

ITC 2/53 Information Technology and Control Vol. 53 / No. 2 / 2024 pp. 355-371 DOI 10.5755/j01.itc.53.2.34525	Efficient Guided Grad-CAM Tuned Patch Neural Network for Accurate Anomaly Detection in Full Images	
	Received 2023/07/01	Accepted after revision 2024/02/08
	HOW TO CITE: Rajkumar, R., Shanthi, D., Manivannan, K. (2024). Efficient Guided Grad-CAM Tuned Patch Neural Network for Accurate Anomaly Detection in Full Images. <i>Information Technology and Control</i> , 53(2), 355-371. https://doi.org/10.5755/j01.itc.53.2.34525	

Efficient Guided Grad-CAM Tuned Patch Neural Network for Accurate Anomaly Detection in Full Images

R. Rajkumar

Dept. of CSE, Chettinad College of Engineering & Technology, Karur - 639114, Tamil Nadu, India

D. Shanthi

Dept. of CSE, PSNA College of Engineering & Technology, Dindigul - 624005, Tamil Nadu, India

K. Manivannan

School of Computer Science and Engineering, Jain (Deemed to be) University, Bengaluru, Karnataka, India

Corresponding author: rraj कुमारres@outlook.com

Deep learning-based anomaly detection in images has recently gained popularity as an investigative field with many global submissions. To simplify complex data analysis, researchers in the deep learning subfield of machine learning employ Artificial Neural Networks (ANNs) with many hidden layers. Finding data occurrences that significantly differ from generalizable to most data sets is the primary goal of anomaly detection. Many medical imaging applications use convolutional neural networks (CNNs) to examine anomalies automatically. While CNN structures are reliable feature extractors, they encounter challenges when simultaneously classifying and segmenting spots that need removal from scans. We suggest a separate and integration system to solve these issues, separated into two distinct departments: classification and segmentation. Initially, many network architectures are taught independently for each abnormality, and these networks' main components are combined. A shared component of the branched structure functions for all abnormalities. The final structure has two branches: one has distinct sub-networks, each intended to classify a particular abnormality, and the other for segmenting various abnormalities. Deep CNNs training directly on high-resolution images necessitate input layer image compression, which results in the loss of information necessary for detecting medical abnormalities. A guided Grad-CAM (GCAM) tuned patch neural network is applied to full-size images for anomaly localization. Therefore, the suggested approach merges the pre-trained deep CNNs with class activation mappings and area suggestion systems to construct abnormality sensors and then fine-tunes the CNNs on picture patches, focusing on medical abnormalities instead of training on whole images. A mammogram data set was used to test the deep patch classifier, which had a 99% overall classification accuracy. A Brain tumor image data set was used to test the integrated detector's ability to detect abnormalities, and it did so with an average precision of 0.99.

KEYWORDS: Abnormality detection, Artificial Neural Networks, CNN, Deep learning, integrated structure, segmentation, GCAM.

1. Introduction

The detection of anomalies is a significant machine-learning challenge. Rather than assuming a static and closed system, as with most current machine learning methods, this study explores how machine learning models may manage unknown and unpredictable input in an open and changing context. Learning systems designed for anomaly detection typically assume an open environment and use that knowledge to deduce the unexpected, i.e., non-typical, out-of-the-ordinary patterns. To find unusual or new ways in data, anomaly detection methods typically first, With the average data at hand, draw out trends, describe them, and model them. Visual or image anomaly detection is used when the data being examined is an image.

Medical photos with poor contrast and excessive noise make CADe of abnormalities a problematic research topic despite its clinical importance. In particular, one of the challenges specific to medical image analysis is that abnormalities tend to manifest in tiny local regions within a high-resolution image. In mammography, for instance, benign calcification is typically quite extensive, whereas suspicious calcification is generally relatively small and requires the study of magnified images for characterization [10]. It is well known that the efficiency and accuracy of traditional machine learning approaches, which rely on constructing sliding window-based detectors, need to be revised [6, 15]. Additionally, they employ labor-intensive and fallible manual feature engineering. New methods for pinpointing abnormalities in medical images have become available thanks to recent advancements in deep learning. At the ImageNet, deep CNNs outperformed humans on picture categorization tasks [25]. However, it remains challenging due to a lack of adequate training data. Transfer learning has been implemented to solve this issue and prevent over-fitting [38]. Acquired deep features from large-scale labeled picture datasets, such as ImageNet, are then utilized for feature extraction on a target data set with an inadequate count of training photos.

Anomaly identification, also known as outlier detection [26] seeks to identify cases that deviate significantly from the norm that are out of the ordinary or unexpected. Recently, anomaly detection in images has become a popular research area due to its wide range of potential applications, from video surveil-

lance to medicine [35] [32]. Anomalies can occur for several reasons, including data errors or noises, but they can also reveal a hitherto unseen process. As a result, detecting anomalies is an essential endeavor, especially in medical image processing. Because deep neural networks have become so popular and produced remarkable results in many different contexts, many researchers have begun using them to spot anomalies in images. By looking at each pixel, it can also handle complex features like ROIs [22], [4]. Deep learning-based anomaly detection has become general. It has been functional for many tasks, with technologies becoming increasingly commonplace in the medical sector [32], [29], [5], [37], [33]. This is so because deep learning gets around the problem of data imbalance, which can lead to a positive-case bias. We conclude that anomaly detection is preferable to binary classification [33] because more negative medical pictures outnumber good ones from five to one.

This work offers a deep learning method for detecting local abnormalities in training deep patch CNNs and merging them with CAM and RPN [24] to recognize regions in medical pictures. When deep CNNs are trained on high-resolution images, the input layer must perform image compression, which results in the loss of information necessary for medical anomaly detection. Consequently, our method fine-tunes the deep CNNs on image patches focused on medical irregularities rather than training on full-size medical images. The deep patch CNNs are educated to distinguish between the predetermined medical abnormalities.

The research questions rely on the following:

- 1 How does the use of guided Grad-CAM (GCAM) tuned patch neural networks impact the localization of anomalies in high-resolution medical images, and what is the trade-off between accuracy and computational efficiency in this context?
- 2 What are the specific challenges and limitations of training deep convolutional neural networks (CNNs) directly on high-resolution medical images, and how does image compression affect the detection of medical abnormalities?

In this study, we provide a new technique for separating and categorizing medical anomalies, which can be used to solve the issues above. The proposed split

neural network uses one branch to segment images while the other uses classification. The availability of similar low-level properties across several anomalies allows the proposed network's central node to be created to be identical in both splits so that they may share computing functions. Following are some ways in which the proposed method contributes:

- 1 We introduced straightforward network architecture that can be used for segmenting and classifying medical image anomalies.
- 2 This research presented unified network whose primary nodes share features helpful in segmenting and classifying various abnormalities.
- 3 The novel Guided Grad-CAM tuned patch neural network is then applied to full-size images for anomaly localization.
- 4 we use patch-based CNN classifiers for the full images for abnormality localization.

The remaining sections of this paper are structured as follows. Multiple categorization schemes for anomalies converse in Section 2. Section 3 describes the GCAM-equipped network proposed for spotting numerous anomalies. The methods and outcomes of the research are discussed in Section 4. Section 5 contains the conversation, followed by the closing thoughts in Section 6.

2. Related Works

In computer vision and image analysis, abnormality detection using deep learning is a hot topic. Multiple studies have proposed deep learning-based methods for anomaly detection across a wide range of fields and industries, from medical imaging and industrial inspection to surveillance and beyond.

In the past few years, CNNs have made great strides toward automating the derivation of generic characteristics from WCE pictures. Jia et al. [12] are one of the earliest publications on CNN-based WCE image classification. They implemented a 5-layer convolutional neural network CNN to extract features from WCE images and detect bleeding. Since then, various CNN architectures for detecting anomalies have been proposed [7]. To learn the features of WCE images, Segui et al. [30] used three parallel CNNs and the Hessian and Laplacian of the images in addition to the original RGB

inputs. Finally, features from all three CNNs are combined before being sent to fully connected layers for image classification. Iakovidis et al. [11] have also used a 5-layer CNN to detect and pinpoint anomalies. By applying this algorithm to the feature maps, we can locate exciting features that may be found in out-of-the-ordinary areas. Goel et al. [8] used a CNN with two forks, one serving as the network's backbone and the other performing a dilated Conv operation that preserves image resolution. The primary CNN uses multiple scales to extract features. The expanded Conv branch broadens the system's sensitivity and facilitates the separation of key features. To get the dominant global features, [2] we finally combine the elements from the backbone CNN and the dilated CNN.

Performance evaluation measures were also calculated using three distinct pattern classifiers: Support Vector Machine [28], [13], [34], Linear Discriminate Analysis, and Bayes Linear Classifier. The other work contributed is dimensionally reducing the obtained features with Principal Component Analysis and then classifying them with the Support Vector Machine technique. The abnormal breast area can be pinpointed with the help of a mammogram. The abnormality in mammograms was detected using a variety of pattern recognition algorithms and decision-making systems. Textural and geometric features are used for feature extraction. A correlation-based Feature Selection (CFS) scheme selects features for further classification into masses or non-masses. Breast disorders such as tumors and microcalcification may be detected using methods for image processing such as Shape derivation and Boundary of group or lesion.

One indicator of breast abnormality detection is the presence of a mass at the periphery of the mass. Breast masses that are round and well-defined on a mammogram are typically benign, while those with an uncertain border indicate cancer. The Digital database for screening mammography (DDSM) is the source for these images, and Verma et al. [3] have developed a computer-aided diagnosis scheme to establish an outline for computerized mammograms that prioritizes a neural hereditary calculation with an acquired precision of 85%. According to Ahuja et al. [23], benign neoplasms are well-marginated and have regular shapes, while malignant neoplasms have indistinct and irregular borders that become more speculative over time. Masses and benign glandular

tissue slightly differ in X-ray attenuation, so a lack of contrast and blurriness characterize their visual presentation. Micro-calcifications, or mammographic “bright spots,” are tiny calcium deposits.

Reconstruction-based or image generation and completion methods are another option for training on one-class data. The benefit of these methods—often based on neural networks—is that they can detect anomalies pixel-by-pixel. After an input has been processed through a bottleneck layer, for instance, auto-encoders can reconstruct it [36], [14]. To avoid having to learn a meaningful representation, they compress the images. Denoising variants solves this issue by introducing noise into the input image, forcing the DAN to reconstruct the original query [19]. However, because such manipulation is typically performed at the pixel level, the DAN does not need to learn much semantic data. A recent image creation method relies on a GAN taught to produce average results.

Yuan et al. [39] have implemented a resemblance restraint in the loss function when training a DenseNet to handle intra-class variations. For low intra-class

variation and high inter-class variations, Yuan et al. [40] proposed a densely connected CNN with unbalanced discriminant loss and category-sensitive loss. To reduce intra-class variance in polyp detection, Guo et al. [9] developed a triple ANET construction using angular contrastive loss and an attention mechanism. Many existing studies concentrate on intra-class variations, a diverse range of anomalies challenging in achieving accuracy and reducing error rate. The popular systems described in the literature focus on intra-class variation caused by a single abnormality type. Unique mechanisms need to be designed to capture the underlying characteristics of various anomalies to deal with and classify them as abnormal. This research presents a parallel CNN framework that utilizes a unique meta-feature retrieval approach to differentiate among multiple anomalies using the underlying statistical trends in the feature maps. Combined with traditional feature maps, these meta-features boosted the accuracy of the classifier. Table 1 shows comparison of conventional model with its unique characteristics.

Table 1

Comparison of characteristics on existing work

Reference	Method	Key Features	Applications	Unique Contributions
Jia et al. [12]	5-layer CNN	Feature extraction from WCE images	WCE image classification	Early use of CNN in WCE image classification
Segui et al. [30]	Three parallel CNNs	Uses Hessian, Laplacian, and RGB inputs	WCE images	Combination of multiple CNNs and image processing techniques
Iakovidis et al. [11]	5-layer CNN	Anomaly detection and localization	Threat classification	Localization of anomalies in feature maps
Goel et al. [8]	CNN with two forks	Dilated Conv operation, multi-scale features	Image processing	Expanded Conv branch for key feature separation
Various [2, 28, 13, 34]	Various methods	SVM, LDA, BLC classifiers	Mammogram analysis	Comparative analysis using different classifiers
Verma et al. [3]	CAD scheme	Neural hereditary calculation	Computerized mammograms	Outline for computerized mammograms
Ahuja et al. [23]	Not specified	Analysis of neoplasm characteristics	Mammogram analysis	Differentiation between benign and malignant neoplasms
Yuan et al. [39, 40]	DenseNet, CNN	Resemblance restraint, unbalanced discriminant loss	Polyp detection	Addressing intra-class variations
Guo et al. [9]	Triple ANET	Angular contrastive loss, attention mechanism	Polyp detection	Reducing intra-class variance

3. Materials and Methods

Deep CNN pre-trained on normal pictures is fine-tuned using extracted image patches centered on medical aberrations. This allows the CNNs to acquire image attributes for local abnormalities.

3.1. Data Collection

The DDSM is a free resource for researchers interested in using digital mammograms to study breast cancer. Computer-aided detection and diagnosis algorithms for breast cancer extensively use a dataset developed by the University of South Florida's Digital Image and Communications in Medicine (DICOM) Standard. The DDSM dataset's Curated Breast Imaging Subset (CBIS) is a subset of the entire dataset that has been processed and cleaned up for research convenience. More than three thousand mammography photos with annotations and information are included, including screening and diagnostic mammography.

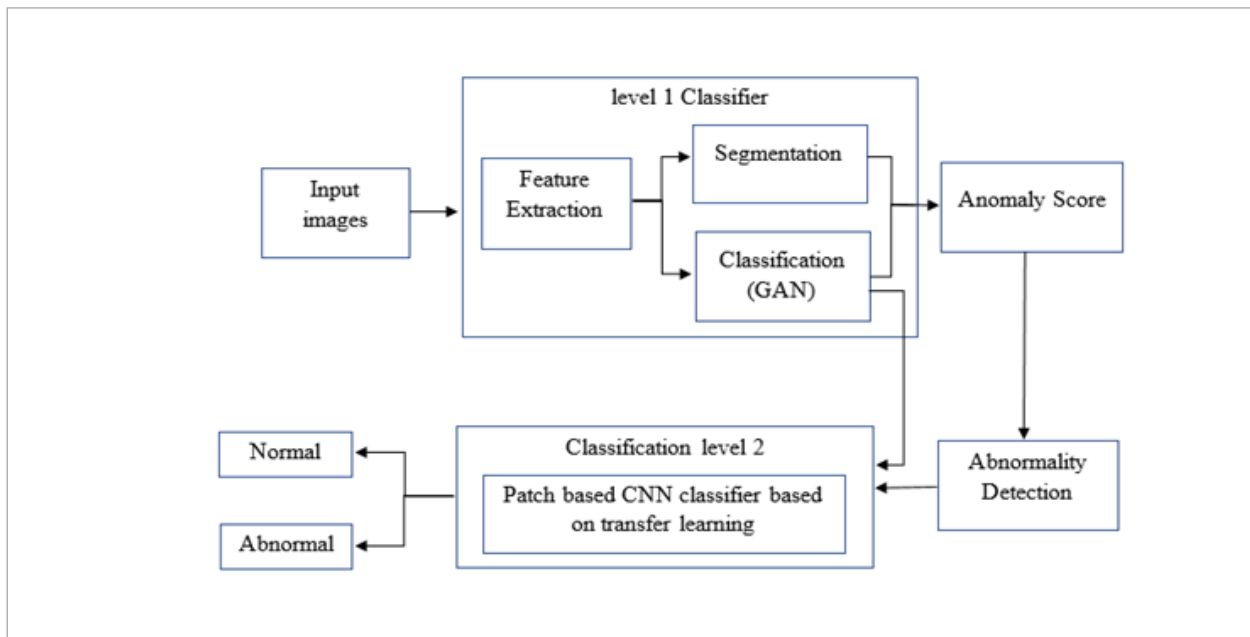
Full-field digital mammography (FFDM) and digital breast tomosynthesis (DBT) pictures are also part of the CBIS collection. Digital breast tomosynthesis (DBT) pictures are 3D digital images that enable more

precise identification of breast lesions than flat-field digital mammography (FFDM) images. Breast lesions are annotated in the CBIS dataset with bounding boxes, benign/malignant labels, and BI-RADS (Breast Imaging Reporting and Data System) scores to help classify mammogram findings according to the level of suspicion for breast cancer. The CBIS dataset may be used to train and evaluate CAD algorithms to identify and diagnose breast cancer. The dataset may be downloaded at no cost and used in various studies. Our studies use the publicly accessible brain MRI dataset from the Kaggle repository. Out of 253 MRI scans, 155 are considered "normal," leaving 98 scans that show abnormalities.

The proposed detector incorporates patch classifiers where only image patches can be obtained during training. Performs a complete medical image abnormality map computation once trained. The object detector includes the patch classifiers. The use of whole pictures and patch locations in different contexts can be obtained during training. Once the sensor has been trained, it will report irregularities as rectangles with likelihoods attached to them. Both methods are proven effective and improve performance when finding local abnormalities in mammograms and ultrasound liver images. Figure 1 depicts the proposed method.

Figure 1

Outline of the Modeling Proposal



3.2. Level 1 Classifier

Each anomaly can have its network trained explicitly on its dataset of abnormal images. Each trained network successfully segments the training-set abnormality. As a result, the architecture of each network can be fine-tuned to address a specific abnormality. There is no coordination between the different constituents of the networks and no use of shared characteristics of the abnormalities, all contributing to the overall complexity of the network. Although these systems are appropriately particular, they need more job duplication. Not only are they useless for classification, but they are only suitable for segmentation. When segmenting and classifying multiple abnormalities, it is essential to focus on individual monsters as required for segmentation to maintain the broader understanding needed for classification.

Common abnormalities in medical images have been linked to severe illnesses. Segmentation and abnormal region classification are both necessary for accurate diagnosis of significant diseases. Segmentation and categorization of aberrant areas are essential for a precise assessment. The primary objective of the first classification stage is to identify aberrant pictures. The normal samples are used to train the model, which then attempts to produce consistent reconstructions of reg-

ular inputs with low anomaly scores and inconsistent reconstructions of aberrant pictures with high anomaly scores. The suggested model's design is depicted in Figure 2. It involves incorporating semi-supervised learning via a memory-augmented deep auto-encoder.

To learn representations in which the input and the output are identical, an AE makes use of feedforward neural networks. The sole components of a standard AE are the encoding and decoding units. The encoder is symbolized by the function $f_{en} \odot$, $A \subseteq R^{(d_a \times 1)}$, $B \subseteq R^{(d_b \times 1)}$ where A and B are the domains of the original data and the latent manifold, respectively. This function maps the original image a onto a latent vector b of lower dimensionality ($d_b < d_a$) by use of a pre-determined mapping.

$$b = f_{en}(a, \theta_{en}), \quad (1)$$

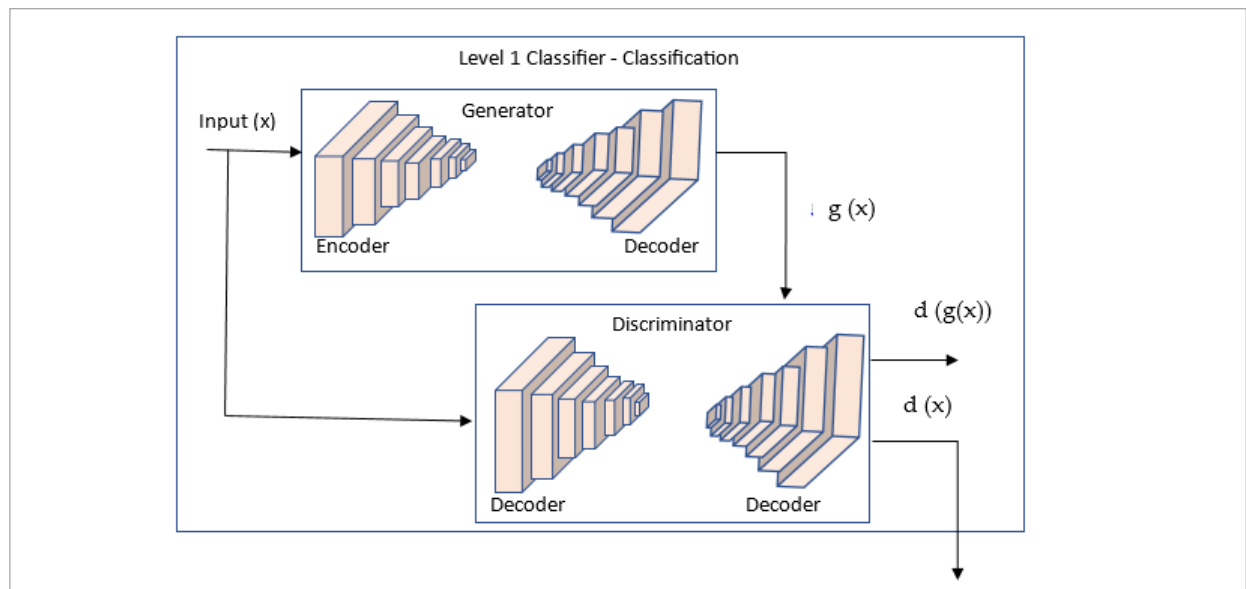
where e is the collection of encoder parameters. For producing the recreated picture from b , the decoder is specified by the function $f_{de} \odot$, which maps b back to its innovative space.

$$\hat{a} = f_{de}(b, \theta_{de}) = f_{de}(f_{en}(a, \theta_{en}), \theta_{de}), \quad (2)$$

where θ_{de} is the decoder's settings variable. The goal of AE is to use the latent-space representation b to successfully compress the input data a and then use

Figure 2

The level 1 classifier architecture – GAN structure



the encoded version to reconstruct a image \hat{a} that is similar to a .

$$E_{ende}(\theta_{en}, \theta_{de}) = \frac{1}{G} \sum_{i=1}^G \|\hat{l} - l\|^2, \tag{3}$$

where $l = a$. A storage record containing a matrix is the basis of the system that stores information of $\Omega \subseteq R^{M_\Omega \times d_b}$, where M_Ω characterizes the number of prototypical routes. The memory component, given a latent variable b , will produce a recovered latent representation in accordance with,

$$\hat{a} = \Omega^T w = \sum_{n=1}^{M_\Omega} w_n \omega_n, \tag{4}$$

where n is the n th memory element in the database and ω is a weight vector where $w_n > 0$ and $\sum_{n=1}^{M_\Omega} \omega_n = 1$ to describe the importance of each prototype vector Ω in generating b . The memory components and the latent query vector b are used to inform the construction of the weight vector w .

$$w_n = \frac{\exp(d(b, w_n))}{\sum_{n=1}^{M_\Omega} \exp(d(b, w_n))}. \tag{5}$$

There are two parts to the GAN: the generator G and the discriminator D. D seeks to categorize the data as

real or artificial, whereas G is meant to develop synthetic samples that are comparable to the real data. As G and D compete, GAN is able to optimize the objective function and so understand the fundamental information distribution, allowing for the generation of high-quality artificial samples.

The subsequent anomaly score is used during the inference phase to determine the level of abnormality present in a fresh query signal a :

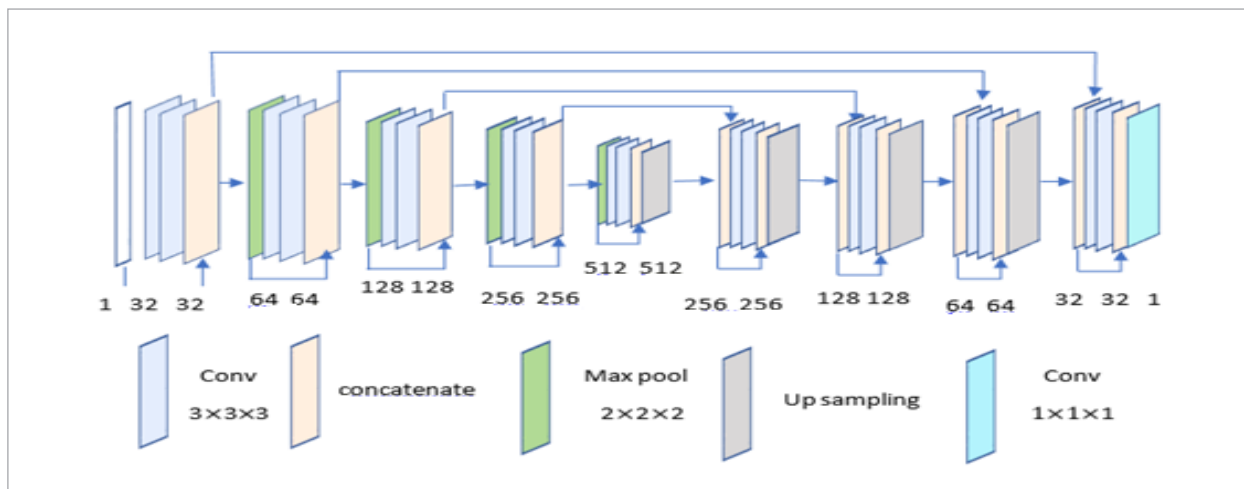
$$s(a) = \|a - G(a)\|. \tag{6}$$

This measures how far off G 's pattern extraction from normal photos is from the actual image. When processing anomalous images, the system is expected to yield high abnormality scores. However, if $s(a)$ is small, it means that the test picture a has patterns that are typical of normal images, and so will be classed as normal by the suggested technique

3.3. Segmentation

The U-Net was designed as a CNN for use in the medical imaging field. The network has an encoder and a decoder with skip links to keep spatial data intact. Class imbalance, limited sample numbers, and complicated anatomical systems are only some of the difficulties this architecture is built to overcome. For segmenting medical images, Figure 3 exposes the technical details of the U-Net architecture.

Figure 3
The UNet architecture for the image segmentation



Algorithm 1: Segmenting using U-Net**Input:** medical image dataset I_d **Output:** Segmented medical images**Step 1:** Pre-processing:

- a Load the medical image dataset I_d .
- b Normalize the intensity values of I_d to a standard range, typically $[0, 1]$.
- c augmentation of data is performed including rotations, scalings, and elastic deformations to increase dataset robustness.

Step 2: Initialize a U-Net model with the following architecture**Step 3:** Encoder operation

- i. input the image with one channel.
- ii. Apply a sequence of convolutional blocks, each consisting of:
 - _ Two $3 \times 3 \times 3$ convolutions with ReLU activation.
 - _ A $2 \times 2 \times 2$ max pooling operation to reduce spatial dimensions.
 - _ Doubling the number of filters starting with 32 and progressing through 64, 128, 256.
- iii. Performing batch normalization after each convolutional layer to stabilize learning and improve convergence.

Step 4: Bottleneck

In the bottom of the U-Net without pooling image patch is executed using two $3 \times 3 \times 3$ convolutions followed by ReLU activation and batch normalization.

Step 5: Decoder operation

- i. Up-sample the image using $2 \times 2 \times 2$ transposed convolutions to increase the spatial dimensions.
- ii. Concatenate the upsampled output with the corresponding encoder feature maps (skip connections).
- iii. two $3 \times 3 \times 3$ convolutions with ReLU activation is used for image segmentation.
- iv. Halve the number of filters after each up-sampling step (starting from 512 and reducing through 256, 128, 64).

Step 6: Output Layer:

$1 \times 1 \times 1$ convolution performed to map the feature maps to the number of segmentation labels.

Step 7: Execute loss function using softmax cross-entropy for multi-class voxel classification and Dice loss for segmentation accuracy. optimize the features using Adam with an initial learning rate and evaluate overall measures

Step 8: Model Training:

- a Train the U-Net model using the preprocessed and augmented dataset I_d .
- b validates the data still stopping and best model checkpointing.
- c learning rate is adjusted and performance is evaluated

Step 9: Test set is executed for the evaluating the performance. Use the performance metrics defined earlier for quantitative assessment.

To extract information from an input picture, the encoder network passes it over a chain of convolutional layers. The construction is built from stacked units of two 3×3 convolutional layers, a rectified linear unit (ReLU) activation function, and a down sampling layer of 2×2 max pooling. Each pooling operation doubles the number of filters used in each convolutional layer. Using batch normalization after each convolutional layer improves the resolution, and overfitting is mitigated.

The encoder network's output is sent into the decoder network, where it is upsampled to the input size and combined with the encoder network's feature maps through skip links. There are two 3×3 convolutional layers with ReLU activations, and one 2×2 transposed convolutional layer for Upsampling, all of which are repeated in blocks throughout the design. After every block, the number of filters in the transposed convolutional layer is halved.

Slices taken at various angles comprise the three-dimensional pictures used in biomedical imaging. The study of biomedical images requires the processing of massive volumes of information. Segmentation-labeled data may be challenging to annotate since computers can only show data in two-dimensional slices. Thus, typical 2D image models often need more performance and context loss while processing 3D images. The solution is a 3D U-Net based on the existing U-Net framework but with a reducing encoder component for full-image analysis and a continuously expanding

decoder section for high-resolution segmentation generation. There are numerous structural similarities between 2D and 3D U-Nets; however, 3D U-Nets use 3D convolution, 3D pooling, and 3D upsampling in place of all 2D processes. To avoid slowdowns in the network, batch normalization (BN) is implemented.

There are four parsing steps per layer in the encoding path, the same as the regular U-Net, and the same number in the decoding road. The layer structure consists of two $3 \times 3 \times 3$ convolutions, a 2-step-sized corrected linear unit (ReLU), and a $2 \times 2 \times 2$ maximum pooling layer. Each ReLU active layer in the synthesis route is sandwiched between two $3 \times 3 \times 3$ convolutions and two $2 \times 2 \times 2$ higher convolutions with two steps in each dimension. The skip links in the equal-resolution attribute map give high-resolution information for decoding. The number of output channels is reduced to the size of the labels, which in this case is 3, thanks to the $1 \times 1 \times 1$ convolution in the last layer. The total number of parameters for the structure is 19069955.

Before training, the data and ground truth labelers are subjected to smooth, dense deformation fields in addition to the rotation, scaling, and Gray value increase. Therefore, B-spline interpolation is applied to random vectors drawn from a uniform spreading with a standard deviation of 4 and a grid spacing of 32 voxels in each direction. We evaluate what the network produces to the ground truth label using the softmax plus weighted cross-entropy loss to strike a better balance among the effect of microscopic blood vessels and backdrop voxels on the loss. This comprehensive learning approach may employ fully or partially automated techniques to segment 3D targets from limited explanations. This network can generalize well from a limited amount of labeled data thanks to its well-designed structure and data improvement features. Reasonable pictures could be produced, its preprocessing approach could be rationalized, and the system construction could be expanded to accommodate any size 3D data collection if suitable rigid alteration and modest elastic distortion submissions were used.

3.4. Level 2 Classifier

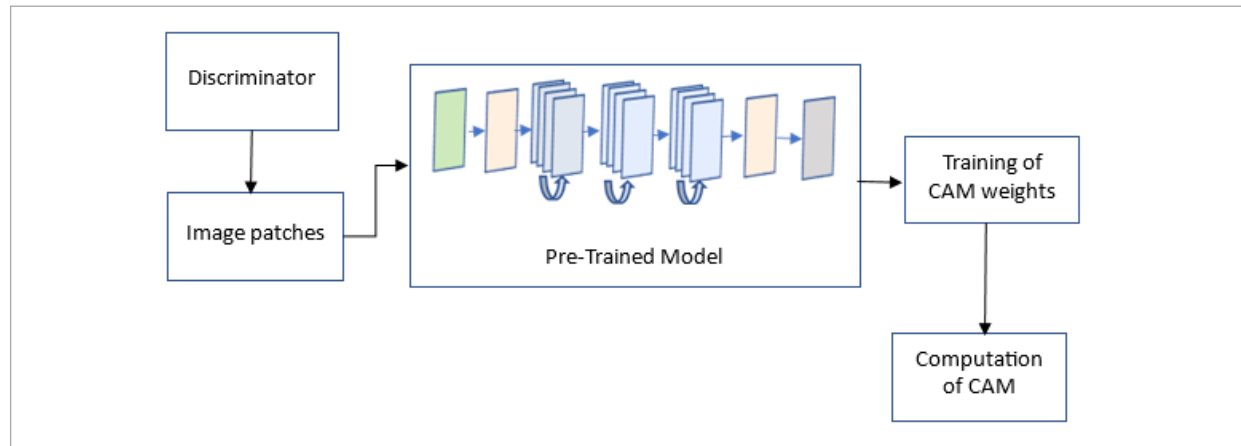
Numerous medical procedures, such as endoscopy, dermoscopy, funduscopy, etc., must classify and segment various disorders. Multiple abnormalities associated with severe illnesses can be seen in diag-

nostic imaging. Current years have seen a rise in the use of CNN for automatic abnormality identification in medical imaging applications because of its efficiency and accuracy. Because mammography abnormalities tend to be present in minimal, local regions of a high-resolution picture, computer-aided identification cannot be approached as a straightforward image classification problem. An abnormality area in a standard mammography scan of 3000×4600 pixels (width by height) is just 80×80 pixels (or even less). Resizing photos to 224×224 (in pixels) at the input layer is necessary for training high-quality deep CNNs. However, this results in a significant reduction in information concerning medical problems. To solve this problem, we first adopt deep CNNs trained on ROIs from trimmed pictures to full mammography images. The deep patch CNN is trained similarly to a window operator, but it may be used on the entire picture for the first time because the convolutional weights are shared in deep CNNs. The maps are calculated using class activation mapping (CAM) [42]. CAM isolates class-related features in a picture. Because of this, classifiers may be reused for localization even if location-specific training data is unavailable. It also shows that CNNs have an inherent capacity for paying attention.

Breast cancer is prevalent among women, and early detection is crucial for improving patient outcomes. Digital mammography is an imaging technique that captures X-ray images of the breast, which computer algorithms can analyze to assist radiologists in identifying potential cancerous regions [18], [41]. Meta-heuristic optimization algorithms are techniques inspired by natural processes such as evolution, swarm intelligence, and simulated annealing. These algorithms search for optimal solutions in complex, high-dimensional problem spaces [16]. Figure 4 depicts the proposed method. Starting with transfer learning helps prevent over-fitting. picture patches centered on medical abnormalities are retrieved from the discriminator's output picture and used to fine-tune deep CNNs, allowing the CNNs to learn image characteristics for local abnormalities. The suggested detector combines patch classifiers with GCAMs for use in situations when only picture patches are available during training. The GCAM-based sensor, once trained, generates an anomaly map for the entire medical picture.

Figure 4

Level 2 classifier with GCAM



In step 1, we use picture patches from calcification and bulk instances to train a two-class classifier utilizing transfer learning and cutting-edge deep CNN architectures. The data flow is based on AlexNet's architecture for illustration purposes. VGGnet, GoogLeNet, and ResNet are the other models used for this purpose. After the last convolution layer in the deep patch, CNN is removed, a global average pooling layer and a fully linked layer are added. The new model requires retraining to acquire the output layer CAM weights $w_k (i = 1, 2, \dots, n)$. After feeding a complete mammography into the input layer, the ResNet is the only one of the four deep CNN designs that is ready to compute CAM. The output of the last convolutional layer is the feature maps, or Tensors, indicated by $F_k (i = 1, 2, \dots, n)$. By reflecting Convolutional maps of features and their weights from the output layer, we can determine which parts of the picture are most significant.

$$CAM(c) = \sum_{k=1}^m w_k F_k. \quad (7)$$

The Guided Grad-CAM tuned patch neural network is applied to full-size mammograms for anomaly localization. Traditional methods employ a sliding window for the classifier to scan the entire image, which is time-consuming and inefficient. In contrast, our process allows anomaly localization with a single forward computing run. When the full-size mammography is fed into the patch classifier, and CAM is computed, a heat map of anomalies is generated.

Guided Grad-CAM (Gradient-weighted Class Activation Mapping) is a method used in deep learning to visualize the sections of an input image that contribute the most to the prediction made by CNN. It combines the concepts of Grad-CAM and guided backpropagation to generate more accurate and detailed visualizations.

The equation for Guided Grad-CAM (GCAM) can be defined as follows:

The gradient of the desired class score relative to the output feature maps from the convolutional layer is computed.

$$\alpha_c^k = \frac{1}{Z} \sum_m \sum_n \frac{\partial y^c}{\partial A_{mn}^k}. \quad (8)$$

Here, y^c represents the target class score, A_{mn}^k represents the activation at position (m, n) in the k -th feature map of the final convolutional layer, and Z is the spatial size of the feature maps.

Calculate the weights (α_c^k) by taking the global average pooling of the gradients obtained in the previous step:

$$\alpha_c^k = \text{ReLU}(\alpha_c^k). \quad (9)$$

The ReLU activation is applied to discard any negative gradients. Calculate the Grad-CAM activation map by taking the weighted sum of the feature maps using the obtained weights:

$$L_G^k = \text{ReLU}(\sum_k \alpha_c^k A^k). \quad (10)$$

L_G^c represents the Grad-CAM activation map for the target class c , and A^k characterizes the k -th feature map of the final convolutional layer. Calculate the guided backpropagation gradient for each pixel:

$$g_{bp} = \frac{\partial y^c}{\partial X} \tag{11}$$

Here, X represents the input image. Calculate the guided Grad-CAM by element-wise multiplying the Grad-CAM activation map with the guided backpropagation gradient

$$L_G^c = L_G^c \odot g_{bp} \tag{12}$$

The symbol \odot denotes element-wise multiplication. The resulting Guided Grad-CAM L_G^c represents the guided Grad-CAM activation map for the target class c , which identifies the input picture regions to be highlighted that contribute the most to the prediction made by the CNN.

4. Experimental Setup

Linux was used for the experiments, and they were run using the Keras framework. Table 2 displays the experiment-specific values for our model’s hyper-parameters. The validation data set was used to determine these hyper-parameters. The validation dataset also served to establish the normal/abnormal cut-off.

Table 2
Hyperparameter tuning for the proposed model

Hyperparameter	Value
Batch size	64,128,512
Epochs	50
Learning rate	0.001,0.0001,0.00001
Optimizer	Adam
Activation	ReLU, softmax, Sigmoid

5. Result and Discussion

This section shows how well the suggested technique determines whether a picture is aberrant (DR). The obtained performance metrics are then evaluated against competing, cutting-edge strategies. When

applied to the proposed image abnormality detection method, the effectiveness metrics of accuracy, sensitivity, specificity, Receiver Operating Characteristic (ROC), and Area Under the Curve (AUC) validate the excellence of the findings and outcome from the actual case.

The input picture was scaled down to 64x64 pixels so that the model could be systematically evaluated. In the first round of testing, we use a 64x64 grid with a 0.1, 0.01, 0.001, and 0.0001 epoch learning rate. After the database was split into testing and training sets, the generated algorithm was put through its paces regarding performance metrics. Performance metrics of 99.12% accuracy, 98.21% sensitivity, and 98.36% specificity on the CBSI-DDSM dataset and 99.35% accuracy, 98.10% sensitivity, and 98.87% specificity were obtained with a learning rate of 0.0001 as in Table 3 and Figure 5. Table 4 shows the standard performance measure of precision, recall and F1-score.

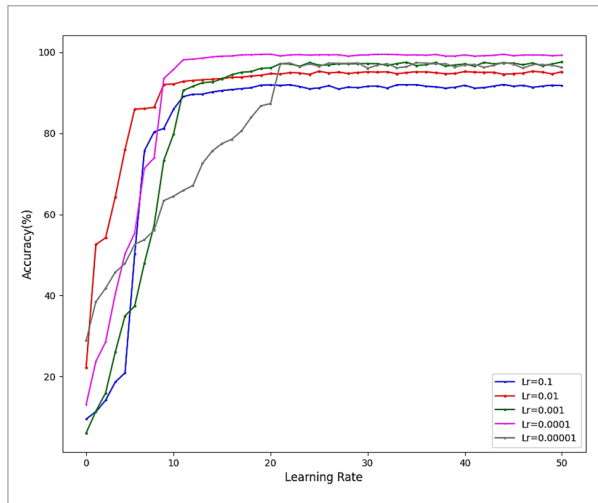
We construct class activation mapping using the fine-tuned ResNet to pinpoint anomalies. The decision to utilize ResNet to compute GCAM was based on the fact that it does not require any further training. Without sacrificing generalizability, we load a single calcification-class complete mammography into ResNet and calculate the GCAM. It is challenging to train classifiers on calcification and mass instances because of mammography pictures’ poor contrast and noise. Input pictures to deep neural networks can

Table 3
Performance measures in terms of Learning rate

Dataset	Learning Rate	Accuracy	Sensitivity	Specificity
CBSI-DDSM	0.1	90.78%	90.12%	90.41%
	0.01	95.32%	94.75%	94.15%
	0.001	97.58%	96.78%	96.23%
	0.0001	99.12%	98.21%	98.36%
	0.00001	98.32%	97.12%	97.54%
Brain MRI	0.1	89.62%	89.32%	88.98%
	0.01	95.47%	95.65%	94.85%
	0.001	96.54%	96.32%	95.10%
	0.0001	99.35%	98.10%	98.87%
	0.00001	97.65%	97.21%	96.54%

Figure 5

Accuracy of the DACN model for Alzheimer's Disease Detection

**Table 4**

Standard performance measures using various Learning rate

Dataset	Learning Rate	Precision	Recall	F1-score
CBSI-DDSM	0.1	90.12%	90.12%	90.12%
	0.01	94.75%	94.75%	94.75%
	0.001	96.78%	96.78%	96.78%
	0.0001	98.21%	98.21%	98.21%
	0.00001	97.12%	97.12%	97.12%
Brain MRI	0.1	89.32%	89.32%	89.32%
	0.01	95.65%	95.65%	95.65%
	0.001	96.32%	96.32%	96.32%
	0.0001	98.10%	98.10%	98.10%
	0.00001	97.21%	97.21%	97.21%

only be 224x224 or 227x227 pixels in size. Fine features essential for categorization will likely be lost if mammography pictures are resized to these sizes. To distinguish between calcification and bulk instances, we suggest training classifiers on cropped batches of images and then applying the resulting deep CNN models to full-size mammograms. We effectively apply the patch classifier to the task of pinpointing anomalies in complete mammography images using a method termed GCAM.

The total accuracy indicates how well the training data was classified. However, val_acc is essential for testing the correctness of the hidden data in the test data. As shown in Figure 6, a neural network model performs well when the val_loss and val_acc go up. However, the model has learned for a non-deterministic batch, as shown by the training epoch. The word batch is employed since the validation loss is determined per sample. val_loss is an excellent indicator of the model's performance on unseen data. Overfitting is not an issue if the val_loss is small. Thus, if the model has been taught significantly on the data, an increase in val_loss indicates overfitting. Line charts for the defined central terms depict an overall assessment of the proposed model. Figures 6-7 depict this phenomenon. It is essential to observe that as more epochs pass, the accuracy metrics improve, and the Loss numbers drop. When val_loss goes down, it means the model has been trained correctly. After the outliers have been filtered out, the dataset of images is identified as normal or abnormal using a basic neural network design.

In Figure 8, the projected result is presented as a confusion matrix. Of the total input images, 99% were correctly labeled as expected, and 100% were labeled as abnormal. The suggested approach addresses limits in the same way that any expert system would. The training data plays a crucial role in the model's recital. Because of this, the algorithm flagged the intake of new information—a typical, healthy cardiac image—as suspicious. This may be an issue when the corpus lacks domain-specific examples. The temporal components of the picture are stripped away during the cleaning process since the suggested model only considers spatial features. As a result, features are occasionally misrepresented because of a failure to account for edges and differences in the local binary pattern of the pictures.

Interestingly, only a small amount of training can get the minimal Loss score down to 0.01. This theory helped reduce the input data length and illustrate the latent space model. Most activations also show a sparse hierarchical architecture, mainly when focusing on the spatial aspects. In addition, the encoder reduces the dimensionality of the complex representation before reconstructing the original picture during decoding. The model may get overfit during this adjustment by memorizing the training data.

The suggested approach typically uses half the input variables to regulate how many neurons may be used in the core layer. By doing so, we can guarantee that the model is picking up on the most relevant and valuable aspects of the data we provide it. It is important to emphasize that the model is unsupervised because it makes no use of training labels. The system is da-

ta-dependent since the buried layers' neurons are all driven by the available data. Therefore, the activation function causes various neurons to fire, resulting in a varied network output when the input characteristics change. Our technique had a tremendous overall classification performance but needed to be more accurate in a small percentage of data.

Figure 6

(a) Accuracy and (b) Loss of the proposed model on the dataset CBSI-DDSM

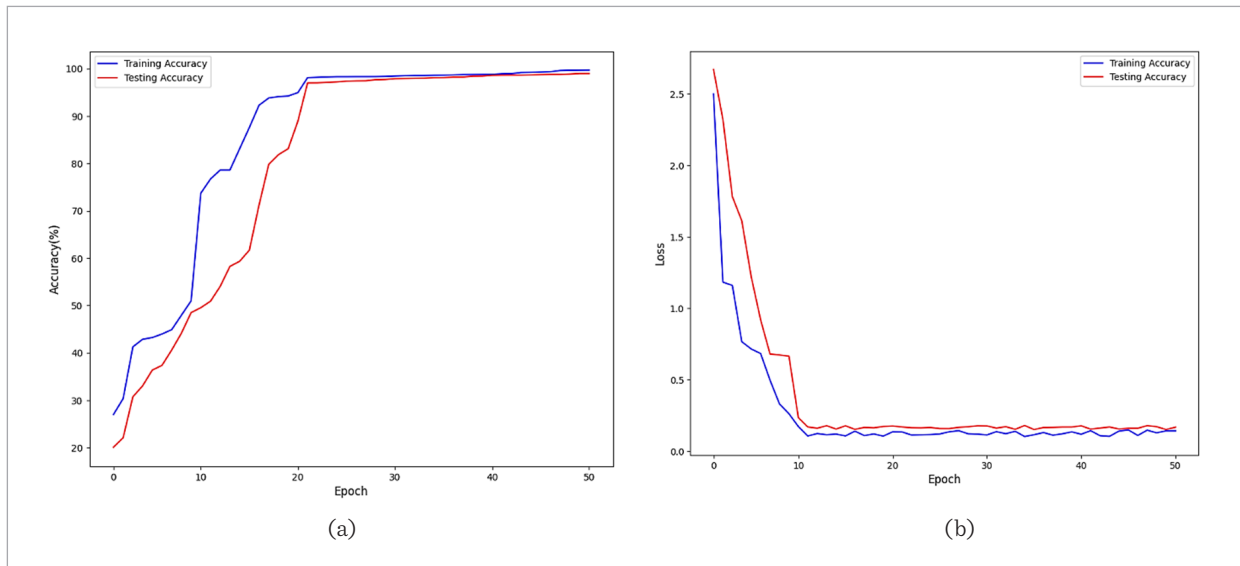


Figure 7

(a) Accuracy and (b) Loss of the proposed model on the dataset Brain MRI

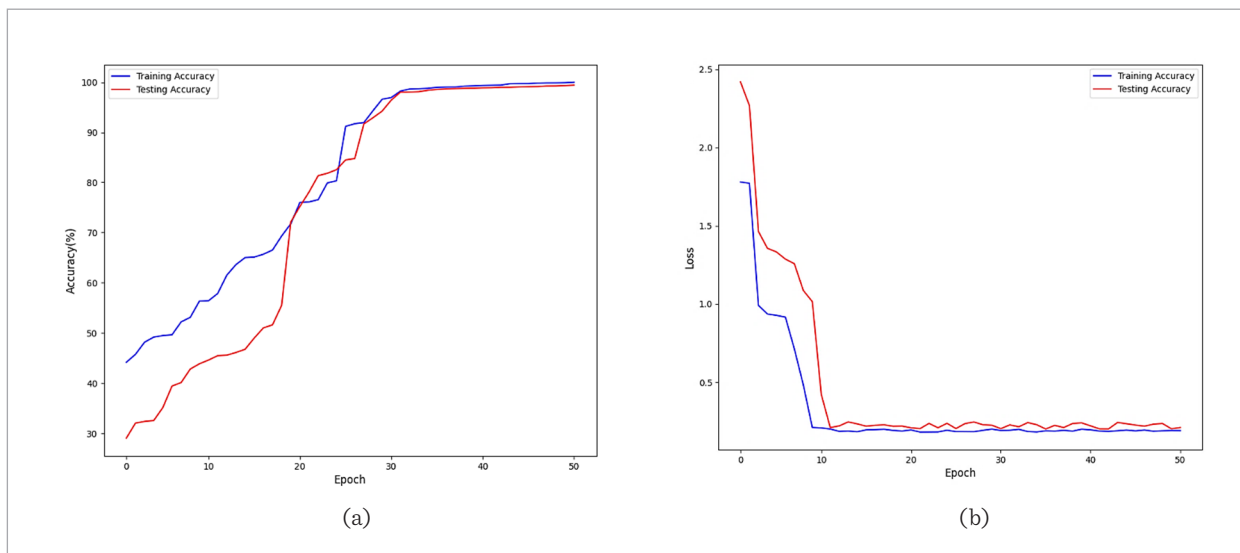
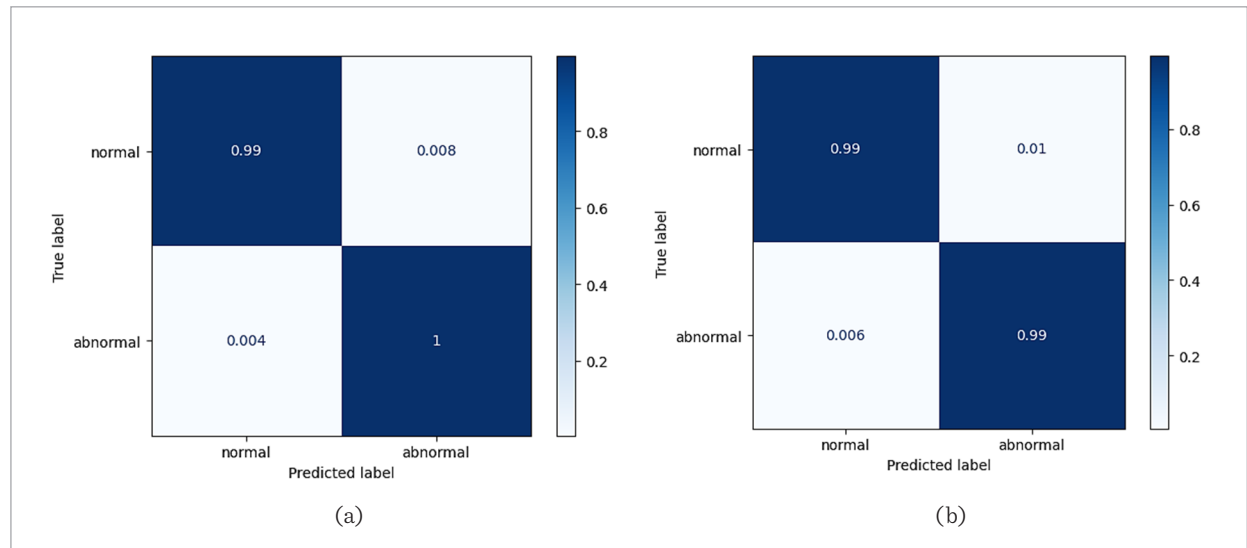


Figure 8

The Confusion matrix of the proposed model on (a) CBSI-DDSM dataset and (b) Brain MRI

**Table 3**

Performance comparison with related work

Related study	Dataset	Accuracy
Montaha et al. [20]	CBSI-DDSM	98%
Zebari [41]	CBSI-DDSM	79.36%
Panambur et al. [21]	CBSI-DDSM	87%
Sarker et al. [27]	CBSI-DDSM	98%
Shen et al. [31]	CBSI-DDSM	98%
Mallick et al. [17]	Kaggle Brain MRI	93%
Alsaif et al. [1]	Kaggle Brain MRI	97%
Proposed methodology	CBSI-DDSM, Kaggle Brain MRI	99%

The table summarizes related studies along with their respective datasets and accuracy rates. Montaha et al. [20] conducted a study using the CBSI-DDSM dataset and achieved an accuracy of 98%. Zebari [41] also utilized the CBSI-DDSM dataset but achieved a slightly lower accuracy of 79.36%. Panambur et al. [21] obtained an accuracy of 87% using the same CBSI-DDSM dataset. Sarker et al. [27] and Shen et al. [31] used the CBSI-DDSM dataset and achieved accu-

racies of 98%. Mallick et al. [20] employed the Kaggle Brain MRI dataset and achieved an accuracy of 93%. Similarly, Alsaif et al. [1] used the Kaggle Brain MRI dataset and achieved an accuracy of 97%. Finally, the proposed methodology utilized both the CBSI-DDSM and Kaggle Brain MRI datasets and achieved an impressive accuracy rate of 99%. Overall, the table highlights various studies, datasets, and their corresponding accuracy rates, providing an overview of the performance of different approaches in the field.

6. Conclusion and Future Work

We offer a method based on deep learning to find anomalies in medical pictures. Transfer learning is used for actual feature learning without over-fitting when available training data is scarce. Deep CNN is taught to recognize certain anomalies in an image. They are used with GCAMs to construct anomaly detectors. Experimental data shows that the integrated strategy is superior to conventional and alternative deep learning strategies. After the database was split into testing and training sets, the generated algorithm was put through its paces regarding performance metrics. Performance metrics

of 99.12% accuracy, 98.21% sensitivity, and 98.36% specificity on the CBSI-DDSM dataset and 99.35% accuracy, 98.10% sensitivity, and 98.87% specificity were obtained with a learning rate 0.0001. The suggested technique has been successfully validated on mammography and ultrasound pictures, indicating tremendous promise for helping clinicians discover local problems in medical imaging. Classifier performance may be measured in terms of accuracy,

f-score, precision, and calculation time and is acquired through a testing procedure that follows the training. The experimental analysis results are superior to those of the current methods.

Shortly, genetic algorithms will be used to improve the deep convolution neural network further. The properties of the sperm pictures enhance feature analysis. An algorithm is being developed to speed up the processing of massive datasets.

References

1. Alsaif, H., Ramzi, G., Badr, A., Tarek, H., Tawfik, G., Alzamil, A., Lamia, B. A Novel Data Augmentation-based Brain Tumor Detection using Convolutional Neural Network. *Applied Sciences*, 2022, 12(8), 3773. <https://doi.org/10.3390/app12083773>
2. Ayyaz, M. I. U., Lali, M., Hussain, H. T., Rauf, B., Alouffi, H., Alyami Wasti, S. Hybrid Deep Learning Model for Endoscopic Lesion Detection and Classification Using Endoscopy Videos. *Diagnostics*, 2022, 12(1), 43. <https://doi.org/10.3390/diagnostics12010043>
3. Brijesh, V., Zakos, J. A Computer-Aided Diagnosis System for Digital Mammograms Based on Fuzzy-Neural and Feature Extraction Techniques. *IEEE Transactions on Information Technology in Biomedicine*, 2001, 5(1). <https://doi.org/10.1109/4233.908389>
4. Chalapathy, R., Chawla, S. Deep Learning for Anomaly Detection: A Survey. *arXiv preprint arXiv:1901.03407*, 2019. <https://doi.org/10.1145/3394486.3406704>
5. Choi, H., Ha, S., Kang, H., Lee, H., Lee, D. S., Initiative, A. D. N. Deep Learning only by Normal Brain Pet Identify Unheralded Brain Anomalies. *EBioMedicine*, 2019, 43, 447-453. <https://doi.org/10.1016/j.ebiom.2019.04.022>
6. Ciresan, D., Giusti, A., Gambardella, L. M., Schmidhuber, J. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. In: Pereira, F., Burges, C. J. C., Bottou, L., Weinberger, K. Q. (eds.) *Advances in Neural Information Processing Systems*, Curran Associates Inc., New York, 2012, 25, 2843-2851.
7. Georgakopoulos, S. V., Iakovidis, D. K., Vasilakakis, M., Plagianakos, V.P., Koulaouzidis, A. Weakly Supervised Convolutional Learning for Detection of Inflammatory Gastrointestinal Lesions. In 2016 IEEE International Conference on Imaging Systems and Techniques (IST). *IEEE*, 2016, 510-514. <https://doi.org/10.1109/IST.2016.7738279>
8. Goel, N., Kaur, S., Gunjan, D., Mahapatra, S. Dilated CNN for Abnormality Detection in Wireless Capsule Endoscopy Images. *Soft Computing*, 2022, 1-17. <https://doi.org/10.1007/s00500-021-06546-y>
9. Guo, X., Yuan, Y. Triple ANET: Adaptive Abnormal-Aware Attention Network for WCE Image Classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, 293-301. https://doi.org/10.1007/978-3-030-32239-7_33
10. Hernández, P. L. A., Estrada, T. T., Pizarro, A. L., Cisternas, M. L. D., Tapia, C. S. Calcificaciones Mamarías: Descripción y Clasificación Según la 5.a Edición Bi-Rads. *Rev. Chilena De Radiol*, 2016, 22(2), 80-91. <https://doi.org/10.1016/j.rchira.2016.06.004>
11. Iakovidis, D. K., Georgakopoulos, M., Vasilakakis, A., Koulaouzidis, Plagianakos, V. P. Detecting and Locating Gastrointestinal Anomalies using Deep Learning and Iterative Cluster Unification. *IEEE Transactions on Medical Imaging*, 2018, 37(10), 2196-2210. <https://doi.org/10.1109/TMI.2018.2837002>
12. Jia, X., Meng, M. Q. H. A Deep Convolutional Neural Network for Bleeding Detection in Wireless Capsule Endoscopy Images. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). *IEEE*, 2016, 639-642. <https://doi.org/10.1109/EMBC.2016.7590783>
13. Jothilakshmi, G. R., Raaza, A., Effective Detection of Mass Abnormalities and Its Classification Using Multi-SVM Classifier with Digital Mammogram Images. *IEEE International Conference on Computer, Communication, and Signal Processing (ICCCSP2017)*. <https://doi.org/10.1109/ICCCSP.2017.7944090>
14. Kholief, E. A., Darwish, S. H., Fors, N. Detection of Steel Surface Defect based on Machine Learning Using Deep Auto-Encoder Network. In *Industrial Engineering and Operations Management*, 2017, 218-229.

15. Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., den Heeten, A., Karssemeijer, N. Large Scale Deep Learning for Computer Aided Detection of Mammographic Lesions. *Medical Image Analysis*, 2016, 35, 303-312. <https://doi.org/10.1016/j.media.2016.07.007>
16. Kurdi, S. Z., Ali, M. H., Jaber, M. M., Saba, T., Rehman, A., Damaševičius, R. Brain Tumor Classification Using Meta-Heuristic Optimized Convolutional Neural Networks. *Journal of Personalized Medicine*, 2023, 13(2). <https://doi.org/10.3390/jpm13020181>
17. Mallick, P. K., Ryu, S. H., Satapathy, S. K., Mishra, S., Nguyen, G. N., Tiwari, P. Brain MRI Image Classification for Cancer Detection Using Deep Wavelet Autoencoder-based Deep Neural Network. *IEEE Access*, 2019, 7, 46278-46287. <https://doi.org/10.1109/ACCESS.2019.2902252>
18. Maqsood, S., Damaševičius, R., Maskeliūnas, R., TT-CNN: A Breast Cancer Detection and Classification Towards Computer-Aided Diagnosis Using Digital Mammography in Early Stages. *Applied Sciences*, 2022, 12(7). <https://doi.org/10.3390/app12073273>
19. Mei, S., Wang, Y., Wen, G. Automatic Fabric Defect Detection with a Multi-Scale Convolutional Denoising Autoencoder Network Model. *Sensors*, 2018, 18(4). <https://doi.org/10.3390/s18041064>
20. Montaha, S., Azam, S., Kalam, A., Rakibul, M., Rafid, H., Ghosh, P., Hasan, Md., Jonkman, M., De Boer, F. BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images. *Biology*, 2021, 10(12), 1347. <https://doi.org/10.3390/biology10121347>
21. Panambur, A., Bhandary, P. M., Maier, A. Effect of Random Histogram Equalization on Breast Calcification Analysis Using Deep Learning. In *Bildverarbeitung für die Medizin 2022: Proceedings, German Workshop on Medical Image Computing, Heidelberg, and Wiesbaden: Springer Fachmedien Wiesbaden, June 26-28, 2022*, 173-178. https://doi.org/10.1007/978-3-658-36932-3_38
22. Pang, G., Shen, C., Cao, L., Hengel, A. V. D. Deep Learning for Anomaly Detection: A Review. *arXiv preprint arXiv:2007.02500*, 2020, 2020.
23. Pavankumar, M., Ahuja, G., R. D., Wakde, D. G. Detection and Diagnosis of Breast Tumors Using Digital Mammogram Images: A Fuzzy Approach. *International Journal of Science and Research (IJSR)*, 2016, 5(12).
24. Ren, S., He, K., Girshick, R., Sun, J. Faster R-CNN: Towards Realtime Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6), 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
25. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., Fei-Fei, L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 2015, 115(3), 211-252. <https://doi.org/10.1007/s11263-015-0816-y>
26. Sarafijanovic, Djukic, N., Davis, J. Fast Distance Based Anomaly Detection in Images Using an Inception-like Auto Encoder. In *International Conference on Discovery Science, Springer, 2019*, 493-508. https://doi.org/10.1007/978-3-030-33778-0_37
27. Sarker, P., Sarker, S., Bebis, G., Tavakkoli, A. Connected UNets++: Mass Segmentation from Whole Mammographic Images. In *Advances in Visual Computing: 17th International Symposium, Cham: Springer International Publishing ISVC 2022, San Diego, CA, USA, October 3-5, 2022*, 419-430. https://doi.org/10.1007/978-3-031-20713-6_32
28. Sarmistha, C., Kanti, M., Bhowmik, A. G., Pal, T., Automated Edge Detection of Breast Masses on Mammograms, 978-1- 5090-2597-8/16, 2016.
29. Seebock, P., Orlando, J. I., Schlegl, T., Waldstein, S. M. Bogunovic, H., Klimescha, S., Langs, G., Schmidt U. Erfurth, Exploiting Epistemic Uncertainty of Anatomy Segmentation for Anomaly Detection in Retinal Oct. *IEEE Transactions on Medical Imaging*, 2019, 39, 1, 87-98. <https://doi.org/10.1109/TMI.2019.2919951>
30. Seguí, S., Drozdal, M., Pascual, G., Radeva, P., Malagelada, C., Azpiroz, F., Vitrià, J. Generic Feature Learning for Wireless Capsule Endoscopy Analysis. *Computers in Biology and Medicine*, 2016, 79, 163-172. <https://doi.org/10.1016/j.combiomed.2016.10.011>
31. Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., Sieh, W. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Scientific Reports*, 2019, 9(1), 12495. <https://doi.org/10.1038/s41598-019-48995-4>
32. Shi, W., Yan, G., Li, Y., Li, H., Liu, T., Sun, C., Wang, G., Zhang, Y., Zou, Y., Wu, D. Fetal Brain Age Estimation and Anomaly Detection using Attention-based Deep Ensembles with Uncertainty. *NeuroImage*, 2020, 223, 117316. <https://doi.org/10.1016/j.neuroimage.2020.117316>

33. Wei, Q., Ren, Y., Hou, R., Shi, B., Lo, J. Y., Carin, L. Anomaly Detection for Medical Images Based on a One-Class Classification. In *Medical Imaging 2018: Computer-Aided Diagnosis*, International Society for Optics and Photonics, 2018, 10575, 105751M. <https://doi.org/10.1117/12.2293408>
34. Wenfeng, H., Dong, J., Guo, Y., Zhang, M., Wang, J. Identification of Masses in Digital Mammogram using an Optimal Set of Features. *IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*, 2011.
35. Xie, X., Wang, C., Chen, S., Shi, G., Zhao, Z. Realtime Illegal Parking Detection System based on Deep Learning. In *Proceedings of the 2017 International Conference on Deep Learning Technologies*, 2017, 23-27. <https://doi.org/10.1145/3094243.3094261>
36. Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., Madabhushi, A. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. *IEEE Transactions on Medical Imaging*, 2016, 35(1), 119-130. <https://doi.org/10.1109/TMI.2015.2458702>
37. Xu, S., Wu, H., Bie, R. Cxnet-m1: Anomaly Detection on Chest X-Rays with Image-based Deep Learning. *IEEE Access*, 2018, 7, 4466-4477. <https://doi.org/10.1109/ACCESS.2018.2885997>
38. Yosinski, J., Clune, J., Bengio, Y., Lipson, H. How are Features in Deep Neural Networks Transferable? In: *Proceedings of the 27th International Conference on Neural Information Processing Systems*, MIT Press, Cambridge, MA, USA, 2014, 2, NIPS'14, 3320-3328.
39. Yuan, Y., Qin, W., Ibragimov, B., Han, B., Xing, L. RI-IS-Densenet: Rotation-Invariant and Image Similarity Constrained Densely Connected Convolutional Network for Polyp Detection. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, 620-628. https://doi.org/10.1007/978-3-030-00934-2_69
40. Yuan, Y., Qin, W., Ibragimov, B., Zhang, G., Han, B., Meng, M. Q. H., Xing, L. Densely Connected Neural Network with Unbalanced Discriminant and Category Sensitive Constraints for Polyp Recognition. *IEEE Transactions on Automation Science and Engineering*, 2019, 17(2), 574-583. <https://doi.org/10.1109/TASE.2019.2936645>
41. Zebari, D. A., Ibrahim, D. A., Zeebaree, D. Q., Mohammed, M. A., Haron, H., Zebari, N.A., Maskeliūnas, R. Breast Cancer Detection using Mammogram Images with Improved Multi-Fractal Dimension Approach and Feature Fusion. *Applied Sciences*, 2021, 11(24). <https://doi.org/10.3390/app112412122->

