ORIGINAL ARTICLE

GSO-CNN-based model for the identification and classification of thyroid nodule in medical USG images

Rajshree Srivastava[1] · Pardeep Kumar[1]

Abstract
Thyroid ultrasonography is one of the widely used techniques for the detection and classification of thyroid nodules. In this paper, grid search optimization (GSO)-based convolutional neural network (CNN), i.e., GSO-CNN model is proposed for thyroid nodule identification and classification. A total of 295 public and 654 collected thyroid USG datasets are considered in this work. The increased datasets size for the proposed model becomes 1770 for the public dataset and 3924 for the collected dataset after applying data augmentation techniques. We experimentally determined the best optimized value using grid search optimization (GSO) technique for learning rate and dropout. The model works in four phases: (i) data collection, (ii) pre-processing, (iii) morphological operation, segmentation and boundary detection and (iv) classification using CNN. The proposed model has achieved an accuracy of 95.30%, sensitivity of 96.66%, specificity of 94.87% and f-measure of 97.20% on the public dataset having 1770 thyroid USG images and an accuracy of 96.02%, sensitivity of 96.70%, specificity of 95% and f-measure of 98.34% on the collected dataset having 3924 thyroid USG images. The proposed model has been compared with popular deep learning techniques like dense neural network (DNN), Alexnet, Resnet-50 and Visual Geometry Group (VGG-16) with and without considering segmentation and boundary detection techniques. The proposed model has shown an improvement of (6.126%, 6.846%), (7.1%, 7.14%), (6.77%, 6.9%) and (7.77%, 8.91%) in terms of accuracy, sensitivity, specificity and f-measure on (dataset -1, dataset-2) against other state of the art models.

Keywords Deep learning · CNN · Active contour · Boundary detection · Augmentation

1 Introduction

Thyroid nodules are one of the most common endocrine malignancies which are characterized by a distinctive appearance. Screening of thyroid nodules is conducted in several ways, among them ultrasonography (USG) is highly recommended. Here, the doctors/radiologist studies the USG images to find the characteristic of benign/malignant thyroid nodules. USG is said to be a popular tool used for clinical practice of medical diagnosis as it is radiation free, low cost, etc. (Zheng et al. 2020). However, owing to human fatigue and habitual behaviour, exact prediction often becomes impractical (Wang et al. 2019). A regular cell could appear like a tumor cell in the case of USG and this is said to be a false positive. This misdiagnosis scenario involves additional testing and medical procedures (Shivhare and Saxena 2022). There is a set of scores based on thyroid nodules characteristics to differentiate benign and malignant thyroid nodules. These can be found in thyroid imaging reporting and data systems (TI-RADS) (Pedraza et al. 2015).

Machine learning (ML) has recently gained a lot of attention due to the explosion of data generation processes in the domain of health care (HealthITAnalytics. 2020). Some of the applications includes heart failure management, medical imaging, clinical decision support, etc. (Awan et al. 2018). There are several ML algorithms that are widely used for the clinical diagnosis of the diseases like support vector machines (SVM), random forest (RF), artificial neural network (ANN), decision tree (DT), logistic regression (LR), etc. (Nagavelli et al. 2022). In the current scenarios, the clinical decision support (CDS) system is being designed to prioritize the prescription checks of the patients. Similarly

✉ Rajshree Srivastava
    rajshree.srivastava27@gmail.com

    Pardeep Kumar
    pardeepkumarkhokhar@gmail.com

1   Department of Computer Science and Engineering, Jaypee University of Information Technology, Dist. Solan, Waknaghat, H.P, India

in the same way, these systems can also be designed for the pharmaceutical company to select research on candidate molecule that works on drug discovery (Lo-Ciganic et al. 2022). Some of the major applications of CDS include early identification of lung cancer and circular ribonucleic acid (RNAs), brain tumor, etc. (Cinaglia et al. 2018). With a recent advancement, the size of the datasets has become too large due to which the traditional data pre-processing and machine learning techniques are not able to handle effectively (Rezvani et al. 2019). Deep learning (DL) algorithms have won many competitions in pattern recognition and disease diagnosis. However, analysing noise-contaminated data sets and complex data, it is very essential to extract relevant information from the data to convert it into a readable form (Wang and Wang 2018). DL techniques are deeper variants of artificial neural network (ANN) with multiple layers that can be linear or non-linear (Wang et al. 2017). Here, each layer is connected with some weights to different hidden layers. This capability of DL leads to learning hierarchical features from different data like, images, audio, video, etc. (Sengupta et al. 2020). Figure 1 shows the architecture of CNN.

In this research work, several deep learning models like deep neural network (DNN), Alexnet, Resnet-50 and VGG-16 are used for the classification of thyroid nodules. DNN is an artificial neural network (ANN) with multiple hidden layers between inputs and outputs layers having nodes in each layer (Mohsen et al. 2018). The hidden layer performs mathematical computation as the data are fed into the input layer. The main advantage of using DNN is integrating the feature and classification tasks to build good decision models (Shaban et al. 2021a). DNN is one of the successful classification models for the diagnosis of disease. Alexnet was coined in 2012 which has shown a great development of CNN (Krizhevsky et al. 2012). In comparison with traditional neural networks (NN), CNN has reduced the number of parameters and complexity and also solves the problem of overfitting. Some of the improvements made in Alexnet was the introduction of de-convolution after convolution operation. Some of the benefits are: (i) overfitting and parameters were reduced, (ii) good results were achieved (Li et al.

2021). It contained eight layers, having five convolutional (Conv) layers, followed by max-pooling layers and three fully connected (FC) layers. Resnet-50 was coined by He Zhang in a 2015 paper entitled "Deep residual learning for image recognition" (He et al. 2016). Resnet-50 is a CNN having 50 deep layers (Boesch 2022). During backpropagation in CNN, the value of gradient decreases significantly known as the vanishing gradient problem (Rastogi 2020). To overcome this issue, Resnet was introduced, making use of skip connection (i.e., adding original input to the output of convolutional block) (Ikechukwu et al. 2021). VGG-16 has 16 layers deep neural network (VGGnet) having 138 million parameters (Shaban et al. 2021b). It was coined by K.Simonyan in 2014 (Pedamkar 2021). This model has 16 layers and uses very few hyperparameters throughout the training process (Llamas et al. 2017). A complete model comprises 12 convolutional (Conv) 2-D layers and four fully connected (FC) layers with the softmax function at the end (Yang et al. 2021). After finding all these advantages of deep learning models, these models are incorporated into the study. The paper is organized as follows: Section 2 provides the related work on the detection and classification of thyroid nodules. Section 3 focuses on methods and materials. Sect. 4 discusses the results and Sect. 5 concludes this manuscript.

2 Related work

The main focus of this section is to throw some light on the various work done so far in the literature on thyroid nodule detection and classification. This will be beneficial to the researchers, practitioners and clinicians to propose a high-performance model for the identification of benign and malignant nodules. Each reported literature is having contributions, limitations/ future work.

Wang et al. (2018) proposed a novel semi-supervised learning method for automatic thyroid nodule classification. In their work, there is a combination of expectation–maximization (EM), i.e., generating a bag of patient's images and instances as nodules of images and CNN to train the model

Fig. 1 Architecture of CNN
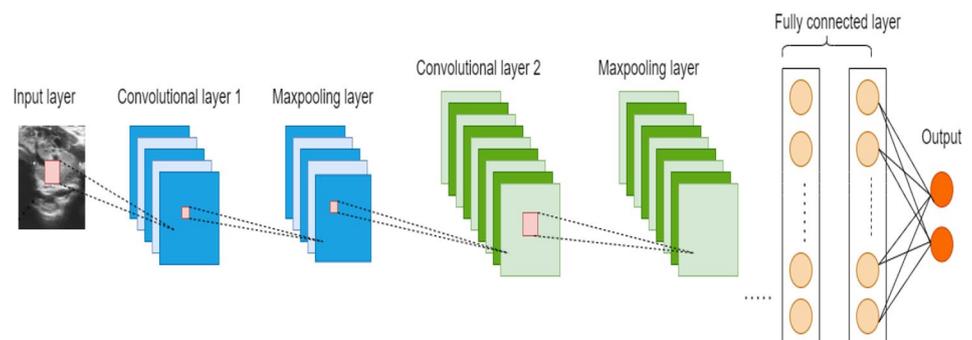
and classify the thyroid nodules. Data augmentation techniques are used to increase the training size of the model. The model has achieved an accuracy of 80.91%, sensitivity of 80% and specificity of 80.91% with a pre-trained VGG-16 model on the thyroid digital imaging database (TDID) dataset. Some of the limitations are less sample size, failed to address the problem of noise and the random selection of training and testing datasets.

Song et al. (2018) developed a multi-task cascade convolution neural network (MC-CNN) for automatic thyroid nodule detection. In their work, multi-scale single-shot detection is used for detecting nodules and MC-CNN algorithm is for the classification. Different classifiers like MC-CNN, Naive Bayes, multi-layer perceptron (MLP), Alexnet, Googlenet and gradient boosting decision tree (GBDT) were used to evaluate the models. On the TDID dataset, the model has achieved an accuracy of 92.1%, sensitivity of 94.1% and specificity of 96.2%. Among all the classifiers VGG-16 has performed better in comparison with other classifiers. Future work can be extended to exploring different types of deep learning techniques.

Nguyen et al. (2019) analysed the problem of frequency and spatial domains. A cascade classifier scheme is used to improve the model accuracy. The model has achieved an accuracy of 90.88% using the TDID open database. Frequency and spatial domain features were extracted for detecting thyroid nodules and CNN is used to train the model. One of the limitations of their work is that their model is quite complicated rather than using a single method and did not consider the problem of the imbalanced dataset.

Ko et al. (2019) proposed a deep convolutional neural network (DCNN) to diagnose thyroid malignancy on ultrasonography (USG) images. Using three trained CNN, the model has achieved an accuracy of 86%, sensitivity of 84% and specificity of 90% on "Our Institution Hospital". They have used imagenet-vgg and imagenet-verydeep techniques for the classification of thyroid nodules. The model is competitive for the identification of thyroid nodules. Some of the limitations are the random selection of training and testing datasets and less sample size. For the capturing of the image, there is an involvement of different types of scanners which lead to slow different diagnostic performances of CNNs. Future work can be extended to explore different types of DL models like Alexnet, Resnet-50, VGG-16, ETC.

Shi et al. (2020) proposed knowledge-guided synthetic image adversarial augmentation (ACGAN) to classify thyroid nodules. The model has achieved an accuracy of 90.63%, sensitivity of 90.63% and specificity of 92.65%

using 1937 USG thyroid images. Here, in this case authors have incorporated domain knowledge and deep learning together to improve the performance of the model. Some of the limitations are the less sample, failed to address the noise problem. Future work can be extended to semi-supervised learning and evaluating the model on large datasets.

Guo et al. (2020) proposed an improved deep learning approach to diagnose thyroid nodules. The squeeze and excitation (SE) modules were used for adaptive feature selection and maximum inter-pixel relations (MPR) for inter-pixel relations problem. The authors have also used data augmentation techniques along with the classification. The model has achieved an accuracy of 90.17%, sensitivity of 86.99% and specificity of 92.35% using 407 thyroid USG images on "Cooperated hospitals" dataset. Some of the limitations are the less sample size, failed to address the noise problem. Future work can be extended to semi-supervised learning and evaluating the model on large datasets.

Ajilisa et al. (2020) used a different pre-trained CNN network (Alexnet, Xception, VGG-16, Inception v-3, Googlenet, VGG-19, Resnet-10 and Resnet-50) to classify thyroid USG images. The K-means clustering technique is used to deal with imbalanced datasets. From their experiment, it is found that a deep neural network (DNN) outperforms from the rest of the networks with an accuracy of 89.93% and a sensitivity of 92.76% using the TDID dataset. Some of the limitations of their model were the small dataset size, skewed dataset distribution and not an automatic detection model to classify thyroid nodules from USG images.

Nguyen et al. (2021) analysed the problem of frequency and spatial domains. In their work, there is a combination of Resnet-50 and Inception-based neural networks to classify thyroid nodules. The model has achieved an accuracy of 92.05% on TDID open database using the weighted binary cross-entropy loss function for the handling of the imbalanced dataset. One of the limitations of their model is the small dataset size and not addressing the problem of noise removal. Future work can be extended to exploring different types of deep learning techniques for the identification of benign and malignant thyroid nodules.

Yang et al. (2021) proposed a multi-task cascade deep learning model (MCDLM) using the integration of domain knowledge (DK) and multimodal ultrasound (USG) images for the diagnosis of thyroid nodules. U-net technique was used for the segmentation of thyroid nodules along with a pre-trained VGG-13 model. The model has achieved an accuracy of 90.01%, sensitivity of 87.47% and specificity of 92.15% using 3090 thyroid USG images. Future work can be

extended to exploring fusion of traditional and deep learning techniques for the identification of malignant nodules.

Zhu et al. (2021) proposed a deep convolutional neural network (DCNN) to classify thyroid USG images and achieved an accuracy of 86.5%, sensitivity of 86.7% and specificity of 87.7% using 719 thyroid USG images. They have collected datasets from the "Shanghai Pundong People's hospital". Future work can be extended to use the fusion of ML/traditional learning and DL techniques.

2.1 Research gaps

This subsection throws light on the various research gaps identified in the existing studies. For evaluating and comparing the proposed model, we have considered those research papers that have used the public TDID dataset in their works. From the review, it is analysed that DL techniques attract wide attention from the researchers for the classification of thyroid nodules. Some of the points noted regarding the DL are summarized as:

1. Optimizing the parameters of the deep learning model plays a significant role to achieve better results.
2. The pre-trained classifiers have an inbuilt feature extraction module. Thus, there is a scope to use detection techniques along with pre-trained classifiers.
3. Pre-processing is an essential step for the development of any model as it enhances the performance of the model.
4. Deep learning techniques sometimes lead to the problem of overfitting; hence, it is essential to address the problem of less sample size of the datasets.

2.2 Research contribution

The following research contribution is incorporated into our work:

   i. GSO technique is used to optimize the learning rate and drop-out factor of DL models.
  ii. In this research work, various DL models like Alexnet, DNN, Resnet-50 and VGG-16 are explored along with the morphological operation, segmentation and boundary detection techniques.
 iii. Gaussian blur function and image enhancement techniques are used to address the problem of noise found in medical USG images.
 iv. For the comparison of the proposed model with the reported literature and DL models two datasets, the public Thyroid Ultrasound Image Database (TDID) and collected datasets from Kriti Scanning Center, Prayagraj, India is considered in this work.

3 Materials and methods

This section covers the various materials, i.e., datasets collected for this experiment and methods (pre-processing, data augmentation, segmentation, boundary detection and classification) adopted to propose this GSO-CNN model. Figure 2 shows the framework of the proposed methodology. A detailed description of the proposed model in four phases is discussed below:

3.1 Materials

In this work, two **datasets** are used to evaluate the proposed model. First public TDID dataset having 295 thyroid USG images, out of which 107 were benign, and 188 were malignant thyroid USG images (Pedraza et al. 2015). Second, **Local dataset** collected from Kriti Scanning Center, Prayagraj, Uttar Pradesh, India duly approved by National Accreditation Board for Hospitals and Healthcare Providers (NABH) (https://www.nabh.co/frmViewCGHSRecommend.aspx?Type=Diagnostic%20Centre&cityID=94 2022). The local/collected has the total number of 654 images, out of which 428 were benign and 226 were malignant thyroid USG images. The duration of dataset collection was from July 2020 to March 2021. Thyroid ultrasound video sequences with Voluson E-10, Mindray-Resona 7, SIEMENS Healthineers, and Voluson E-8 Ultrasound devices with 3–11 MHz convex and linear transductors captured the thyroid images. In both cases, Thyroid Imaging Reporting and Data system (TI-RADS) scores have been used to pre-classify the thyroid USG images (Richman et al. 2020).

3.2 Methods

Methods subsection covers the phases adopted for thyroid nodule classification after data collection (i.e., pre-processing phase, segmentation and boundary detection for the detection and deep learning techniques for classification) in this research work.

3.2.1 Data pre-processing and data augmentation techniques

Data pre-processing is essential for the development of any model. Initially, the images were of size $560X360$ pixels; it has been resized with $256X256$ pixels using Eq. 1:

$$B = \text{Imresize}(A, \text{scale}) \tag{1}$$

To reduce the complexity of the images, images are converted to grayscale from RGB scale using Eq. 2:
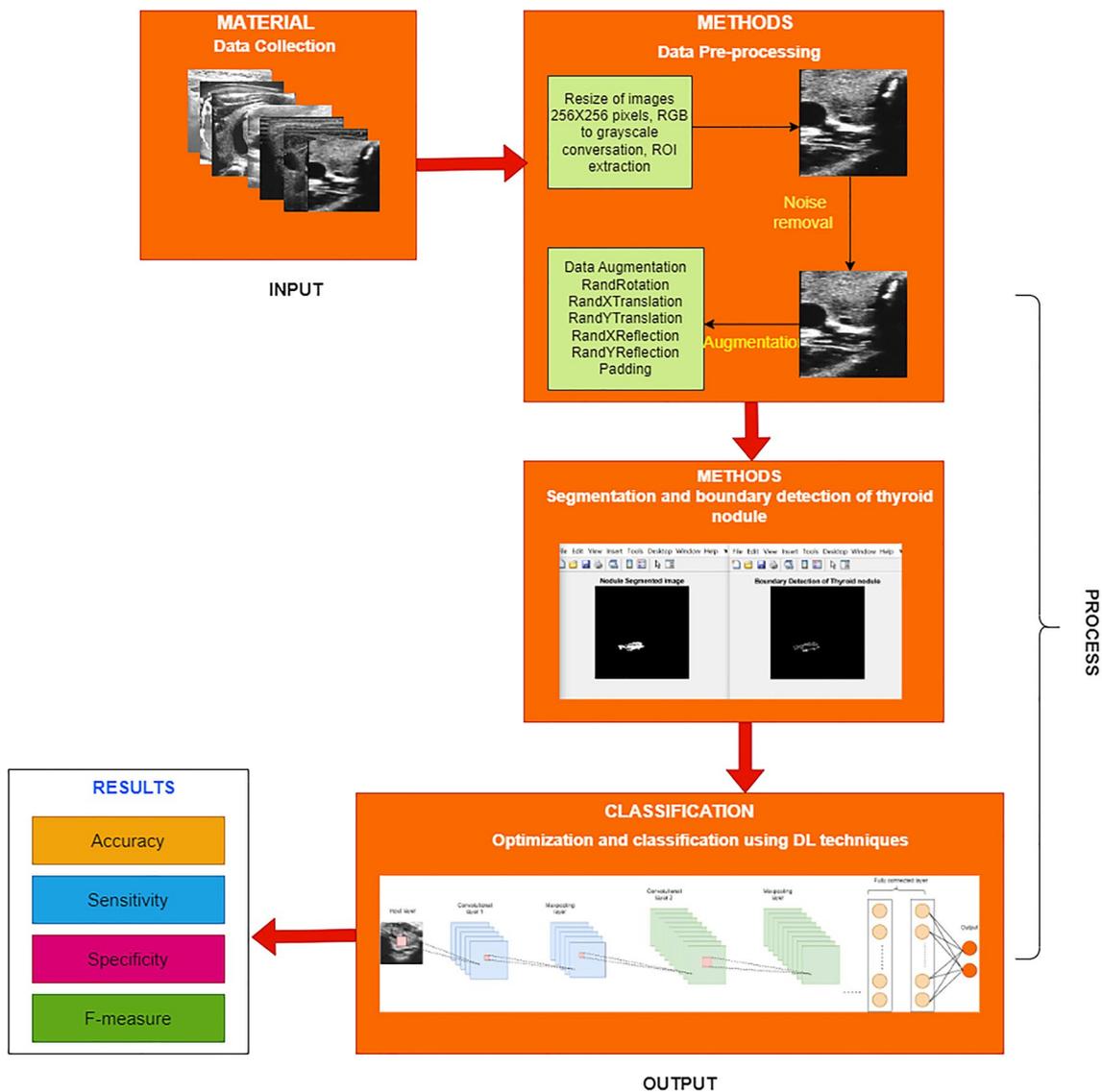
$$I = rgb2gray(RGB) \tag{2}$$

Fig. 2 Framework of proposed methodology

Two different noise removal techniques like median blur and Gaussian blur were considered for this work. After the analysis, it is found that Gaussian blur performs better. It can be computed using Eq. 3:

$$h = f\text{special(type)} \tag{3}$$

where $f$special: function returns the kernel correlation, h: creates a 2D filter.

Later on, sharpening and enhancement techniques were applied on the images. To reduce complexity, images were converted to grayscale. Sometimes network gets overfitted, due to the availability of small datasets. Data augmentation techniques are applied to increase the sample size and it

performs a geometric transformation to the image datasets to reduce the overfitting problem. It is applied using Eq. 4:

$$aug = \text{imageDataAugmenter(Name, Value)} \tag{4}$$

where Name: name of the techniques, Value: values assigned to techniques.

3.2.2 Morphological operations, segmentation and boundary detection techniques

These are a collection of non-linear operations related to the shape or morphology features of an image. It is applied on the binary image to find out the exact region of the thyroid
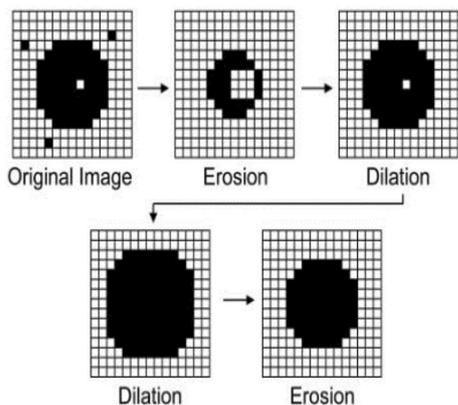
Fig. 3 The morphological operation (Merzban and Elbayoumi 2019)

nodule within the image using some basic operations (Ilhan and Ilhan 2017). The most commonly used operations are dilation and erosion. The structure used for dataset-1 and dataset-2 in our proposed model is "disk." Figure 3 shows the morphological operation.

The dilation function can be performed using Eq. 5:

$$I(\text{Dilate})S = \{I + S; \text{for all pixels in } I \in S\} \qquad (5)$$

where S: structuring element.

The erosion function can be performed using Eq. 6:

$$I(\text{Eroded})S = \{I(\text{Errored})S\}(\text{Eroded})S \qquad (6)$$

The morphological gradient function is performed for the enhancement of the edges. The subtraction function for this function can be performed using Eq. 7:

$$G = \{I(\text{Dilate})S - \{I(\text{Eroded})S\} \qquad (7)$$

3.2.2.1 Segmentation using active contour Image segmentation is defined as a technique used to derive information from the images whereas segmentation refers to the separation or segregation of information from the required region of an input image. Active contour is one of the types of image segmentation techniques which separate boundary or curvature for the regions of the input image (Hemalatha et al. 2018). If segmentation is successfully applied, the after stages of image analysis are made easier. It is also known as the snake method coined by M. Kass (Kass et al. 1988). It is basically the curve in the image space whose deformation is based on energy minimization. Let x and y be the position of co-ordinates of a 2D image I(x,y), and the curve be:

$$V(s) = (x(s), y(s)) \qquad (8)$$

where sϵ [0,1].

The energy function of the active contour is defined as Eqs. 9 and 10:

$$E_{\text{total}} = \int_0^1 E(V(s))ds \qquad (9)$$

$$E_{\text{total}} = \int_0^1 [E_{\text{int}}(V(s)) + E_{\text{exl}}(V(s))]ds \qquad (10)$$

where E_{int}: internal energy, E_{ext}: external energy.

To control the control deformation and displacement, these energy components (internal and external) are converted into internal and external forces.

The internal and external energy components can be defined as:

$$E_{\text{int}}(V(s)) = \left[\frac{1}{2}\alpha(s)\backslash V_s(s)|^2 + \frac{1}{2}\beta(s)|V_{ss}(s)|^2\right] \qquad (11)$$

where α: weight factor, β(s): control factor

$$E_{\text{exl}}(V(s)) = \gamma|\nabla[G_\sigma(x, y) * I(x, y)]|^2 \qquad (12)$$

where γ: weight factor, $G_\sigma(x, y)$: 2D gaussian function, σ: standard deviation, ∇: gradient operator, *: convolution operator.

Using Eq. 13, the E_{int} and E_{ext} attains the minimum value.

$$F_{\text{int}+}F_{\text{ext}} = 0 \qquad (13)$$

At the Eq. 13, the contour deformation is stop; hence we say that the goal is reached.

3.2.2.2 Boundary detection Previously in medical imaging, identified tumors and their boundaries were drawn by experienced experts, and contained some manual errors. Hence, to remove this error, the morphological operation 'remove' is used to remove all the interior pixels and leave the boundary pixels (Saba et al. 2019). Boundary detection finds the boundaries between the light and dark pixels in an image or we can say it finds the semantic boundaries between what humans would consider to be different objects or regions of the image (http://cs.brown.edu/courses/cs143/2011/results/proj2/edwallac/#:~:text=Boundary%20detection%20is%20an%20important,or%20regions%20of%20the%20image. 2022).

Boundary detection of an image is computed using Eq. 14:

$$B(A) = A - (A \ominus B) \tag{14}$$

where A: Matrix of an image M x N, B: Structuring element, $\ominus$: Erosion function, B(A): Boundary detection.

3.2.3 Classification

CNN or ConvNets has achieved a brilliant achievement in the field of image recognition, pattern analysis, medical image classification, etc. The main reason behind the success of CNN is the architecture design that is capable of considering local and global characteristics of the input image data (Pinaya et al. 2020). It is mainly a feedforward neural network (FNN), having a convolutional and pooling layer, and automatic feature extraction method (i.e., helpful in manual investments) (Li et al. 2021). Figure 4 shows the basic architecture of CNN. ConvNets performs in four layers, namely convolutional layer, pooling layer, fully connected layer and output layer. A detailed description is below:

i.   **Convolutional layer**
It consists of a set of filters which performs convolution operation between filters, thereby creating a feature map. Convolution operations are a combination of filter sliding and dot products. It also allows ConvNet to be spatial invariant (Truong et al. 2018). Whereas the feature maps store the result of the convolution operations. Figure 5 shows the convolution process. Equation 15 represents the convolution operation:
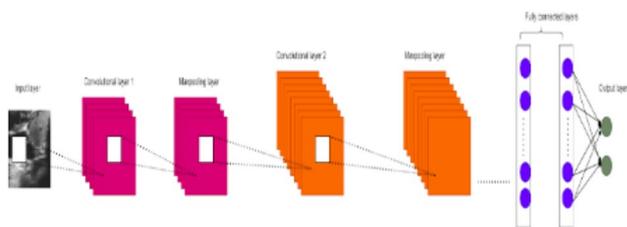
$$x_j = I * W_j + b_j, j = 1, 2 \dots f \tag{15}$$

where I: input, x_j: output, W_j: weights of the jth filter, b_j: jth bias.

ii. **Pooling layer**
The main aim of this layer is to reduce the spatial size of the image captured by the previous layer. In this work, we have used the max-pooling layer which takes the maximum value in the window, and neglecting the rest of the values (Pinaya et al. 2020). It can be computed using Eq. 16:

$$f_M = \text{Max}(C_y) \tag{16}$$

In this layer, stride and window size is also specified. Stride is defined as a number of pixels over the input matrix. If stride 1: it moves to filters to 1 pixel, if 2: it moves to filters to 2. Figure 6 shows the operation of max-pooling layer.

iii. **Fully connected layer**
It is abbreviated as the FC layer. In this layer, all the input is flattened into a matrix vector and then feed into an FC layer like a neural network (NN). Figure 7 shows the FC layer.

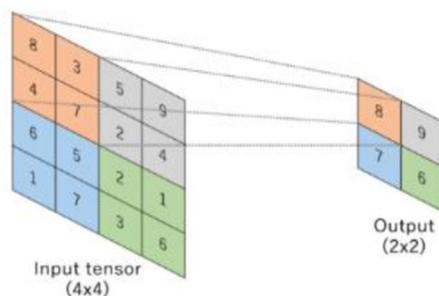**Fig. 4** Basic architecture of CNN (Gu et al. 2019)

Fig. 6 The operation of max-pooling layer (https://medium.com/@RaghavPrabhu, understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148. 2022)

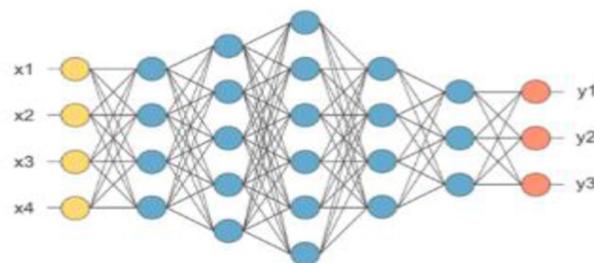**Fig. 5** Convolution process (Esteva et al. 2017)

Fig. 7 The fully connected layer (Pinaya et al. 2020)

iv.   **Output layer**

At the end, the classified output is obtained. In this research work, the softmax function is used for the classification which can be computed using Eq. 17:

$$S(x_i) = \frac{e^x i}{\sum_{j=1}^{n} e^x j} \qquad (17)$$

where x: input; x_i: i-th element of the input vector; e^x_i: exponential; $\sum_{j=1}^{n} e^x_j$: normalized term.

Algorithm 1: Classification of thyroid nodule using GSO-CNN model
Input: Image dataset
Output: Prediction

Step 1	Upload the image dataset
Step 2	Perform pre-processing step like image resizing, RGB to grey-scale conversion, noise removal, data augmentation using equation 1, 2, 3,4
Step 3	Perform morphological operation using equation 5, 6, 7. Use active contour segmentation technique for the segmentation of nodules using equation 13.
Step 4	Use boundary detection to find boundary of the thyroid nodules using equation 14.
Step 5	Set the training and testing ratio
Step 6	Initialize the parameters of CNN.
Step 7	Apply GSO technique to find the best parameter of learning rate and drop-out factor.
Step 8	Train the model
Step 9	Prediction

3.2.3.1 Hyper-parameters optimization using GSO This section focusses on the tuning of the learning rate and drop-out factor to enhance the performance of the model.

Let the hyperparameters be $hyp_1,\ldots\ldots hyp_k$ of the DL model and $\alpha_1,\ldots\ldots\alpha_n$ be their respective domains. The GSO-CNN model is trained with hyp (DL_{train}) on the training data of thyroid USG images. The testing part of the model is significant as λ (hyp, D_{train}, D_{test}). The main motive behind the hyp is to find best parameter of model and is denoted by hyp*.

Algorithm 2: GSO stage 1
Input: hyp
Iteration per stage $\rightarrow Y=<Y_1, Y_2\ldots\ldots Y_K>$,
Total no of stages$\rightarrow$ T
Training data DL_{train} $\rightarrow < DL_{train1, \ldots\ldots} DL_{trainz} >$
Test data $\rightarrow DL_{test}$
Accuracy$\rightarrow \lambda$
Output: hyp*

    for each stage y=1 to T do
    for i=1 to Í
        λ_i= evaluate λ (hyp_i, DL_{train} , DL_{test})
    end
    for Í+1 to Y_k
        g= grid_search(hyp_i, λ_i)
        h_j= maxargs $_{h\epsilon\alpha}$ a (hyp, g)
        λ_i= evaluate λ (hyp_i, DL_{train} , DL_{test})
    end
    Reset hyp= best hyp config ϵ (hyp_1, $\ldots\ldots hyp_n$)
    end
    Return: hyp*

3.2.3.2 GSO stage 1 In the first stage, a small subset of the DL_{train} is trained on DL_{test} for quick identification of an initial stage. These hyp settings are further unitized on the GSO algorithm. Let $hyp_1,\ldots hyp_n$ be the initial arrangement of hyp with many assignments. From this, the assessment of hyp values is recorded based on accuracy. Algorithm 2 explains the steps involved for stage 1 of the GSO technique.

3.2.3.3 GSO stage 2 The above figure explains the complete workflow of the GSO technique. The first stage is the initial stage where the input is the pre-processed, segmented and boundary of the thyroid nodules. The classification is done using GSO-CNN method. As the hyp-like epoch, batch size, optimizer, dropout factor, learning rate and activation function help to build a good performance of the model. In this work, while using GSO, some parameters like batch size, optimizer, epoch, activation function are kept constant. The learning rate and drop-out factor are set according to the GSO result. Here, in this stage learning rate is optimized and is stored in hyp_1*. The algorithm 3 explains the complete process of how the learning rate is optimized.

Algorithm 3: GSO stage 2
Input: Img→ image dataset
Output: hyp$_1$*
    create dlmodel ()
    ml.add(values of learning rate, optimization algo)
    gd= GdSearchCV(par gd=par, gd,est=ml)
    gd_result=gd.fit( )
    Print ("best learning rate, optimization para");
    Return hyp$_1$ *

Algorithm 4: GSO stage 3
Input: hyp$_1$*
Output: hyp$_1$*, hyp$_2$*
    create dlmodel ()
    ml.add(hyp$_1$*, values of drop-out factor, optimization algo)
    gd= GdSearchCV(par gd=par, gd,est=ml)
    gd_result=gd.fit( )
    Print ("best drop-out factor, best learning rate, optimization para");
    Return hyp$_1$ *, hyp$_2$ *

3.2.3.4 GSO stage 3 The parameters hyp like epoch, batch size, optimizer, dropout-factor, learning rate and activation function help to build a good performance of the model. In this stage 3, some parameters like batch size, optimizer, epoch and activation function are kept constant. Once, we have the learning rate values generated by GSO, the learning rate along with drop-out factor is optimized together. The learning rate and drop-out factor are set according to GSO result. The optimized dropout factor is stored in hyp$_2$*. The algorithm 4 explains completely how the drop-out and learning rate are optimized. Figure 8 shows the complete workflow of all the GSO stages used in this research work to find the best parameters of learning rate and dropout factor.

Fig. 8 The complete workflow of the GSO used in this research work

4 Results and discussion

This section discusses the simulation results of the proposed model. The complete experiment has been performed on MATLAB 2019b, 64bit operating system (OS), Intel (R) Core (TM) i5-8250U CPU, 1 TB solid-state drive (SSD), NVIDIA GeForce and 16 GB RAM. Some parts of the experiment are also performed on python Go-laboratory with TensorFlow and keras as the backend. Based on all these four outcomes, the following four parameters are computed in our work:
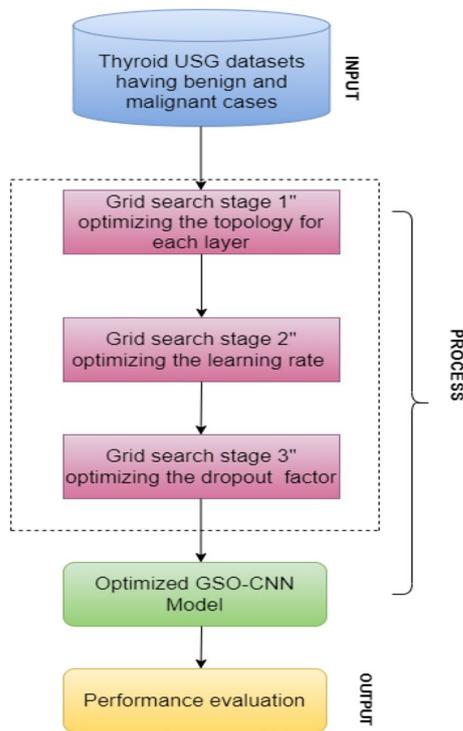
$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{TN}} \tag{18}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{19}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{20}$$

$$F - \text{measure} = 2 * \frac{\text{Precision} * \text{recall}}{\text{Precision} + \text{recall}} \tag{21}$$

The proposed model has been compared with the various state-of-the-art models and popular deep learning techniques like DNN, Resnet-50, Alexnet and VGG-16. For better representation, the public TDID dataset is renamed as dataset-1 and the collected dataset as dataset-2. In most of the thyroid nodule detection and classification problems, the authors have used local/private datasets. Therefore, we have decided to use public TDID and local datasets, i.e., the collected dataset in our work. The main reason for using the TDID dataset in our work is that authors in the previous studies of thyroid nodule problems have used this dataset (Wang et al. 2018; Song et al. 2018; Nguyen et al. 2019, 2021; Ajilisa et al. 2020). An 8:2 ratio is set for all the models on dataset-1 and dataset-2. Various data augmentation techniques have

been utilized to enhance the generalization capability of the proposed model. This helped us to evaluate and compare the performance of the proposed model with the previous studies.

Initially, the sample size for dataset-1 and dataset-2 was 295 and 654, after applying six different data augmentation techniques the size of dataset-1 becomes 1770 i.e., 295X6 and dataset-2 becomes 3924 thyroid USG images. 8:2 ratio is set for training and testing of the model. Table 1 shows the comparison of the proposed model and popular DL models with and without segmentation, boundary detection and classification techniques on dataset-1 and dataset-2. The GSO-CNN model has achieved (94.11%, 95.65%, 93.33%, 96.20%); (95.33%, 96.66%, 95%, 97.20%) in terms of (accuracy, sensitivity, specificity, f-measure) on (dataset-1, dataset-2). After apply case-2, the GSO-CNN model has achieved (95.30%, 96.66%, 94.87%, 97.20%); (96.02%, 97.02%, 95%, 98.34%) in terms (accuracy, sensitivity,

specificity, f-measure) on (dataset-1, dataset-2). It is noticed that the proposed GSO-CNN model is giving good results in comparison with case-2. Further, it is also inferred from the table that the proposed model and other DL models are performing better in case 2 with an improvement of 1% to 2%.

Table 2 shows the performance comparison of the various state-of-the-art models for detecting and classifying thyroid nodules on dataset-1 and dataset-2.

The proposed GSO-CNN model has achieved (95.30%, 96.66%, 94.87%, 97.20%) on dataset-1 and (96.02%, 96.70%, 95%, 98.34%) on dataset-2 on the parameters (accuracy, sensitivity, specificity, f-measure). However, an increment of 3%, 4%, 8% is seen in accuracy, sensitivity, f-measure parameters with the reported literature. But a decrement of 2% is reported in comparison with the reported literature in specificity.

Table 3 shows the parameter settings for dataset-1 and 2. The best result is obtained with drop-out factor:0.2, batch

Table 1 Comparison of the proposed model and popular DL models with and without segmentation, boundary detection techniques on dataset-1 and dataset-2

Dataset	Techniques	Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	*F* measure (%)
Dataset-1	CASE 1 Without segmentation and boundary detection techniques	Proposed (GSO-CNN) model	94.11	95.65	93.33	96.20
		DNN	90.90	92.63	89.85	89
		Alexnet	82.98	83.96	80.98	83.96
		Resnet-50	86.48	87.25	84.52	87.25
		VGG-16	85.71	86.40	83.90	86.40
Dataset-2		Proposed (GSO-CNN) model	95.33	96.66	95	97.20
		DNN	91.81	92.63	90.78	92.14
		Alexnet	82.28	83.51	81.92	82.60
		Resnet-50	90.11	91.66	89.33	93.61
		VGG-16	86.63	87.25	84.88	87.25
Dataset-1	CASE 2 With segmentation and boundary detection techniques	Proposed (GSO-CNN) model	95.30	96.66	94.87	97.20
		DNN	91.01	92.55	90.27	93.54
		Alexnet	83.85	84.76	81.81	84.76
		Resnet-50	87.91	89	86.58	88.55
		VGG-16	86.66	87.12	83.95	87.56
Dataset-2		Proposed (GSO-CNN) model	96.02	97.02	95	98.34
		DNN	92.30	93.68	90.16	93.19
		Alexnet	83.08	84.76	82.97	85.98
		Resnet-50	91.76	92.63	90.66	92.14
		VGG-16	87.16	88.11	86.04	87.68

Table 2 The performance comparison of the various state-of-the-art models on dataset-1 and dataset-2

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	*F* measure (%)
Wang et. al (2018)	80.91	81.82	80	
Song et. al (2019)	92.1	94.1	96.2	
Nguyen et. al (2019)	90.88	–	–	
Ajilisa et. al (2020)	89.93	92.76	–	89.43
Nguyen et. al (2021)	92.05	–	–	
Proposed (GSO-CNN) model on dataset-1	**95.30**	**96.66**	**94.87**	**97.20**
Proposed (GSO-CNN) model on dataset-2	**96.02**	**96.70**	**95**	**98.34**

Table 3 Parameter settings for datatset-1 and 2

Parameters	Dropout Factor	Max Epoch	Batch size	Learning rate	Activation function	Optimizer	Max-pooling
Values	0.2	10	15	0.01	SoftMax	SGDM	2×2

**Table 4** Comparison of the proposed model with the popular DL techniques on dataset-1 and dataset-2

Techniques	Accuracy (%)	Sensitivity (%)	Specificity (%)	F-measure (%)
DNN dataset-1	91.01	92.55	90.27	93.54
DNN dataset-2	92.30	93.68	90.16	93.19
Alexnet dataset-1	82.84	83.17	81.63	83.17
Alexnet dataset-2	83.08	84.76	82.97	85.98
Resnet-50 dataset-1	86.01	87.25	85.55	92.70
Resnet-50 dataset-2	91.76	92.63	90.66	92.14
VGG-16 dataset-1	86.66	87.12	83.95	87.56
VGG-16 dataset-2	87.16	88.11	86.04	87.88
Proposed (GSO-CNN) model dataset-1	**95.30**	**96.66**	**94.87**	**97.20**
Proposed(GSO-CNN) model dataset-2	**96.02**	**96.70**	**95**	**98.34**

Table 5 Results of GSO on different parameters set for dataset-1 and 2 on proposed model and DL models

Para-meters								
	Learning rate	0.1	0.04	0.05	0.07	0.08	0.09	**0.01**
	Dropout factor	0.1	0.02	0.04	0.06	0.08	0.09	**0.2**
Accuracy dataset-1 models	Proposed (GSO-CNN) Model	76.06	80.90	82.32	85.16	87.19	89	**95.30**
	DNN	62.3	69	71.1	76.88	81.1	82.76	**91.01**
	Alexnet	58	59.99	64	68.3	73.6	77.9	**82.84**
	Resnet-50	54	62.7	69.6	64	74.29	78.41	**86.01**
	VGG-16	59	68.2	73.3	78.55	80.3	81.27	**86.66**
Accuracy Dataset-2 models	Proposed (GSO-CNN) Model	77.39	84.96	86.40	88.86	90.81	92.70	**96.02**
	DNN	59.5	65.2	72.2	79.27	85.1	88.9	**92.30**
	Alexnet	57.11	64.44	69.9	73.23	75	78.9	**83.08**
	Resnet-50	72.6	76	81.49	86	88.1	91	**91.76**
	VGG-16	60	64.7	71.34	75	79.81	81.9	**87.16**

size: 15, learning rate: 0.01, activation function: soft-max, optimizer: stochastic gradient descent (SGDM), max-pooling:2X2. Table 4 shows a comparison of the proposed model with the popular deep learning techniques on dataset-1 and dataset-2. Further, it can be seen that the proposed GSO-CNN model has shown an improvement of 4% in comparison with the DNN, Resnet-50, VGG-16 and Alexnet models.
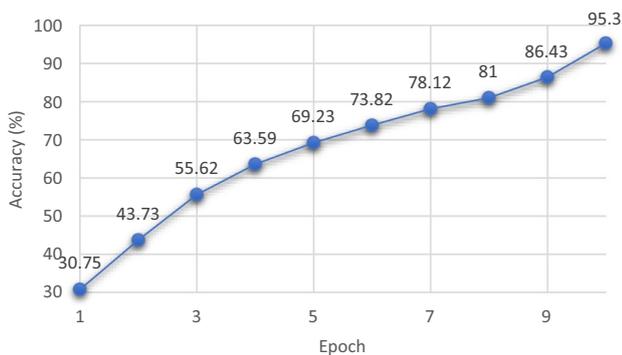
Fig. 9 Accuracy epoch graph for dataset-1
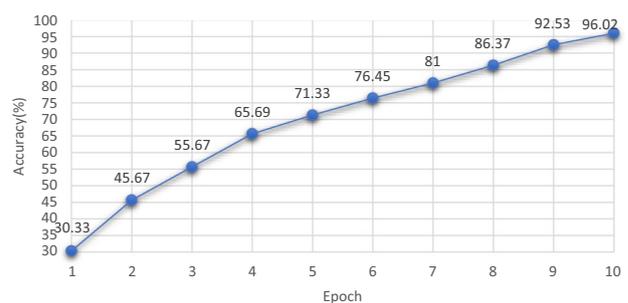
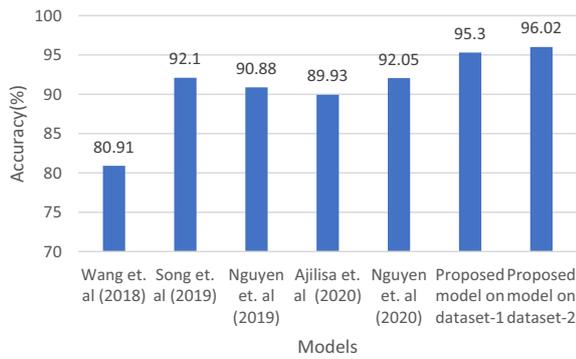

Fig. 10 Accuracy epoch graph for dataset-2

Fig. 11 Comparative analysis of the various state-of-the-art models

Fig. 12 Comparative analysis of the various state-of-the-art models based on accuracy on dataset-1 and 2 based on sensitivity, specificity and f-measure on dataset-1 and 2

Table 5 shows the results of GSO on different parameters set for learning rate and dropout factor on dataset-1 and dataset-2 on proposed (GSO-CNN) model and DL models. In this research work, different values of learning rate (0.1, 0.04, 0.05, 0.07, 0.08, 0.09, 0.01) and dropout-factor (0.1, 0.02, 0.04, 0.06, 0.08, 0.09, 0.2) is optimized using GSO technique. The model performs better with the learning rate of 0.01 and dropout factor of 0.2 on all the DL models like DNN, Alexnet, Resnet-50, VGG-16 and the proposed (GSO-CNN) model. Figure 9 shows the accuracy epoch graph for dataset-1. From the figure, it can be analysed that the epoch and accuracy of the models are increasing simultaneously. The best accuracy of the models is achieved on the 10th epoch with an accuracy of 95.30% on dataset-1. Figure 10 shows the accuracy epoch graph for dataset-2. The best accuracy is achieved on the 10th epoch with 96.02%. Figure 11 shows the comparative analysis of the various state-of-the-art models
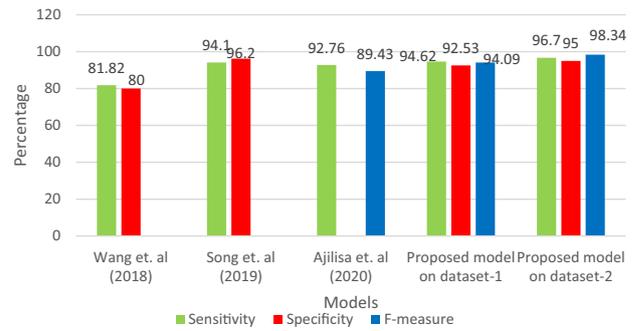
and proposed model accuracy on dataset-1 and 2. It can be seen that a 3% to 4% improvement is reported by the proposed (GSO-CNN) model in comparison with Wang et al., Song et al., Nguyen et al., Ajilisa et al., Nguyen et al. models. Figure 12 shows the comparative analysis of the various state-of-the-art models and proposed model sensitivity, specificity and f-measure on dataset-1 and 2. The proposed model has achieved sensitivity of 96.66%, specificity of 94.87% and f-measure of 97.20% on dataset-1. While the model has achieved sensitivity of 96.70%, specificity of 95% and f-measure of 98.34% on dataset-2. Figure 13 shows the performance comparison of the popular deep learning techniques DNN, Alexnet, Resnet-50, VGG-16 and the proposed model on dataset-1 and dataset-2 respectively. The proposed (GSO-CNN) model has achieved an accuracy of 95.3%, sensitivity of 96.66%, specificity of 94.87% and f-measure of 97.2% on
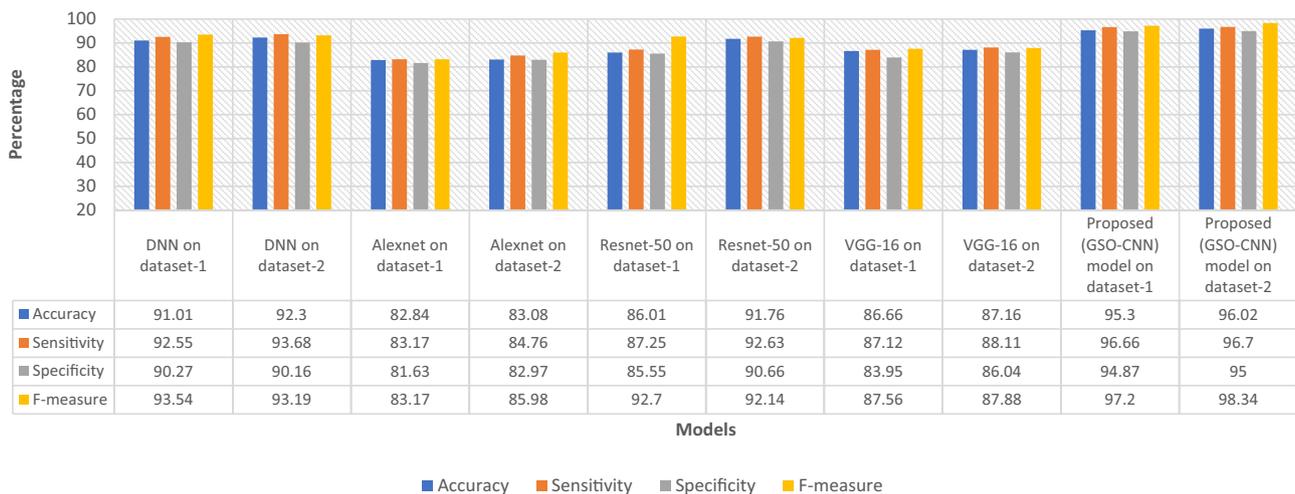

	DNN on dataset-1	DNN on dataset-2	Alexnet on dataset-1	Alexnet on dataset-2	Resnet-50 on dataset-1	Resnet-50 on dataset-2	VGG-16 on dataset-1	VGG-16 on dataset-2	Proposed (GSO-CNN) model on dataset-1	Proposed (GSO-CNN) model on dataset-2
Accuracy	91.01	92.3	82.84	83.08	86.01	91.76	86.66	87.16	95.3	96.02
Sensitivity	92.55	93.68	83.17	84.76	87.25	92.63	87.12	88.11	96.66	96.7
Specificity	90.27	90.16	81.63	82.97	85.55	90.66	83.95	86.04	94.87	95
F-measure	93.54	93.19	83.17	85.98	92.7	92.14	87.56	87.88	97.2	98.34

Fig. 13 Performance comparison of the popular deep learning techniques DNN, Alexnet, Resnet-50, VGG-16 and proposed model on dataset-1 and dataset-2

dataset-1. While on dataset-2, GSO model has achieved 96.02% accuracy, 96.7% sensitivity, 95% specificity and 98.34% f-measure. It is clear from the figure that the proposed (GSO-CNN) model has performed better in comparison with other deep learning techniques followed by DNN having a performance less than 95%, Resnet-50 having performance less than 93%, VGG-16 having performance less than 89% and Alexnet having performance less than 87% with respect to the accuracy, sensitivity, specificity and f-measure parameters.

5 Conclusions

In this work, a GSO-CNN-based model for thyroid nodule classification is proposed using data augmentation, segmentation and boundary detection techniques. The model is evaluated on public (TDID) and collected datasets collected from Kriti Scanning Center, Prayagraj, India, duly approved by NABH. The GSO technique is used to optimize the learning rate and drop-out factor parameters for training the proposed model. The proposed model has shown an improvement of (6.126%, 6.846%), (7.1%, 7.14%), (6.77%, 6.9%) and (7.77%, 8.91%) in terms of accuracy, sensitivity, specificity and f-measure on (dataset-1, dataset-2) against other state of the art models. The proposed approach works well even when there is a limited number of training images. It has shown an improvement of 3% to 5% in the performance evaluation compared with deep learning techniques like DNN, Resnet-50, Alexnet and VGG-16 for thyroid nodule classification with considering segmentation and boundary detection techniques. It can be concluded that our proposed model could give a more accurate prediction to the clinicians for detection and classification of thyroid nodules.

Declarations

Conflict of interest We declare no conflict of interest.

Availability of data Data cannot be made publicly available due to privacy concern.

References

Ajilisa OA, Jagathyraj VP, Sabu MK (2020) Computer-aided diagnosis of thyroid nodule from ultrasound images using transfer learning from deep convolutional neural network models. In 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA). IEEE. pp. 237–241

Awan SE, Sohel F, Sanfilippo FM, Bennamoun M, Dwivedi G (2018) Machine learning in heart failure: ready for prime time. Curr Opin Cardiol 33(2):190–195

Boesch G (2022) "Deep residual networks (Resnet-50), Published 2022. https://viso.ai/deep-learning/resnet-residual-neural-network/. Accessed 15 Sep 2022

Cinaglia P, Tradigo G, Cascini GL, Zumpano E, Veltri P (2018) A framework for the decomposition and features extraction from lung DICOM images. In Proceedings of the 22nd international database engineering & applications symposium. pp. 31–36.

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118

Gu H, Wang Y, Hong S, Gui G (2019) Blind channel identification aided generalized automatic modulation recognition based on deep learning. IEEE Access 7:110722–110729

Guo X, Zhao H, Tang Z (2020) An improved deep learning approach for thyroid nodule diagnosis. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE. pp. 296–299.

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778.

HealthITAnalytics. Over 80% of health execs have artificial intelligence plans in place. HealthITAnalytics. Published November 2, 2020. Accessed February 20, 2022. https://healthitanalytics.com/news/over-80-of-health-execs-have-artificial-intelligence-plans-in-place.

Hemalatha RJ, Thamizhvani T, Dhivya AJA, Joseph JE, Babu B, Chandrasekaran R (2018) Active contour based segmentation techniques for medical image analysis. Med Biol Image Anal 4(17):2

http://cs.brown.edu/courses/cs143/2011/results/proj2/edwallac/#:~:text=Boundary%20detection%20is%20an%20important,or%20regions%20of%20the%20image. Accessed 19 Oct 2021

https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148. Accessed 4 Mar 2018

https://www.nabh.co/frmViewCGHSRecommend.aspx?Type=Diagnostic%20Centre&cityID=94. Accessed 4 Sept 2022

Ikechukwu AV, Murali S, Deepu R, Shivamurthy RC (2021) ResNet-50 vs VGG-19 vs training from scratch: a comparative analysis of the segmentation and classification of pneumonia from chest X-ray images. Global Trans Proc 2(2):375–381

Ilhan U, Ilhan A (2017) Brain tumor segmentation based on a new threshold approach. Procedia Computer Science 120:580–587

Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput Vision 1(4):321–331

Ko SY, Lee JH, Yoon JH, Na H, Hong E, Han K, Kwak JY (2019) Deep convolutional neural network for the diagnosis of thyroid nodules on ultrasound. Head Neck 41(4):885–891

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 25:1097–1105

Li S, Wang L, Li J, Yao Y (2021) Image classification algorithm based on improved AlexNet. In Journal of Physics: Conference Series. IOP Publishing. 1813(1):012051

Llamas J, Lerones MP, Medina R, Zalama E, Gómez-García-Bermejo J (2017) Classification of architectural heritage images using deep learning techniques. Appl Sci 7(10):992

Lo-Ciganic WH, Donohue JM, Yang Q, Huang JL, Chang CY, Weiss JC, Gellad WF (2022) Developing and validating a machine-learning algorithm to predict opioid overdose in Medicaid beneficiaries in two US states: a prognostic modelling study. The Lancet Digital Health 4(6):e455–e465

Merzban MH, Elbayoumi M (2019) Efficient solution of Otsu multilevel image thresholding: a comparative study. Expert Syst Appl 116:299–309

Mohsen H, El-Dahshan ESA, El-Horbaty ESM, Salem ABM (2018) Classification using deep learning neural networks for brain tumors. Future Comput Inf J 3(1):68–71

Nagavelli U, Samanta D, Chakraborty P (2022) Machine learning technology-based heart disease detection models. J Healthcare Eng. https://doi.org/10.1155/2022/7351061

Nguyen DT, Pham TD, Batchuluun G, Yoon HS, Park KR (2019) Artificial intelligence-based thyroid nodule classification using information from spatial and frequency domains. J Clin Med 8(11):1976

Nguyen DT, Kang JK, Pham TD, Batchuluun G, Park KR (2021) Ultrasound image-based diagnosis of malignant thyroid nodule using artificial intelligence. Sensors 20(7):1822

Pedamkar P, "Introduction to DNN network", Published August, 6, 2021, https://www.educba.com/dnn-neural-network/

Pedraza L, Vargas C, Narváez F, Durán O, Muñoz E, Romero E (2015) An open access thyroid ultrasound image database. In 10th International Symposium on Medical Information Processing and Analysis. International Society for Optics and Photonics. 9287:92870W

Pinaya WHL, Vieira S, Garcia-Dias R, Mechelli A (2020). Convolutional neural networks. In Machine learning. Academic Press. pp. 173–191

Rastogi A, "Resnet-50", Published March, 14, 2020, Accessed June 11, 2022. https://blog.devgenius.io/resnet50-6b42934db431

Rezvani S, Wang X, Pourpanah F (2019) Intuitionistic fuzzy twin support vector machines. IEEE Trans Fuzzy Syst 27(11):2140–2151

Richman DM, Benson CB, Doubilet PM, Wassner AJ, Asch E, Cherella CE, Frates MC (2020) Assessment of American college of radiology thyroid imaging reporting and data system (TI-RADS) for pediatric thyroid nodules. Radiology 294(2):415–420

Saba T, Khan MA, Rehman A, Marie-Sainte SL (2019) Region extraction and classification of skin cancer: a heterogeneous framework of deep CNN features fusion and reduction. J Med Syst 43(9):1–19

Sengupta S, Basak S, Saikia P, Paul S, Tsalavoutis V, Atiah F, Peters A (2020) A review of deep learning with special emphasis on architectures, applications and recent trends. Knowl-Based Syst 194:105596

Shaban WM, Rabie AH, Saleh AI, Abo-Elsoud MA (2021a) Detecting COVID-19 patients based on fuzzy inference engine and deep neural network. Appl Soft Comput 99:106906

Shaban WM, Rabie AH, Saleh AI, Abo-Elsoud MA (2021b) Detecting COVID-19 patients based on fuzzy inference engine and deep neural network. Appl Soft Comput 99:106906

Shi G, Wang J, Qiang Y, Yang X, Zhao J, Hao R, Kazihise NGF (2020) Knowledge-guided synthetic medical image adversarial augmentation for ultrasonography thyroid nodule classification. Comput Methods Programs Biomed 196:105611

Shivhare E, Saxena V (2022) Optimized generative adversarial network-based breast cancer diagnosis with wavelet and texture features. Multimed Syst 28:1–17

Song W, Li S, Liu J, Qin H, Zhang B, Zhang S, Hao A (2018) Multitask cascade convolution neural networks for automatic thyroid nodule detection and recognition. IEEE J Biomed Health Inform 23(3):1215–1224

Truong ND, Nguyen AD, Kuhlmann L, Bonyadi MR, Yang J, Ippolito S, Kavehei O (2018) Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. Neural Netw 105:104–111

Wang Z, Wang X (2018) A deep stochastic weight assignment network and its application to chess playing. Jparallel Distrib Comput 117:205–211

Wang XZ, Zhang T, Wang R (2017) Noniterative deep learning: Incorporating restricted Boltzmann machine into multilayer random weight neural networks. IEEE Trans Syst Man Cybern Syst 49(7):1299–1308

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

Wang J, Li S, Song W, Qin H, Zhang B, Hao A (2018) Learning from weakly-labelled clinical data for automatic thyroid nodule classification in ultrasound images. In 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE. pp. 3114–3118

Yang W, Dong Y, Du Q, Qiang Y, Wu K, Zhao J, Zia MB (2021) Integrate domain knowledge in training multi-task cascade deep learning model for benign–malignant thyroid nodule classification on ultrasound images. Eng Appl Artif Intell 98:104064

Zheng J, Lin D, Gao Z, Wang S, He M, Fan J (2020) Deep learning assisted efficient AdaBoost algorithm for breast cancer detection and early diagnosis. IEEE Access 8:96946–96954

Zhu YC, AlZoubi A, Jassim S, Jiang Q, Zhang Y, Wang YB, Hongbo DU (2021) A generic deep learning framework to classify thyroid and breast lesions in ultrasound images. Ultrasonics 110:106300