological slides may now be scanned and stored as digital pictures thanks to the invention of imaging sensors [2]. Because the size of digital images grows considerably with magnification, developing approaches for image processing as well as analysis, would be excellent if classification could be utilized in breast cancer computer-aided diagnostics (CAD). Scale-invariant feature transform, histogram of oriented gradient, and gray-level co-occurrence matrix are examples of hand-crafted features, and grain approaches

presented With breast cancer histopathology image analysis tasks of detection and classification [3, 4]. Other Support vector machines, for example, are well-known classifiers that have also been mentioned. Because of their remarkable outcomes in Deep learning methods like CNNs have lately attracted a lot of attention for a variety of histopathological image processing applications for breast cancer research. [5].

Recent developments and advancements in image processing have paved the way for the creation of computer-assisted diagnostic (CAD) systems that use histopathological pictures to detect and diagnose breast cancer more quickly and accurately [6]. Histopathology images vary greatly when compared to radiology images (x-ray, ultrasound, computed tomography, positron emission tomography, and magnetic resonance image) and cytopathology images. The CAD system evaluates the sample tissue's histopathological images and determines the histopathological patterns that distinguish between carcinogenic versus noncancerous conditions, as well as malignancy pictures [7]. The intrinsic complexity of histological images, like cell overlapping, minor variations between images, and uneven color allocation, makes them difficult to interpret among the fundamental issues with breast cancer histopathology image categorization [8, 9].

One of the most extensively utilized procedures is the biopsy method for detecting women with breast cancer, tissue is collected, analyzed, and examined under a microscope by a pathologist for abnormalities [10]. This method is difficult, error-prone, and time-consuming yielding varying findings determining the amount of expertise of the pathologist's competence. Not only can an automated and efficient method will not only aid in the diagnosis of breast cancer but also help to reduce human effort [11]. Cancer is a major public health issue that affects people all over the world. Breast cancer is the second leading cause of mortality among women's cancer kinds after lung cancer. Furthermore, as compared to other cancers, breast cancer has an extremely high death rate [12]. Despite fast improvements in medicine, histological image analysis is still the most often utilized approach for diagnosing breast cancer [13]. The classification job is the most essential of all the responsibilities for histopathology image analysis [14, 15], which is the automated and accurate categorization of high-resolution histopathological images.

Clinical histopathological imaging practice is now primarily reliant on pathologists' qualitative analysis by hand [16]. This technique of analysis has at least three flaws. Major, pathologists are in limited supply across the world, particularly in developing countries besides slight organizations. The shortage of resources and unequal distribution of those resources is a significant issue that must be addressed. Second, whether the histological diagnosis is valid or erroneous is determined by significant professional skill and long-term clinical knowledge of the pathologist [17, 18]. A spate of diagnosis blunders has occurred from pathologist subjectivity. Next, pathologists are prone to tiredness and inattention due to the intricacy of histological pictures. To address these issues, it is critical to create automated and accurate histopathological image analysis tools, particularly classification approaches [19]. The usual method for histopathological object recognition splits a big image split smaller patches, utilizes a CNN to categorize every reinforcement, and afterward aggregates the classification performance of these patches, such as by majority vote, to provide classification accuracy [20]. Contribution to the paper and its structure:

1. In the following steps, we provide a novel categorization framework in our paper: Using the Fourier ptychographic (FP) technique to calculate high-resolution computer-generated integral holograms.
2. The second feature extraction process includes entropy, geometrical features, and textural features.
3. Finally, it attains the classification process of the proposed entropy-based normalization deep neural network (ENDNN) classifier that classifies the breast cancer images into normal or abnormal.
4. Each stage of a methodical approach working process was explained in depth in the subsequent sections.

The other sections of the paper are divided as follows: Section 2 explains its suggested methodology based on relevant research, and Section 3 discusses the proposed FP and deep learning using breast cancer histopathological image classification approach. In Section 4, the experimental results are analyzed, and in Section 5, the conclusion is offered.

2 | LITERATURE SURVEY

Edson D. Carvalho et al. [21] propose a technique for developing agile processes to promote sustainable development phylogenetic diversity indexes and histological breast images were divided into four groups: invasive carcinoma, in situ carcinoma, normal tissue, and benign lesion. In addition, bandwidth imagery retrieval was employed to verify the recognition accuracy and provide a rating for unlabelled image sets in this study. The results were quite consistent, and they were beneficial in the development of a CADx platform for aid experts in large health centers.

Ala'a El-Nabawy et al. [22] a feature fusion approach has been developed for extracting the finest features from various datasets. The METABRIC datasets, which also include clinical data, gene expression, CNA as well as

CNV data, and histopathology images, are used in the proposed method. Several machine learning classifiers were used to create and test different data profiles. The purpose of this study is to show that combining features from different METABRIC datasets enhances breast cancer subtype classification accuracy. Furthermore, histopathological images show encouraging findings for Pam50 subtypes, and when applied to a larger population, it is predicted to increase the accuracy of IntClust subtyping.

Yun Jiang et al. [10] construct a CNN with such a convolutional layer, a small SE-ResNet component, as well as a fully connected layer using a convolutional layer, a tiny SE-ResNet module, and a completely connected layer. In addition, a smaller SE-ResNet module is developed, that progresses the coupling of the outstanding component with both the squeeze-and-excitation chunk besides obtaining equal performance through small datasets. A new learning rate scheduler has also been included, allowing for outstanding performance without the need to fine-tune the learning rate. This approach was used to divide histological images of breast cancer into benign, malignant, and eight subcategories automatically.

Yun Jiang et al. [23] for grouping two CNN-based approaches are suggested for analyzing Histopathology photos of breast cancer stained with H&E. Histopathology image dataset for breast cancer, the first method's CNN demonstrates that the HEBCNet obtains better classification accuracy. The transfer learning hybrid model structure is used in the second technique. The complexity and limitations of manual feature extraction are avoided, and classification accuracy is enhanced, by employing separate pre-training models to extract features and training two classifiers.

An Pan et al. [24] using Euler's formula, establish a relationship between FPM with patterned irradiance microscopy are two different types of microscopy. The research concludes with a discussion of the difficult challenges and potential applications. FPM may be used as a framework for dealing with phase loss and system limitations in imaging systems. Speckle imaging, incoherent imaging for retina imaging, and large-field-of-view fluorescence imaging can all benefit from this information.

Yi Xiao et al. [25] demonstrate a laser-illumination Fourier ptychographic microscopy (FPM) technique based on a digital micro-mirror device (DMD) for high-speed and high-resolution label-free imaging applications. We may extract phase and intensity images of materials, such as stained and unstained cancer tissue slices and cells, using this approach. A 532 nm laser provides system illumination, and two DMDs are used to achieve active illumination angle scanning and dynamic filtering. This device has achieved real-time imaging at 42 frames per second with a resolution of about 1 m. The comparison analysis of existing papers and their advantages and disadvantages is given in Table 1.

3 | PROPOSED METHODOLOGY

This paperwork is to develop holography using FPM and breast cancer classification using ENDNN. The FP technique uses high-resolution computer-generated integral holograms that were originally used. The approach starts to create a high-resolution complicated hologram with a random guess, then uses iterative retrieval with FP constraints

TABLE 1 Comparison analysis of existing papers and their advantages and disadvantages.

Author	Year	Technique	Advantages	Disadvantages
Edson et al. [21]	2020	Textural features and CBIR	Best efficiency	Improve robustness and making the method more generic are needed
Ala'a El-Nabawy et al. [22]	2020	METABRIC breast cancer subtype classification	Highest accuracy achieved	A small number of samples
Yun jiang et al. [10]	2019	Convolutional neural networks with a small SE-ResNet module	Achieves accuracy	Improve reliability is needed
Yun Jiang et al. [23]	2019	Convolutional neural network	Help reduce costs and increase the efficiency of the process	Better accuracy is needed
An Pan et al. [24]	2020	Computational imaging	High-throughput imaging	The optimization problem in nature
Dr. Yi Xiao et al. [25]	2021	Digital micro-mirror device-based laser-illumination Fourier ptychographic microscopy	High-speed and high-resolution	Achieving active illumination angle is needed improvement

to stitch to restore the elevated hologram by combining limited multi-view images obtained from either the elemental images acquired by integral imaging. Then the second step the feature extraction process includes entropy, geometrical features, and textural features. Then entropy-based normalization is used to optimize the features. Finally, it attains the classification process of the proposed ENDNN classifier that classifies the breast cancer images into normal or abnormal. A block diagram of the system is shown in Figure 1 suggesting the categorization technique.

As indicated in Figure 1 above; using FPM and breast cancer classification using ENDNN. Based on this, classifies the breast cancer images into normal or abnormal which is explained detailedly in the following sections;

3.1 | Holography process

Holography is the technique of creating a hologram. Light from a laser records an image of the desired item on film or a photographic plate when a hologram is created. An embossed hologram is created by using a mirror-like material to back a transmission hologram, allowing it to be seen when illuminated from the front. Holography is important to modern technology because it allows us to regulate the flow and direction of light. To generate 3D expansion, we employ holographic methods. We make use of a tiny projector with a small pupil. The following is a diagrammatic representation of the holography process;

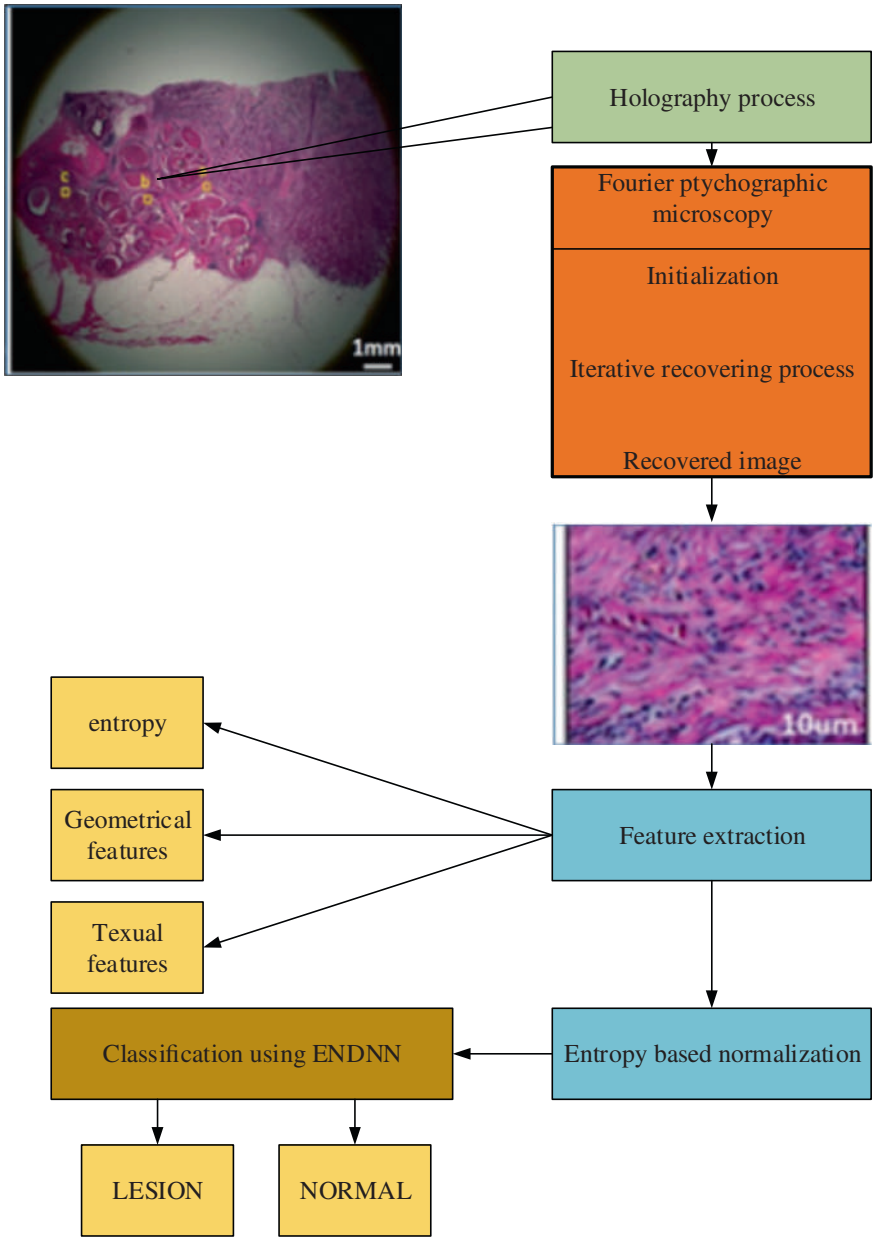


FIGURE 1 Proposed classification technique.

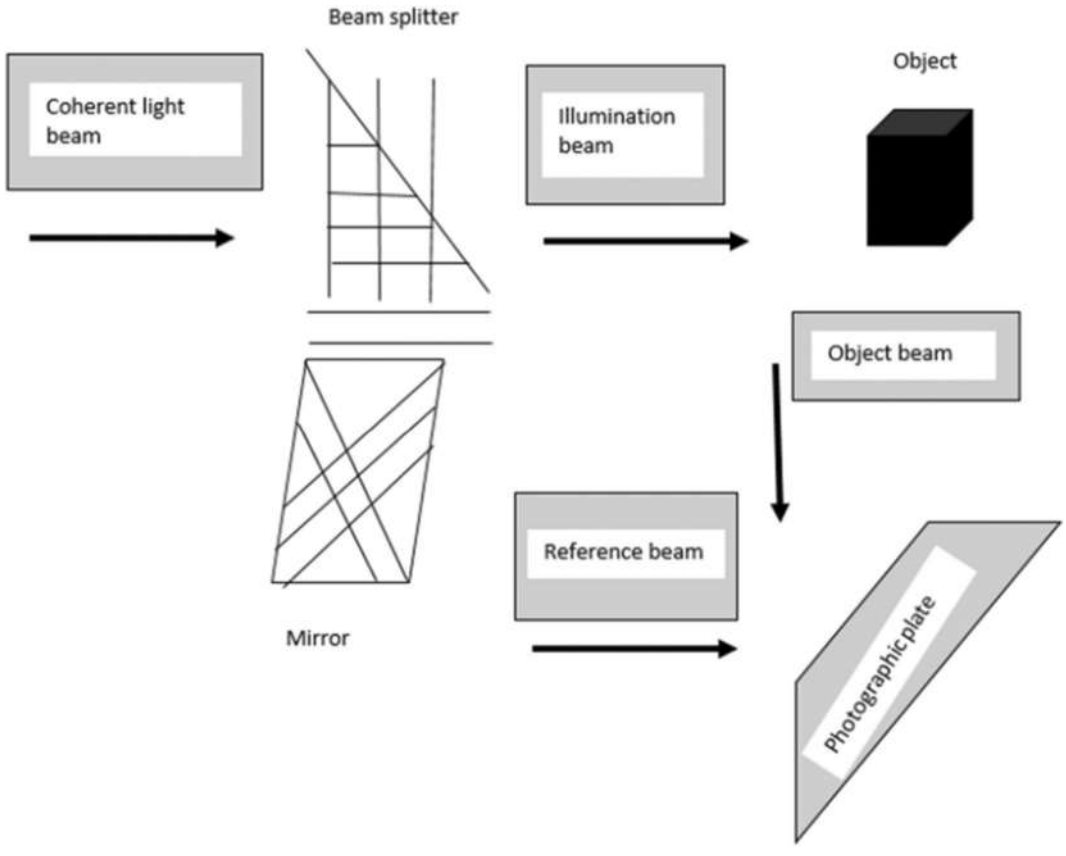


FIGURE 2 Recording holography.

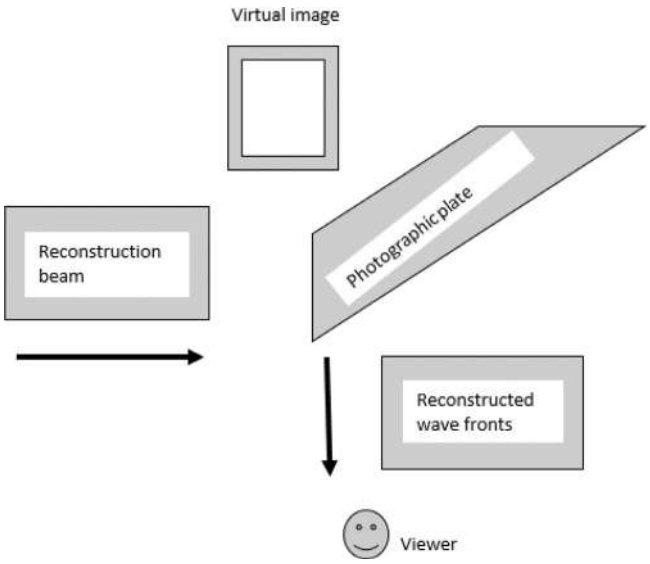


FIGURE 3 Reconstructing holography.

Holography is indeed a technique for capturing as well as reconstructing a light field whenever the original light field is not any longer visible due to the removal of the original items, as seen in the bow Figures 2 and 3. The light waves of the two laser beams collide and interfere as they reach the recording medium. The recording medium bears the impression of this interference pattern. The pattern seems random because it depicts how light inside the scene is associated with both the original light source, but not by the light source itself. It is possible to think about the interference pattern dynamically encoded representation of the environment that requires a certain key for accessing their elements (the actual light source) [27]. A laser, comparable to recording the hologram, is then flashed across the produced film to provide the missing key. The holograms diffract this beam surface pattern when it lights it. This scatters a light field onto the hologram that is identical to the one created by the scene (Figure 4) [28].

A photograph of a small section of an unbleached transmission hologram examined under a microscope is shown here. After completing the holography process, features are extracted. The feature extraction process includes entropy, geometrical features, and textural features, it will be explained in full as follows:

3.2 | Fourier ptychographic microscopy

FP is a computational imaging approach based on optical microscopy that involves generating a larger numerical

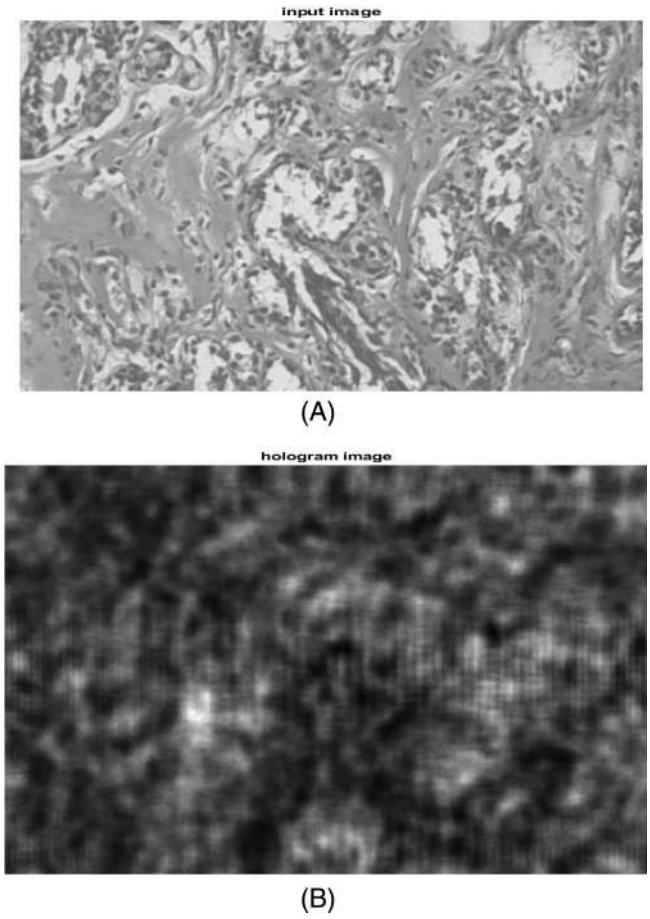


FIGURE 4 (A) Input image. (B) Hologram reconstructed image.

aperture (NA) from a sequence of full-field images recorded at different coherent illumination angles and velocities, resulting in higher resolution than a standard microscope. The suggested high-resolution FPM system is roughly depicted in Figure 3. An illuminating system and an imaging system make up the system. The following are thorough explanations of the working processes.

3.2.1 | Fourier ptychographic setup details

The illumination is provided by a coherent light source, a He–Ne laser of 5 mW, and a wavelength of 632.8 nm. The dual-axis galvanometer provides the sample with illuminations at different incidence angles and is focused on the sample by condenser lenses. FP is a technique for stitching together a succession of lighted, low-resolution intensity photographs in Fourier space to create a high-resolution complex sample image. This technique reconstructs the complex image of the object with quantitative phase information using retrieval of phases that are iterative algorithms. Finally from the reconstructed phase, the refractive index of the specimen is determined. After this, holography starts its process which is explained detailedly as follows.

The PSI-FPM system's configuration is shown schematically in Figure 5. An image patch illumination system and a detecting system make up the system. A laser beam with a wavelength of $\lambda = 632.8$ nm, a beam expanded and collimated instrument, phase-only liquid crystal on silicon (LCoS) with a polarization beam splitter (PBS), a 4-f system, a conder lens, and a sample make up the illumination system. A $4\times/0.13$ objective, a $0.75\times$ tube lens, and a monochrome CCD camera are included in the detention system, which is standard biological microscopy equipment (MI52 Mshot).

The phase-only spatial light modulator employed in the picture patch lighting system is Holoeye LETO LCoS, which has a resolution of 1920×1080 pixels and a pixel pitch of 6.4 μm. After a PBS, which ensures the polarization state of the beam is aligned with that of the LCoS, the collimated and enlarged laser beam illuminates the image patch loaded on the LCoS. The patch picture patterns are chosen to be 512×512 pixels in size for each illumination. The projected beam from LCoS is transmitted through a 4-f system consisting of identical-focal-length lenses 1 and 2. Because the condenser's back focal length is short, a 4-f method is used to communicate the loaded image patch patterns to the condenser's rear focal plane. Furthermore, the 4-f system makes it simple to insert a spatial filter between lenses 1 and 2 to suppress diverse noises, which improves experimental operability. To project the dense image patch pattern onto the sample, an inverted $4\times/0.13$ objective with infinity-corrected is chosen as the condenser.

Finally, the detection mechanism captures the beam with the sample information. To make the sample information image on CCD, the detecting system's objective is infinity-corrected and should be utilized with an $f = 180$ mm tube lens. With the $4\times/0.1$ objective and $0.75\times$ tube lens, the CCD pixel size is 2.4 μm, which equals 0.8 μm in the sample plane. For two objectives, the sample is put at a working distance; no defocus is required because the image patch pattern employed in our method is known. The image patch pattern is loaded with one-pixel scanning in the x-direction and three-pixel scanning in the y-direction sequentially, allowing the monochromatic CCD camera to capture a series of images with different speckle illumination, which can then be used to reconstruct a high-resolution complex image. The following is a diagrammatic representation of the laser-overall FPM's flow chart.

Figure 6 represents the overall flowchart of laser FPM. The next stage of the image reconstruction process is explained detailedly as follows.

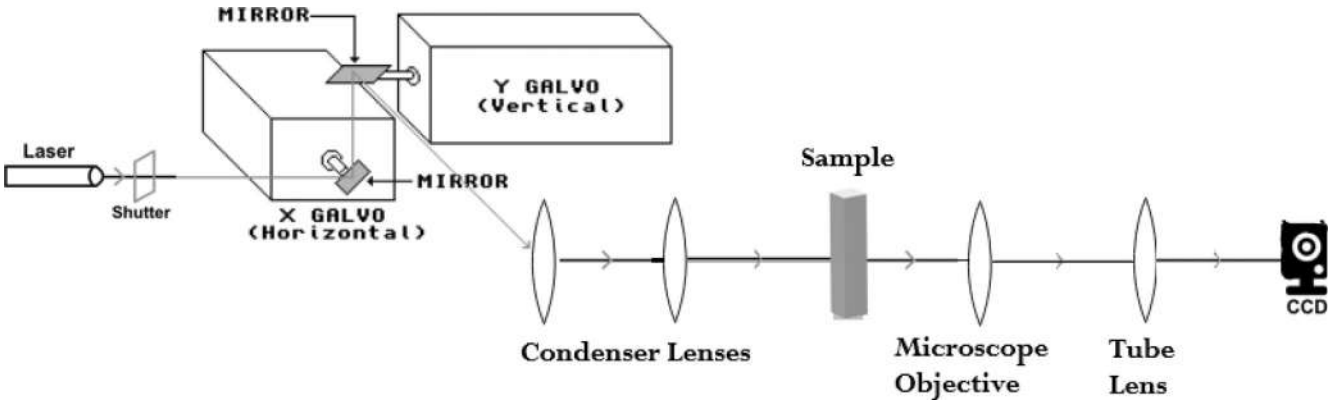


FIGURE 5 Structure of Fourier ptychographic set up detail.

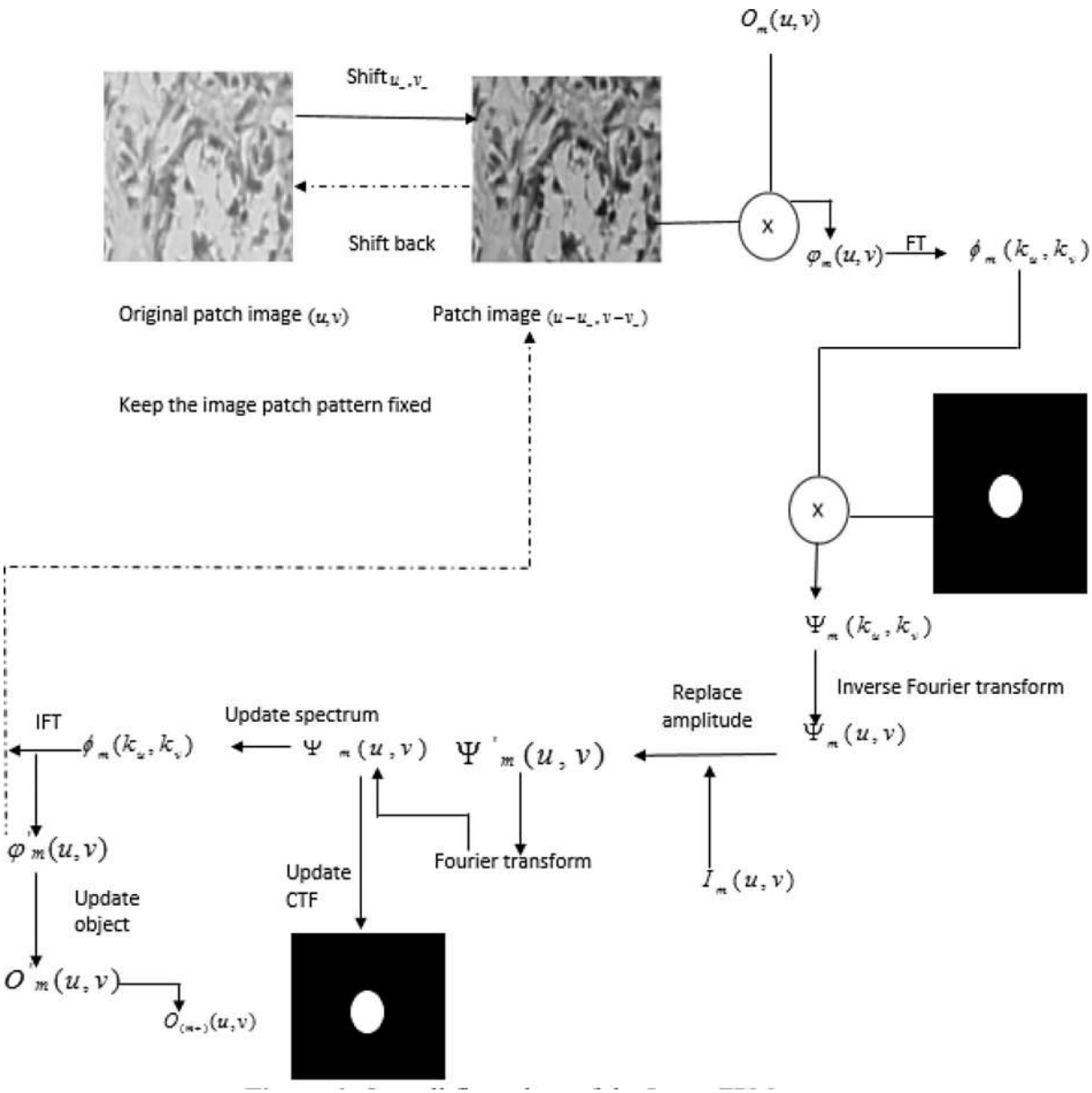


FIGURE 6 Overall flow chart of the laser-FPM.

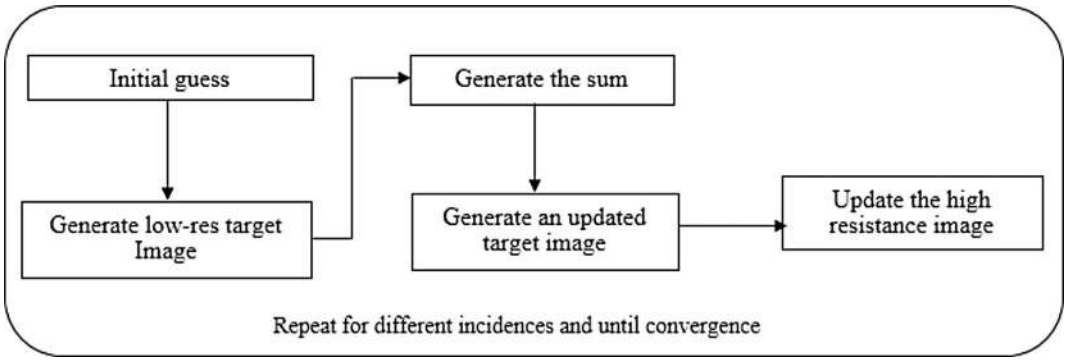


FIGURE 7 Recovery procedures of the FP scheme.

3.2.2 | Image reconstruction

High-frequency components, on the other hand, reflect the sample's features, while low-frequency components indicate the sample's brightness and shape. According to the imaging principle, the optical system serves as a low-pass filter, preventing high-frequency information from going through. As a result, turning on the center micromirror arrays allows for the acquisition of the sample's outline while the details are blurred. It is possible to modify the state of the micromirror array and modulate the reflected beam into a parallel beam that illuminates the sample at varying angles using programming controls, therefore increasing the system's NA. To generate a parallel beam perpendicular to the sample plane, the first stage in the imaging process is to turn on the micromirrors array from the center of the DMD chip. Following that, the rest of the micromirror arrays are turned on in spiral order. Finally, a series of raw photos matching various illumination angles are recorded. These unprocessed pictures are then used to create a high-resolution image. The following are the recovery techniques and algorithms for the FP scheme:

The procedure for recovering from Figure 7 and Algorithm 1 depicts the FP system; It begins with only an estimate of the sample profile with great resolution, comparable to the single state methodology: $\sqrt{I_h}e^{i\phi_h}$ lots of low target images to various coherent states are produced using this sample estimate. Second, to form the incoherent mixture, the target images' intensity elements are combined I_t. The third factor is the proportion of reality to potential measurements I_m besides I_t updating the intensity aspects of the target pictures while leaving the phase components alone. The modified target images are then utilized to adjust the sample estimate's corresponding spectral regions. Finally, the method is iterated numerous times until the answer converges for all intensity data. The total number of coherent states is proportional to the computational cost of the state-multiplexed FP technique. Two-state multiplexing, for example, takes twice as long to compute as single-state multiplexing does.

Algorithm 1 **Fourier ptychography scheme**

Begin
- **Step 1:** Initialization
- **Step 2:** Generate low-resistance target images corresponding to different coherent state
- **Step 3:** sum up the intensity components of target images
- **Step 4:** low resistance images
- **Step 5:** update the corresponding spectral Region of the high-resistance image

End

The substitution of intensity procedure is the main distinction between the disclosed method and the FP with a single state. The objective image's intensity element is immediately substituted as determined by the actual measurement in the single-state FP, but the phase component remains unaltered. Using the ratio of the incoherent mixture to the actual measurement and on the other hand, the presented state-multiplexed technique updates the targeted pictures' luminance components. The intensity accumulation of several coherent phases matches the actual coherent combination as a result of the current updating mechanism, while the amplitude of individual modes is retained.

The micromirror arrays are turned on one by one in this experiment. The first micromirror array is chosen, which is concentric with the zoom mechanism and allows the beam to illuminate the sample normally. Each micro mirror's position is shown here $\text{array}_{a,b}$(row a column n) and relates to a distinct angle of illumination $(k_{u,a,b}, k_{v,a,b})$ of the parallel wave generated by the zoom system and the condenser lens. The algorithm's flow

chart is presented in Figure 8 below. Intensity image with a low resolution $P_{a,b}$ with A CCD camera captures particular angle information. The starting value of the high-resolution complex amplitude image is approximated as a random value before the FPM iteration process starts $o^o(u,v)$ thus the pupil function is calculated as follows: $q^o(k_u,k_v)$. Fourier transformation can be used to acquire the spectrum, and the auxiliary function for each micromirror array can be built using the spectrum and pupil function.

$$O(k_u,k_v)=\delta\{o(u,v)\}, \tag{1}$$

$$\Psi^j_{a,b}(k_u,k_v)=o^j(k_u-k_{u,a,b},k_{v,a,b})q^j(k_u,k_v). \tag{2}$$

Anywhere $\Psi^j_{a,b}$ in the Fourier spectrum just after the micro mirror's pupil array$_{a,b}$ enlightenment. After that, an inverse Fourier transform is applied to the acquired auxiliary function to produce an estimate of the complex field of the low-resolution image. $\varphi_{a,b}$.

$$\phi^j_{a,b}(u,v)=\delta^{-1}\left\{\Psi^j_{a,b}(k_u,k_v)\right\}. \tag{3}$$

The actual image intensity takes its place after that $Z_{a,b}$ while leaving the phase component alone the low-resolution image's modified Fourier spectrum is as follows;

$$\varphi^j_{a,b}(u,v)=\sqrt{\frac{Z_{a,b}}{\left|\phi^j_{a,b}(u,v)\right|^2}}\phi^j_{a,b}(u,v), \tag{4}$$

$$\varphi^j_{a,b}(k_u,k_v)=\delta\left\{\phi^j_{a,b}(u,v)\right\}. \tag{5}$$

Lastly, the aperture position is used $(k_{u,a,b},k_{v,a,b})$ as well as the new Fourier spectrum $\varphi^j_{a,b}(k_u,k_v)$, the object and pupil functions are updated using two functions.

$$\begin{aligned}O^{j+1}(k_u,k_v)=O^j(k_u,k_v)+\beta\frac{|q^j(k_u+k_{u,a,b},k_v+k_{v,a,b})|\cdot[q^j(k_u+k_{u,a,b},k_v+k_{v,a,b})]^*}{|q^j(k_u,k_v)|_{\max}\cdot\left(\left|q_j(k_u+k_{u,a,b},k_v+k_{v,a,b})\right|^2+\gamma_1\right.}\\ \cdot\left[\phi^j_{a,b}(k_u+k_{u,a,b},k_v+k_v+k_{v,a,b})\right.\\ \left.-O^j(k_u,k_v)q^j(k_u+k_{u,a,b},k_v+k_{v,a,b})\right],\end{aligned} \tag{6}$$

$$\begin{aligned}P^{j+1}(k_u,k_v)=P^j(k_u,k_v)+\alpha\frac{\left|O^j(k_u-k_{u,a,b},k_v-k_{v,a,b})\right|\cdot\left[O^j(k_u-k_{u,a,b},k_v-k_{v,a,b})\right]^*}{\left|O^j(k_u,k_v)\right|_{\max}\cdot\left(\left|O_j(k_u-k_{u,a,b},k_v-k_{v,a,b})\right|^2-\gamma_1\right.}\\ \cdot\left[\varphi^j_{a,b}(k_u-k_{u,a,b},k_v-k_{v,a,b})\right.\\ \left.-O^j(k_u,k_v)q^j(k_u+k_{u,a,b},k_v+k_{v,a,b})\right].\end{aligned} \tag{7}$$

Wherever γ_1 and γ_2 (set as $\gamma_1=1$ and $\gamma_2=10$) regularization constants are used in FPM to maintain numerical stability, α and β (both sets to 1) are the coefficient parameters of the FPM iteration's step size. The next micromirror array is turned on when a micromirror is changed. After updating all of the micromirror arrays, the aforementioned stages are continued until the FPM converges, and the following is the entire flowchart of laser FPM.

3.3 | Feature extraction

Histopathological image classification for breast cancer, feature extraction is crucial. For categorization purposes, the advantageous aspects of histopathological pictures are taken from the image. It is difficult to get a nice feature out of an image. There are a variety of feature extraction approaches to choose from. We extract entropy, geometrical characteristics, and textural features from photos in this paper. These features are explained detailedly as follows.

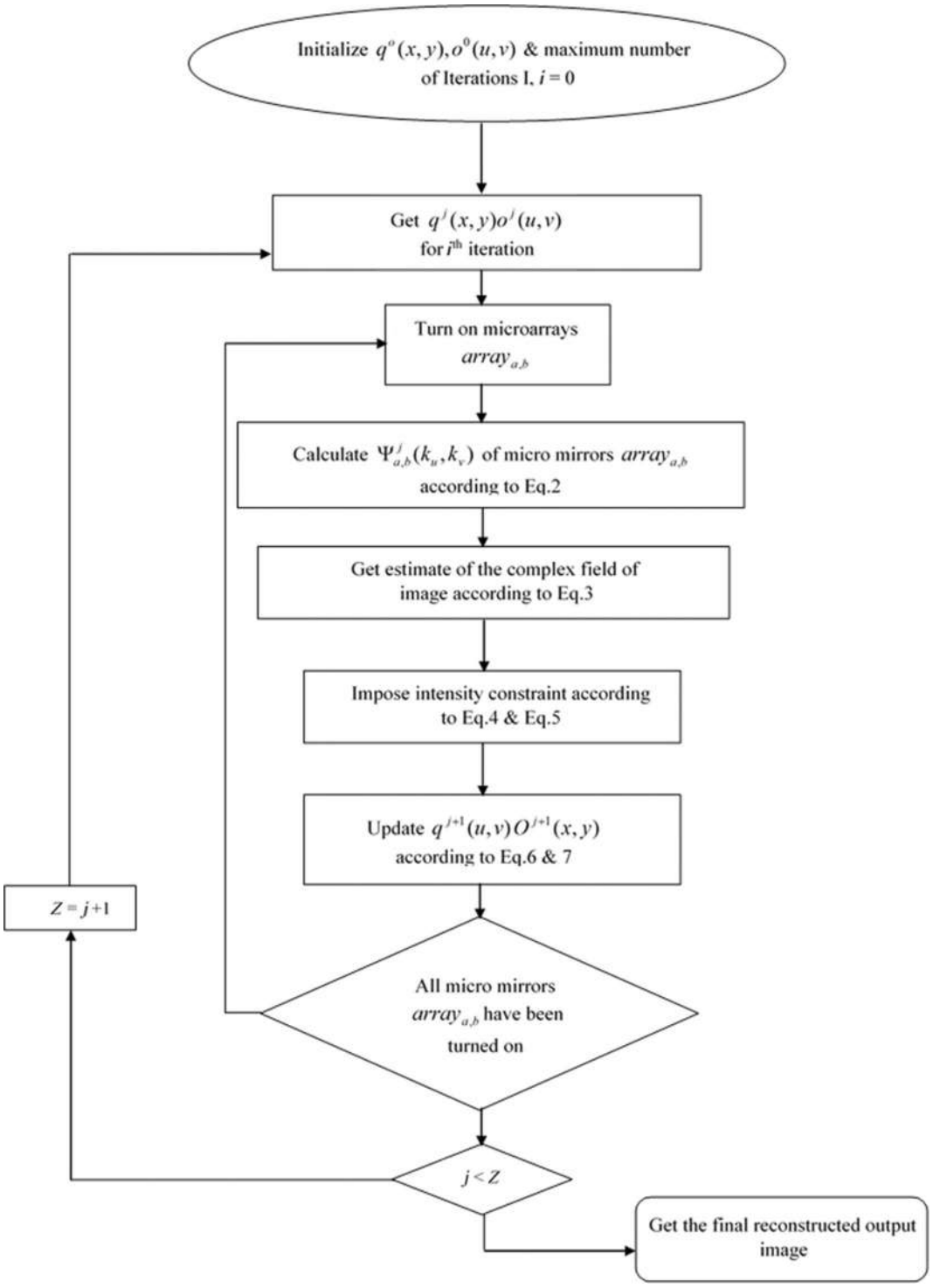


FIGURE 8 Overall flowchart of the laser-FPM.

3.3.1 | Feature extraction based on entropy

Entropy is a measure of the degree of uncertainty in a system. When all distribution points have the same probability, the entropy of the distribution is maximized. The higher the entropy, the less predictable the occurrences are, and the lower the entropy, the more predictable the events are. It is a good idea to choose the attribute with the maximum information acquisition. It does so by lowering entropy, which improves predictability. When the entropy decreases as a result of representation selection, information gain is positive. A reduction in entropy indicates a reduction in unpredictability, as well as an improvement in predictability.

The sum of every image makes a random variable for which the expected esteem or average is the entropy. Entropy is evaluated for the informational collections by utilizing conditions (8)

$$E_y(\text{set}) = \sum_{i=1}^{M} P(\text{val}_i)\log_2\{P(\text{val}_i)\}, \tag{8}$$

where, $P(\text{val}_i)$ is the probability of choosing the ith feature.

3.3.2 | Feature extraction based on geometrical features

For several types of cell differentiation, morphological features are used to characterize the nucleus shape. Various morphological characteristics Due to the nucleus's diversity of morphologies, information such as surface, exclusion zones, equivalent diameter, stability, irregularity, form identity, boundary, density, breadth, massive duration, as well as minor-axis length is obtained. The following is a thorough explanation:

Nucleus area

The nucleus pixel region is used to calculate the area of an item in a 2D image. The following formula can be used to compute it:

$$Z = \sum_{x=1}^{u}\sum_{y=1}^{v} Q(x,y), \tag{9}$$

where the nucleus area is represented by Z, while the segmented brest cancer-based histopathological image picture or region of interest (ROI) is represented by Q, which has u rows and v columns.

Perimeter

On the nucleus's boundary, the perimeter is the spacing amongst each adjacent pair of pixels. Counting the total number of edge pixels associated with the item is the easiest approach to determining the nucleus' perimeter. The following mathematical formula demonstrates this:

$$p_i = \sum_{u,v\epsilon B_i} Q(u,v), \tag{10}$$

where p_i seems to be the circumference, Q remains the image that has been split or a ROI through u rows as well as v columns, besides B_i characterizing the boundary pixels.

Aspect ratio

This factor is especially useful for discriminating between spherical versus noncircular objects, as well as needle-like shapes. The aspect ratio ranges from 0 to 1 in value. The higher the number, the more elongated the cells are, whereas the lower the number, the more benign the cells are.

Solidity

The ratio of ROI's area A to its convex hull is known as solidity. It is a necessary feature, and it is calculated as follows:

$$\text{Solidity} = \frac{\text{area}}{\text{convexhull}}. \tag{11}$$

3.3.3 | Feature extraction based on textual features

The fundamental idea of implementing FP and deep learning using classification of histopathological images of breast cancer. This project's goal is to acquire texture value. Every image is encoded after being separated into "n" sub-blocks. Histopathological Image assesses whether the phases of neighboring pixels are homogeneous or heterogeneous for each sub-block. As a result, the texture features of the image are more easily found during image retrieval, making it easier to detect similarities with the texture features.

The following objectives were pursued in the experiments:

1. To establish if the conceptually hypothesized topographical assets feature relationship aligns with both the practical scenario by investigating the degree to which each of the five textural characteristics is related to each of the five textural properties.
2. To find out how closely the features are linked to one another, as well as how the qualities are associated with each other;
3. To see if particular configurations of features might suggest resemblance among distinct textural patterns, and thus to see how close the features are to human perception of textures.

After the extraction of features, all of the extracted features are then moved on to the next phase of classification, which is described in full below;

3.4 | Entropy-based normalization deep neural network

ENDNN classifier classifies the breast cancer images into normal or lesion.

3.4.1 | Entropy-based DNN

Before flowing via a max-pooling layer to the convolutional layer, the input is first transmitted through a DNN convolutional layer, which performs entropy-based normalization. The method is then repeated for the

max-pooling layer, with a layer that is entirely coupled to the Softmax regressor in the convolutional layer. The structure of the building projected ENDNN is depicted in Figure 9.

3.4.2 | Entropy-based normalization

Normalization is a technique for altering the range of pixel force values. To improve the range of Histopathological Images, an entropy-based normalizing computation is performed. The following condition represents entropy-based normalization mathematically:

$$\mathrm{EBN} = \text{Entropy based value} \times \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \tag{12}$$

where, $X_{\min}$ along with $X_{\max}$ the minimum as well as data's maximum values X, wherein EBN stands for entropy based on esteem normalized or else standardized histopathological image.

Convolution layer: The first layer of the network is utilized to retrieve the image in its original clear format using a matrix or kernel. By understanding the pixels, the link between the histopathological image characteristics is maintained. Analyze the data fields thoroughly for the following layers of convolution. The requirement obtained in the equation is satisfied by this layer (5). In either instance, the element map is the result of the convolution

$$y_k = \sum_{n=0}^{N-1} x_n h_{k-n}. \tag{13}$$

Wherever x the input images, h are fp, as well as N is the number of essentials in x. The output is y. The subscripts denote information nth as a constituent of histopathological image.

Max-pooling layer: The purpose of the pooling layer is to minimize the system's boundary calculations. To minimize the image's dimensionality, the max-pooling layer is utilized as well as the histopathological image, and it is also known as the downsampling layer since it contributes to the contribution of the following layer.

Fully connected layer: The previous layer's input is received by each neuron, which is beneficial for developing the most neurons from the day before levels.

Softmax: It is used to denote the different digits of {0, 1}, to handle numerous classes, the assumption of labels logistic regression is used.

$$p_i = \frac{e^{x_i}}{\sum_1^k e^{x_i}}. \tag{14}$$

Anywhere, x is the network input, and ENDNN is used to classify the input images based on the entropy value and the outcomes to determine whether they are injured or normal. We applied the entropy-based deep neural architecture in a precise order. Furthermore, its parameter learning includes both the pre-training as well as fine-tuning stages.

3.4.3 | Pre-training stage

The DBN paradigm allows relies mostly on the network it provides observable activations states the network belief is all about the hidden units of a network. We used RBM to solve the problem described above.

Restricted Boltzmann machine: BM is a type of Markov random sequential manner with one layer of hidden stochastic units and one layer of visible or observable stochastic units.

Step 1: We begin by initializing the visible units, which are the features that have been chosen f_{si} for the vector of training

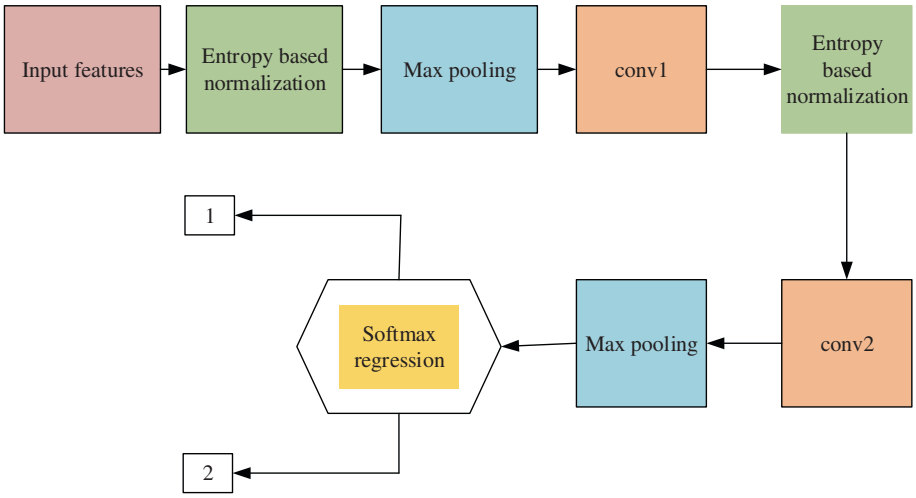


FIGURE 9 Architecture of the proposed EDNN.

$$E(x,y) = -\sum_{i=1}^{I}\sum_{j=1}^{J} Q_{ij} f_{si} y_j - \sum_{i=1}^{I} \alpha_i f_{si} - \sum_{j=1}^{J} \beta_j y_j. \quad (15)$$

Anywhere, Q_{ij} reflects the observable unit's symmetrical interaction term f_{si} as well as the secret unit y_j, α,β is the bias term, I,J which seems to be the total number of units in the system, including visible as well as hidden. Defining the gradient of a training vector's predictive relevance in terms of weight is abnormally simple. That there were in an RBM, there are no direct influences between hidden units; obtaining an impartial sample of them is extremely simple $\left(f_{si}, y_j\right)_{data}$

$$\rho\left(y_j = 1 \,|\, f_{si}\right) = \zeta\left(\sum_{i=1}^{I} Q_{ij} f_{si} + \alpha_j\right). \quad (16)$$

Anywhere $\zeta(x)$ is the logistic sigmoid function $\frac{1}{(1+\exp(x))}, f_{si}, h_j$ which is the unbiased sample.

Step 2: Offered units that are visible and those that are hidden are both updated at the same time. The stochastic steepest ascent throughout the log possibility of the training data leads to a considerably simpler learning method

$$W_{ij} = \theta\left(f_{si} y_j\right)_{data} \left(f_{si} y_j\right)_{reconstruction}. \quad (17)$$

Anywhere, W_{ij} represents the updated weight gained as a result of weight changes in the buried layer; Whenever the RBM is activated, trained a dissimilar RBM is to be able to "stacked" to create a multilayer model on top of it The final layer of the previously trained layers is used as a data source for the revolutionary RBM. The finished deep network resources are currently being fine-tuned.

3.4.4 | Fine-tuning phase

The algorithm is fine-tuned using the backpropagation algorithm. An output layer is meant to evaluate the network efficiency somewhere at the peak of the DNN. In addition, the training dataset is skilled till the ideal weight for optimum performance is achieved. Finally, the ENDNN classifier classifies the breast cancer images into normal or lesion.

4 | RESULTS AND DISCUSSION

Evaluate the overall consequence of the proposed method in this section based on FP and deep learning using breast cancer histopathological image classification. In this section, the method introduced in MATLAB applies to a system with 6-GB-RAM and an Intel I-7 processor. The accuracy and performance of the method were evaluated at 2.6 GHz.

4.1 | Dataset description

The summary of our dataset is given in Table 2. And then the description of histo-pathological images is given in Table 3. The dataset is described in [34].

4.2 | Performance metrics

The effectiveness of the suggested strategy is assessed using a set of performance metrics. To evaluate our proposed methodology for efficient categorization, we need to compute various assessment metric values. Three indicators are used to assess the effectiveness of our presented design accuracy, precision, recall, F-measure, true positive rate (TPR), and false positive rate (FPR). The equations below show how to demonstrate these assessment indicators.

4.2.1 | Precision

Precision is defined as the ratio of the number of normal images identified to the estimated number of regular and abnormal people lesion images detected, as given in Equation (18)

$$P = \frac{TP}{TP + FP}. \quad (18)$$

4.2.2 | Recall

Recall is defined as the proportion of normal people to the total number of people images detected as a percentage of the total amount of photos available in the dataset (Equation (19))

TABLE 2 Summary of our dataset.

Dataset	Normal	Benign	In situ carcinoma	Invasive carcinoma	Total
Initial	55	69	63	62	249
Extended	299	1106	1066	1300	3771
Overall	354	1175	1129	1362	4020

TABLE 3 Description of pathological images in our dataset.

Color model	Red green blue
Size	2048 × 1536 pixels
Memory space	3–20 MB (approx.)
Type of label	Image-wise

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{19}$$

4.2.3 | F-measure

The harmonic mean of recalls and precision metrics, which is given in the equation, is the F-measure (Equation (20))

$$F = \frac{2PR}{P + R}, \tag{20}$$

where TP is the true positive, FP is the false positive, and FN is the false negative.

4.2.4 | True positive rate (TPR)

The TPR is the fraction of positive experiment outcomes that are considered

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{21}$$

4.2.5 | False positive rate (FPR)

The number of inaccurate positive predictions divided by the total number of negatives yields the FPR. 1—Specificity is another way to calculate it

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}. \tag{22}$$

4.3 | Confusion matrix

Confusion distribution is a tool often used to assess the effectiveness of the algorithm on even a set of test data for which the true values have been determined. The confusion matrix is straightforward, however, the vocabulary used to describe it can be perplexing. In a nutshell, the confusion matrix is based on the recommended accuracy. The confusion matrix tables and their values are represented as follows in Table 4.

TABLE 4 Confusion matrix of the proposed method with existing methods.

(a) Proposed **Accuracy 98.81%**		
Benign	98.67% 592	1.33% 8
Malignant	1.17% 7	98.83% 593
	Benign	Malignant
Target class		
(b) CNN		
Accuracy 95.41%		
Benign	95.83% 575	4.17% 25
Malignant	5.00% 30	95.00% 570
	Benign	Malignant
Target class		
(c) Decision tree		
Accuracy 91.6%		
Benign	92.00% 552	8.00% 48
Malignant	8.83% 53	91.17% 547
	Benign	Malignant
Target class		
(d) KNN		
Accuracy 90.5%		
Benign	90.67% 544	9.33% 56
Malignant	9.67% 58	90.33% 542
	Benign	Malignant
Target class		
(e) SVM		
Accuracy 89.2%		
Benign	89.67% 538	10.33% 62
Malignant	11.17% 67	88.83% 533
	Benign	Malignant
Target class		

4.4 | Sample outputs

The following Figures 10 and 11 specifies the sample outcomes of amplitude and phase as well as zoomed-in patch and fpm image. Here the amplitude and phase of an image are viewed. Then the patch image and the extracted image are obtained.

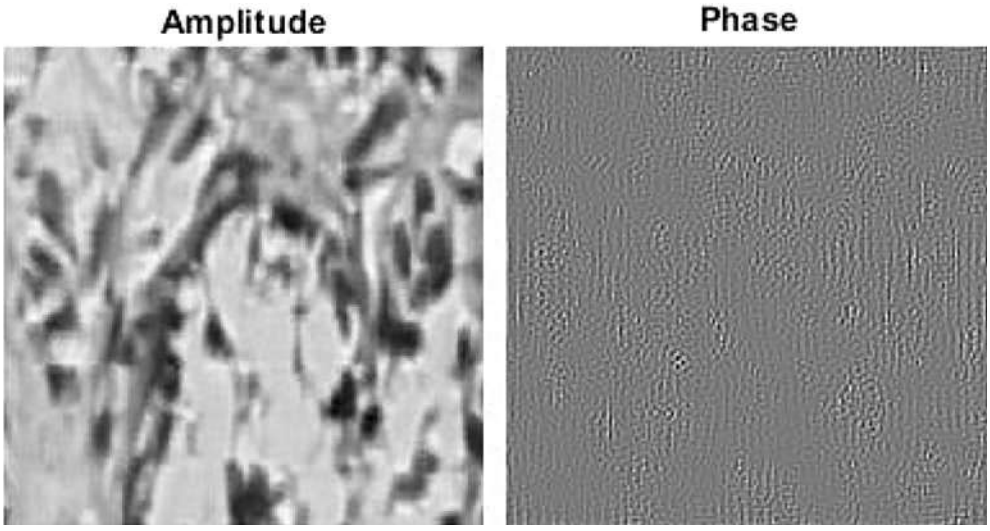


FIGURE 10 Sample outcomes of amplitude and phase

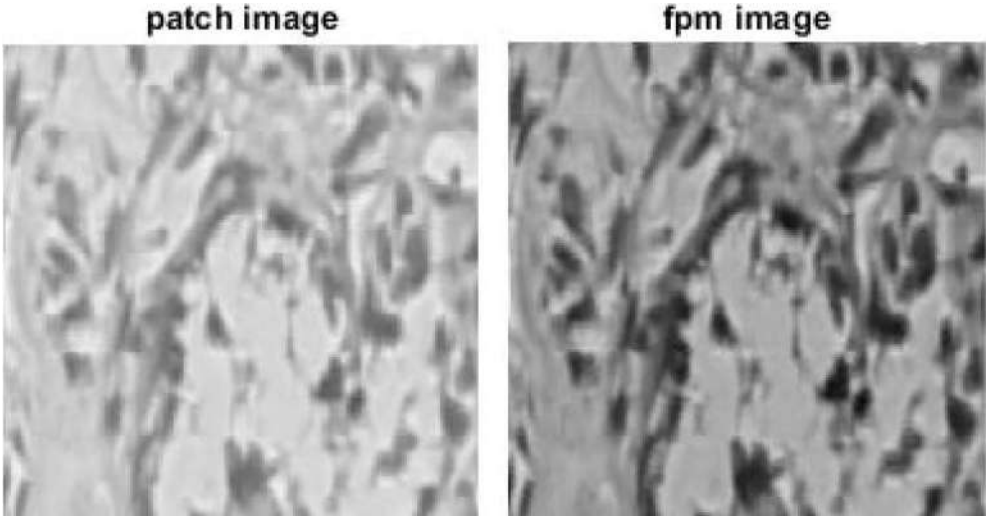


FIGURE 11 Sample outcomes of a zoomed-in patch image.

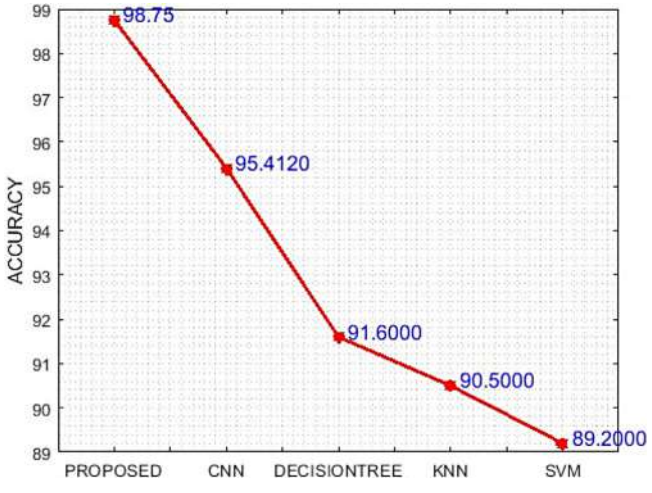


FIGURE 12 Performance and comparative analysis of accuracy.

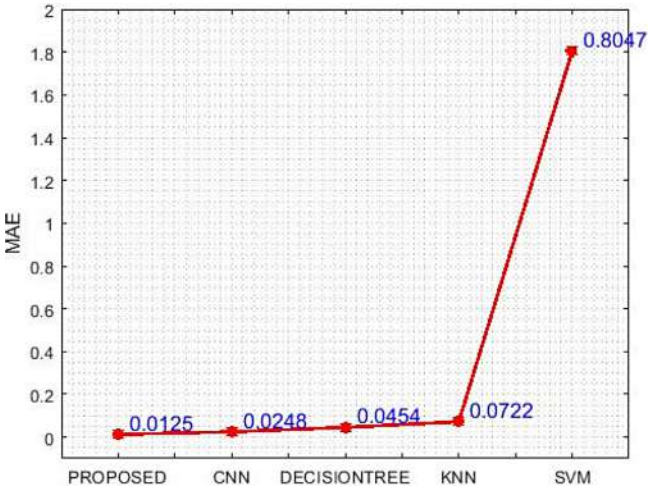


FIGURE 13 Performance and comparative analysis of MAE.

4.5 | Experimental results and analysis

Figure 1 depicts the approach's effectiveness 8–11 which is graphically represented as follows;

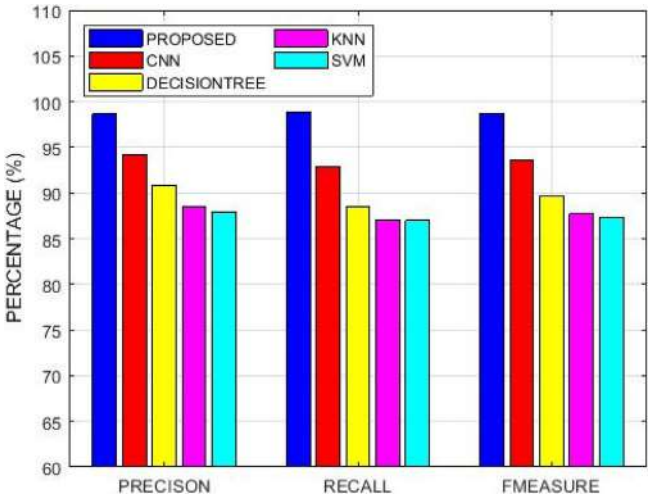


FIGURE 14 Performance and comparative analysis of precision, recall, F-measure.

Figure 12 depicts the approach's effectiveness. The accuracy plot's performance is depicted in the diagram above. When compared to other methodologies, the accuracy gradually improves, according to the analysis. Figure 9 shows the highest level of precision 98.75 for using the proposed. Here, our methodology is compared with existing CNN [28], decision tree [26], KNN [27] as well as SVM [26]. Comparing these other existing approaches proposed to obtain the maximum accuracy and produce higher outcomes.

The above Figure 13 shows the performance and comparative analysis of mean absolute error. From Figure 13 we obtain the maximum MAE of 0.0125 for using the proposed. Here, our methodology is compared with existing CNN, decision tree, KNN as well as SVM. As per the analysis, a comprehensive understanding of our suggested approach to achieving the maximum mean absolute error rates compared with other methods from the result.

The above Figure 14 shows the performance and comparative analysis of precision, recall, and F-measure plots. From Figure 14 we obtain the maximum precision of 98.66666667, recall of 98.83138564, and F-measure of 98.74895746 for using the proposed. Here, our methodology is compared with existing CNN, decision tree, KNN as well

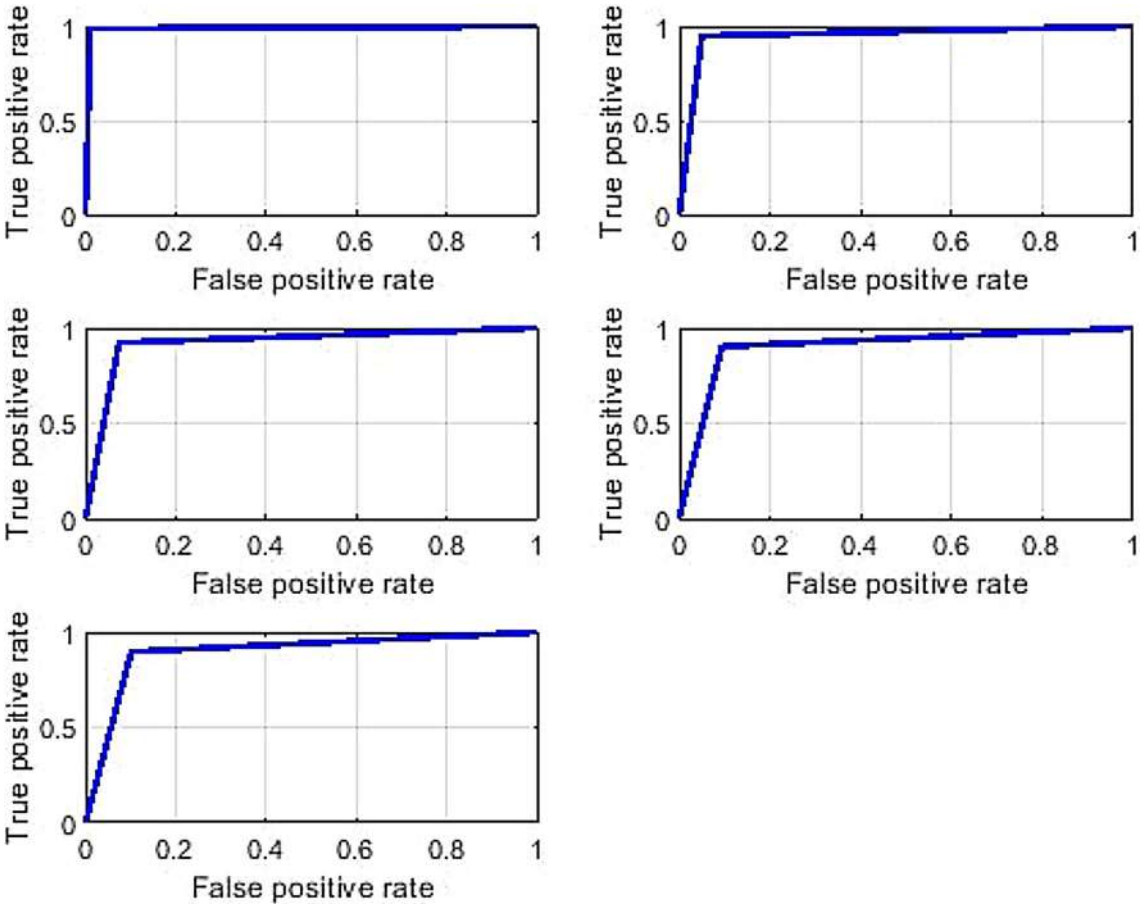


FIGURE 15 Performance analysis of TPR, FPR.

as SVM. As per the analysis, we have a comprehensive understanding of our suggested approach to achieving the maximum precision, recall, and f-measure rates compared with other methods from the result.

FPR is estimated as the ratio among the number of undesirable actions incorrectly considered as FP as well as the entire number of authentic undesirable actions. When analyzing Figure 15 proposed obtains the maximum TPR. Existing techniques are not identifying the TPRs clearly but the proposed correctly identifies the TPRs which are specified in the above graphical representation.

In Figure 16 the loss function for the proposed with the existing methods is given. Here the proposed algorithm has less loss compared to the existing CNN, SVM, KNN, and decision tree. The loss function is calculated in the classification stage to prove the classification algorithm for the proposed is better than the existing one achieved.

In Table 5 the comparison of existing methods with the proposed in terms of accuracy, precision, recall, F-measure, and mean absolute error is determined. The proposed methods of accuracy precision, recall, F-measure and mean absolute error are 98.75, 98.66667, 98.83139, 98.74896, and 0.0125. Compared to the existing methods the proposed algorithm provides better results.

4.6 | Comparison analysis with existing papers

The comparison of related methods and classification with the proposed in terms of accuracy, precision, recall, F-measure, and mean absolute error is described in

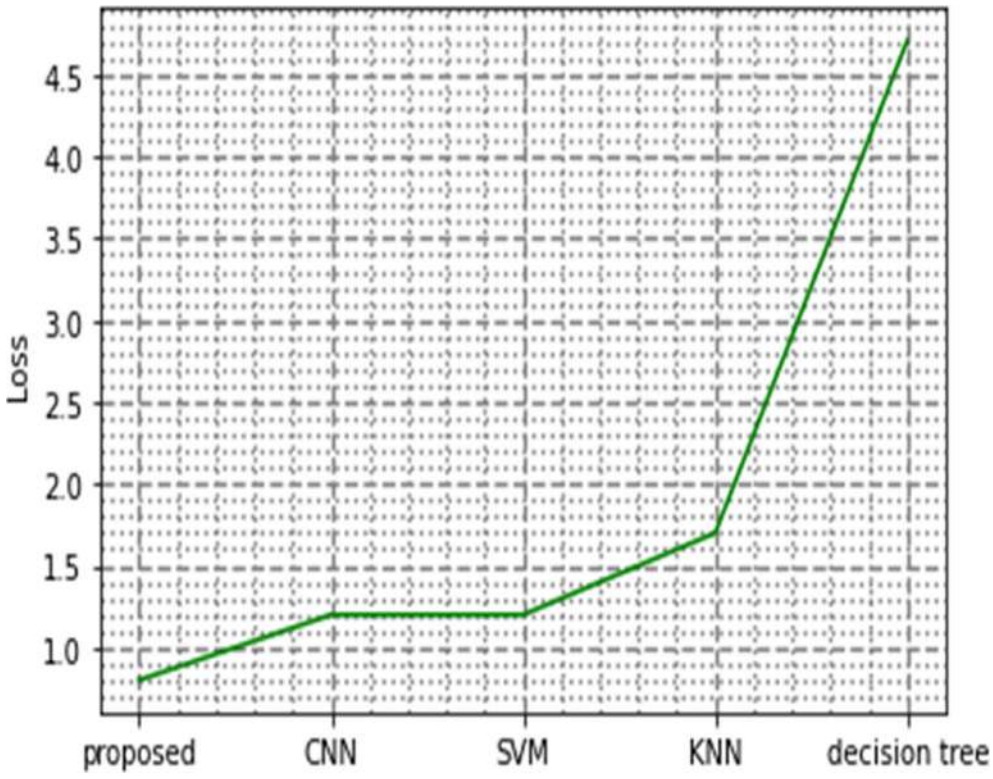


FIGURE 16 Loss function for the proposed with the existing methods.

TABLE 5 Comparison of existing methods with the proposed in terms of accuracy, precision, recall, F-measure, and mean absolute error.

Methods	Accuracy	Precision	Recall	F-measure	Mean absolute error
Proposed	98.75	98.66667	98.83139	98.74896	0.0125
CNN	95.412	94.227	92.846	93.53	0.0248
Decision tree	91.6	90.784	88.562	89.66	0.04536
KNN	90.5	88.5	87	87.7	0.07224
SVM	90.5	87.9	87	87.4	0.80475

TABLE 6 Comparison analysis of related methods and classification.

Paper	Method	Year	Accuracy	Precision	Recall	F-measure	Mean absolute error
Edson et al. [21].	Textural features and content-based image retrieval (CBIR)	2020	97.5%	97.0%	-	-	-
Ala'a El-Nabawy et al. [22]	METABRIC breast cancer subtype classification	2020	95.65	95.82	9565	0.9557	-
Bruno C et al. [29]	Deep convolutional neural network	2021	-	89.57	94.20	-	-
Nouman Ahmad [30]	Transfer learning-based approach	2021	-	97.8	97.58	97.67	-
Hela Elmannai et al. [31]	Deep learning models	2021	97.29%	-	-	-	-
Yao Guo et al. [32]	Hybrid convolutional neural network	2018	87.5%	-	-	-	-
Puspanjali Mohapatra et al. [33]	Deep learning	2019	89%	-	-	-	-
Proposed	ENDNN	98.75	98.66667	98.83139	98.74896	0.0125	98.75

Table 6. Here the related papers to the proposed method are described in [21, 22, 29–33].

5 | CONCLUSION

In this work, the FP and deep learning using breast cancer histopathological image classification are presented. This paper initially extracts the image in high-resolution computer-generated integral holograms using the FP method. Here, we have to recover the high-resolution hologram through an iterative retrieval with FP constraints. Therefore, the feature extraction process includes entropy, geometrical features, and textural features. And then the entropy-based normalization is used to optimize the features. After that, the classification process of the proposed ENDNN classifier classifies the breast cancer images into normal or lesion. The proposed model gives high classification accuracy with optimal features and achieves optimal performance metrics compared to other existing CNN, decision trees, KNN, and SVM methods in terms of accuracy, precision, recall, F-measure, and mean.

FUNDING INFORMATION
In this research article has not been funded by anyone.

CONFLICT OF INTEREST STATEMENT
All authors do not have any conflict of interest.

ORCID
Leena Thomas https://orcid.org/0000-0003-4511-6421

REFERENCE

[1] S. R. Lakhani, I. O. Ellis, S. Schnitt, *International Agency for Research on Cancer*, WHO Press, Lyon **2012**.
[2] M. Veta, J. P. W. Pluim, P. J. van Diest, *IEEE Trans. Biomed. Eng.* **2014**, *2*, 1400.
[3] H. Chen, Q. Dou, X. Wang, *30th AAAI Conference on Artificial Intelligence*, Vol. 30, AAAI Press, Phoenix, Arizona **2016**, p. 1160.
[4] D. Wang, A. Khoslam, R. Gargeya. Deep learning for identifying metastatic breast cancer. arXiv preprint:1606.05718 **2016**.
[5] Y. Huang, H. Zheng, C. Liu, *IEEE J. Biomed. Health Inform.* **2017**, *21*, 1625.
[6] M. Gour, S. Jain, K. T. Sunil, *Int. J Imaging Syst. Technol.* **2020**, *30*, 621.
[7] J. Angel Arul Jothi, V. M. A. Rajam,A survey on automated cancer diagnosis from histopathology images. *Artif. Intell. Rev.* **2017**, *48*, 81. https://doi.org/10.1007/s10462-016-9494-6
[8] B. Du, Q. Qi, H. Zheng, Y. Huang, X. Ding, *International Conference on Artificial Neural Networks*, Springer, Cham **2018** Ch. 4, p. 109.
[9] S. Angara, M. Robinson and P. Guillén-Rondon, *Convolutional Neural Networks for Breast Cancer Histopathological Image Classification*, 4th International Conference on Big Data and Information Analytics (BigDIA), Houston, TX, **2018**, p. 1–6. https://doi.org/10.1109/BigDIA.2018.8632800
[10] Y. Jiang, L. Chen, H. Zhang, X. Xiao, *PLoS One* **2019**, *14*, e0214587.
[11] G. Mahesh, J. Sweta, T. Sunil, *Int. J. Imaging Syst. Teachnol.* **2020**, *30*, 621.
[12] V. Gupta and A. Bhavsar, *Partially-Independent Framework for Breast Cancer Histopathological Image Classification*, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, **2019**, pp. 1123–1130, https://doi.org/10.1109/CVPRW.2019.00146
[13] A. Kumar, S. K. Singh, S. Saxena, K. Lakshmanan, A. K. Sangaiah, H. Chauhan, S. Shrivastava, R. K. Singh, *Inf. Sci.* **2020**, *508*, 405.

[14] C. Zhu, F. Z. Song, Y. Wang, H. H. Dong, Y. Guo, J. Liu, *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 198.
[15] D. Lichtblau, C. Stoean, *PLoS One* **2019**, *14*, e0209274.
[16] T. Dai, J. Cai, Y. Zhang, S. -T. Xia and L. Zhang, *Second-Order Attention Network for Single Image Super-Resolution*, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, **2019**, pp. 11057–11066, https://doi.org/10.1109/CVPR.2019.01132
[17] T. Y. Lin, S. Maji, *British Machine Vision Conference (BMVC)*, Hammamet, Tunisia, **2017**, p. 1.
[18] D. Bardou, K. Zhang, S. M. Ahmad, *IEEE Access* **2018**, *6*, 24680.
[19] Y. Benhammou, S. Tabik, B. Achchab, F. Herrera, *Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications*, Association for Computing Machinery, New York, NY, **2018**, p. 47.
[20] Z. Y. Han, B. Z. Wei, Y. J. Zheng, Y. L. Yin, S. J. Li, *Sci. Rep.* **2017**, *7*, 4172.
[21] E. D. Carvalho, O. C. Antonio Filho, R. R. Silva, F. H. Araujo, J. O. Diniz, A. C. Silva, A. C. Paiva, M. Gattass, *Artif. Intell. Med.* **2020** May, *1*, 101845.
[22] N. El-Bendary, N. A. Belal, *Appl. Soft Comput.* **2020** Jun, *1*, 106238.
[23] Y. Jiang, L. Chen, H. Zhang, X. Xiao, *Journal of Physics: Conference Series* **2019**, *1302*, 032018.
[24] A. Pan, C. Zuo, B. Yao, *Rep. Prog. Phys.* **2020**, *83*, 096101.
[25] Y. Xiao, S. Wei, R. Zhou, *Emerging Digital Micromirror Device Based Systems and Applications XIII*, Vol. 11698, International Society for Optics and Photonics, **2021**, 116980D. https://doi.org/10.1117/12.2579176
[26] V. K. Lam, T. C. Nguyen, V. Bui, B. M. Chung, L.-C. Chang, G. Nehmetallah, C. B. Raub, *J. Biomed. Opt.* **2020**, *25*, 026002.
[27] L. Strbkova, D. Zicha, P. Vesely, R. Chmelik, *J. Biomed. Opt.* **2017**, *22*, 086008.
[28] N. Thanh, Y. Xue, Y. Li, L. Tian, G. Nehmetallah, Deep learning approach to Fourier ptychographic microscopy. *Opt. Express* **2018**, *26*(20), 26470–26484. https://doi.org/10.1364/OE.26.026470
[29] B. C. da Silva, R. T. Gregorio, R. J. Ferrari, *Comput. Biol. Med.* **2021**, *129*, 104133.
[30] N. Ahmad, S. Asghar, S. A. Gillani, *Vis. Comput.* **2022**, *38*, 2751.
[31] H. Elmannai, M. Hamdi, A. AlGarni, *Int. J. Comput. Intell. Syst.* **2021**, *14*, 1003.
[32] Y. Guo, H. Dong, F. Song, C. Zhu, J. Liu, *Image Analysis and Recognition: 15th International Conference, ICIAR*, Springer International Publishing, **2018**, Póvoa de Varzim, Portugal, June 27–29, 2018, p. 827.
[33] P. Mohapatra, B. Panda, S. Swain, *Int. J. Innov. Technol. Exploring Eng.* **2019**, *8*, 2024.
[34] R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang, Y. Liu, X. Rao, C. Zheng, F. Zhang, Breast cancer histopathological image classification using a hybrid deep neural network. *Methods* **2020**, *173*, 52–60. https://doi.org/10.1016/j.ymeth.2019.06.014

How to cite this article: L. Thomas, M. K. Sheeja, *J. Biophotonics* **2023**, e202300194. https://doi.org/10.1002/jbio.202300194