Neural Computing and Applications
https://doi.org/10.1007/s00521-021-06020-8

ORIGINAL ARTICLE

An optimal segmentation with deep learning based inception network model for intracranial hemorrhage diagnosis

Romany F. Mansour[1] · Nojood O. Aljehane[2]

Received: 5 October 2020 / Accepted: 6 April 2021

Abstract
Traumatic Brain Injury (TBI) leads to intracranial hemorrhages (ICH), which is a severe illness resulted in death if it is not properly diagnosed and treated in the earlier stage. Presently, computer tomography (CT) images are widely used by radiologists to identify and locate the regions of ICH. But it is a tedious task and mainly depends on the professional radiologists. This paper develops new deep learning (DL)-based ICH diagnosis and classification (DL-ICH) model using optimal image segmentation with Inception Network. The proposed DL-ICH model involves preprocessing, segmentation, feature extraction, and classification. Firstly, the input data undergo format conversion where the NIfTI files are converted into JPEG format. Next, Kapur's thresholding with an elephant herd optimization (EHO) algorithm called KT-EHO is employed for image segmentation. Then, DL based Inception v4 network is applied as a feature extractor to extract a useful set of features and a multilayer perceptron (MLP) is used for classification. An extensive set of simulations takes place to ensure the effective diagnostic performance of the DL-ICH model and the results are investigated under diverse dimensions. The experimental results achieved better accuracy rate.

Keywords Deep learning · Kapur's thresholding · EHO algorithm · Inception network · Intracranial hemorrhages

1 Introduction

In past decades, Traumatic Brain Injury (TBI) is the main reason for increasing mortal rate and disability in the US. Around 30% of injury deaths were recorded [1]. Followed by, TBI, extra-axial intracranial tumors like intracranial hemorrhages (ICH), might occur. The ICH disease is a root cause for death all over the world which occurs for all age groups. Initially, the disease is caused in the brain because of the leakage in blood vessels and eliminates the path of interaction (follow the brain instructions and function accordingly) and internal organs that lead to inactive body functions like loss of eyesight, speech, memory loss, and so on. The major risk factors such as head trauma, high Blood Pressure (BP), infected blood vessel walls, leakage in veins are correlated with ICH. In order to examine these disorders, the screening modalities like X-ray, Magnetic resonance imaging (MRI), Computed tomography (CT), positron emission tomography (PET), single-photon emission computed tomography (SPECT) are accessible to brain hemorrhage imaging. When compared with other models, a CT scan is applied extensively in hemorrhage detection as it is inexpensive, easily accessible, a limited duration for imaging. Thus, a CT scan is majorly preferred for ICH diagnosis. The manifestations of ICH clots on CT images are based on the external aspects such as position, density, volume, slice intensity, and scientific factors like scanning angles.

The earlier prediction of ICH is essential in order to provide better treatment and sufficient scheduling of scanning. Hence, massive developers have deployed computer-based diagnosis (CAD) model for ICH segmentation. The newly presented computer-based CAD model of ICH depends upon factors like automated segmentation of hemorrhage which is predicted with no professional

✉ Romany F. Mansour
romanyf@aun.edu.eg

Nojood O. Aljehane
noaljohani@ut.edu.sa

[1] Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt

[2] College of Computer Science and Information Technology, Tabuk University, Tabuk, Saudi Arabia

Published online: 22 April 2021

contribution, Manual segmentation where the human expert has to provide an appropriate input for segmentation. Here, the entropy relied automated unsupervised brain intracranial hemorrhage segmentation has been presented by applying brain CT images. The latest deployment in deep learning (DL) and convolutional neural networks (CNNs) served a tremendous performance in automated image classification as well as segmentation operations. Therefore, the DL method is capable to make automatic ICH prediction and segmentation. Then, a fully convolutional network (FCN), named as U-Net has been projected to the segmentation of ICH regions in all slices of CT. The automatic device for ICH prediction and classification has been applied for guiding radiologists while no professionals are available in emergency times, mostly in developing countries and remote areas. This kind of machine helps to limit the duration and failure in ICH diagnosis effectively.

Various conventional and Deep Learning (DL) methodologies are deployed in this study. Based on the classical Machine Learning (ML) approaches, Yuh et al. [2] established a threshold-based model for ICH prediction. This approach is mainly employed for predicting ICH subtypes according to the position, structure, and quantity. Alternatively, Li et al. [3] presented 2 approaches for SAH segmentation and applied the segmented sites for predicting the SAH hemorrhage. Initially, flexible registration with SAH space atlas, while extracted distance transform features as well as presented a Bayesian decision approach for computing the delineation. Once the SAH segmentation is completed, mean gray value, variance, entropy, and energy have been obtained for training a Support Vector Machine (SVM) classification model and predict the SAH hemorrhage.

Based on the DL models, many other approaches were relied on CNNs and corresponding variants by excluding the methods in [4], which depends upon the FCN technique. Here, a spatial dependence among the neighboring slices was assumed under the application of the second method like Random Forest (RF) or RNN. Moreover, CNNs are applied to compute the entire CT scan or applied an interpolation layer. Prevedello et al. [5] introduced 2 models according to CNNs. The first method is concentrated on predicting ICH, mass effect, and hydrocephalus while scanning, whereas the second one was deployed for examining the malicious acute infarcts.

Chilamkurthy et al. [6] implied 4 methods for predicting the sub-types of ICH, calvarial fractures, midline shift, and mass impact. Here, the model was trained and verified under the large-scale dataset with massive CT scans, correspondingly. In this study, 2 datasets have been employed for testing. A medical radiology report is applied as gold-standard for labeling the training and verification of CT scans. Such reports are utilized for labeling every scan and using the natural language processing technique. The testing scan undergoes annotation by a major number of ICH sub-types addressed by professional radiologists. A distinct deep method is deployed to 4 prediction classes. ResNet18 was equipped with 5 parallel FC layers as output layers. Finally, models relied on CNN with RNN have been projected for ICH prediction. Grewal et al. [7] presented a 40-layer CNN, termed as DenseNet, through the bidirectional long short-term memory (LSTM) layer to ICH prediction. Moreover, 3 auxiliary operations were developed after dense convolutional block for computing binary segmentation of the ICH area. These operations are composed of a convolution layer, deconvolution layer for upsampling the feature maps to actual image size.

Ye et al. [8] implied a 3D joint CNN-RNN for detecting and classifying ICH sites. VGG-16 has been utilized as CNN approach, and the bidirectional Gated Recurrent Unit (GRU) has been employed as an RNN approach. This RNN layer performs a similar action with a slice interpolation model; however, it is flexible interms of adjacent slices used in the classification. Among the 3 methods, the CNN approach was extended for computing CT slices simultaneously. Jnawalia et al. [9] developed an ensemble of 3 diverse CNN approaches for proceeding the ICH forecasting. Some of the CNN schemes are relied on the structures of AlexNet and GoogleNet which is modified into a 3D model under the consideration of CT slides. Moreover, the least count of variables was applied by limiting the number of layers as well as filters. Chang et al. [10] established a DL model for detecting ICH and the subtypes (excluding IVH) with the capability of segmenting ICH regions and quantify the ICH quantity. It is relied on a region-of-interest (ROI) CNN which evaluates the regions with ICHs to 5 CT slices and produces segmentation masks to those sites.

Though several models are available in the literature, there is still a need to develop ICH diagnosis model using effective segmentation and DL architectures. This paper designs a novel DL-based ICH diagnosis and classification (DL-ICH) model by optimal image segmentation with Inception Network. The proposed DL-ICH model comprises preprocessing, segmentation, feature extraction, and classification. Initially, the input data perform format conversion where the NIfTI files are converted into JPEG format. Then, image segmentation takes place using kapur's thresholding with elephant herd optimization (EHO) algorithm called KT-EHO. Afterward, DL based Inception v4 network is employed as a feature extractor to extract a collection of features and multilayer perceptron (MLP) is used for classification. Extensive set of simulations takes place to ensure the effective diagnostic performance of the DL-ICH model.

2 The proposed DL-ICH model

The working process involved in the proposed DL-ICH model is illustrated in Fig. 1. As shown in the figure, the input data are preprocessed in the initial phase to convert the data into a compatible JPEF format. Besides, KT-EHO algorithm is applied as an image segmentation technique to indicate the diseased portions. Followed by, Inception v4 is applied to identify and classify the ICH. At last, MLP is utilized as a classifier to detect the multiple classes of ICH.

2.1 KT-EHO algorithm based image segmentation

Kapur's image segmentation was deployed for gray-scale image segmentation under the application of histogram entropy [11]. Here, T is a parameter that maximizes the entire entropy. Assume, $T = [t_1, t_2, \ldots, t_{k-1}]$ implies a vector of image thresholds. Afterward, Kapur's entropy is;

$$J_{max} = f_{kapur}(T) = \sum_{j=1}^{k} H_j^C \; for \; C\{1, 2, 3\} \tag{1}$$

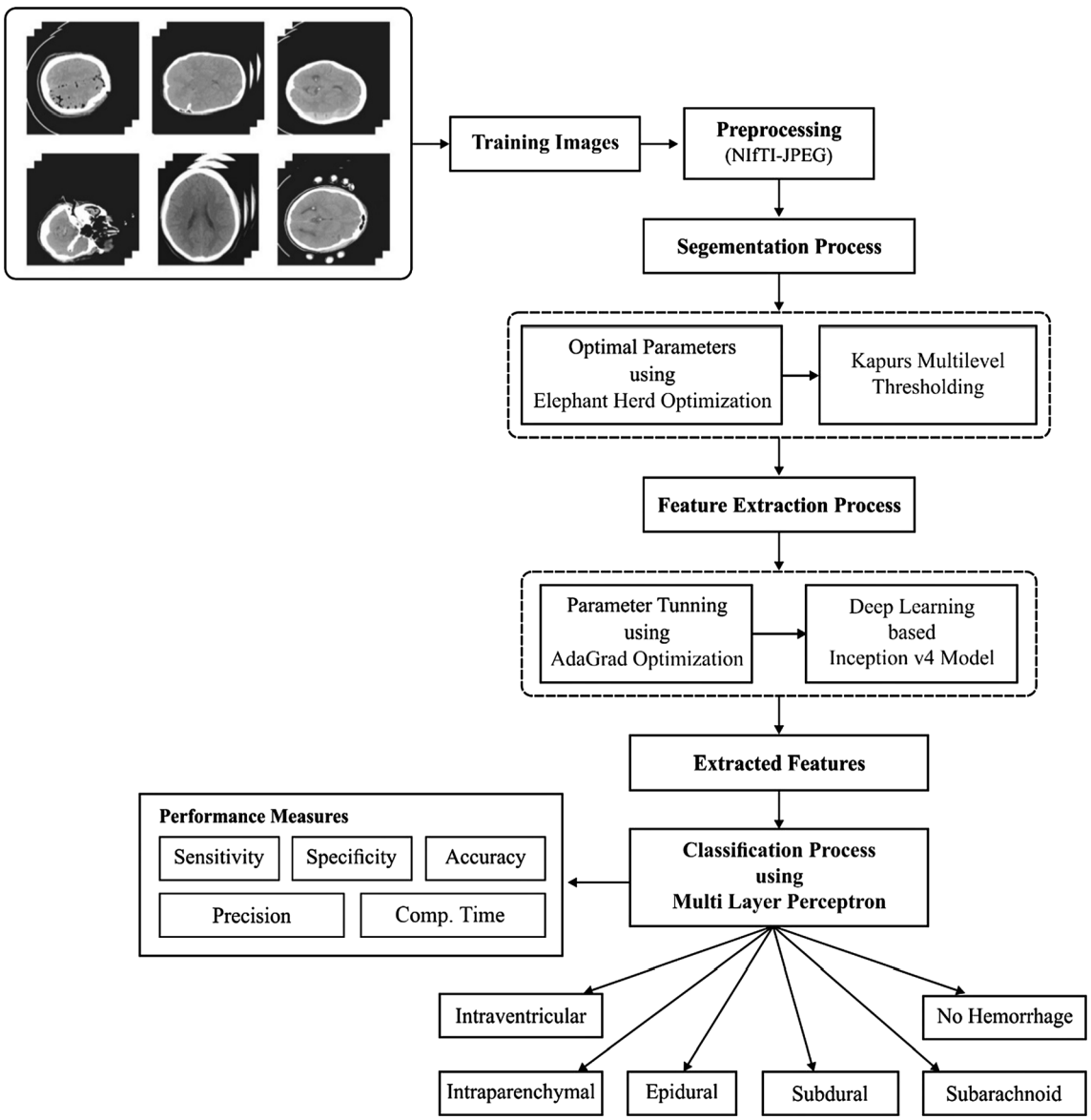


Fig. 1 Block diagram of DL-ICH model

In general, all entropy's is calculated independently depends on the specific t value. To multi-level thresholding problem, it is written as;

$$\begin{aligned} H_1^C &= \sum_{j=1}^{t_1} \frac{Ph_j^C}{\omega_0^C} \ln\left(\frac{Ph_j^C}{\omega_0^C}\right), \\ H_2^C &= \sum_{j=t_{l+1}}^{t_2} \frac{Ph_j^C}{\omega_1^C} \ln\left(\frac{Ph_j^C}{\omega_1^C}\right), \\ &\vdots \\ H_k^C &= \sum_{j=t_{k+1}}^{t_1} \frac{Ph_j^C}{\omega_{k-1}^C} \ln\left(\frac{Ph_j^C}{\omega_{k-1}^C}\right) \end{aligned} \tag{2}$$

where Ph_j^C is the possibility of distribution of the intensity levels and $\omega_0^C, \omega_1^C, .., \omega_{k-1}^C$ possibility occurrence for k levels.

To determine the optimal threshold value, EHO algorithm is used. In general, an elephant is a social animal which always stays in groups with diver's clans of Female Elephants (FE) and small calves. It is ruled by a leader elephant termed as matriarch, in which the movement of an elephant is monitored by a dominant elephant. Also, FE resides in teams, whereas the Male Elephant (ME) lives alone once they become adults and communicate with a family member with lower frequency signals. Some of the considerations made in the herding are given in the following:

- The group of elephants is divided into clans in which all clans have a certain number of elephants.
- A ME lives separately from the clan.
- Each clan is ruled by a matriarch.

The set of matriarchs are composed of the best solution, while the poor solution is decoded from a group of ME. Hence, the step of EHO is updated as given in the following:

In this approach, the place of an elephant is determined by excluding the matriarch and ME holds a supreme and inferior result correspondingly. For clan C_i elephants; a clan is comprised of 'P' elephants. The place of ith elephant $i = 1, 2, \ldots E$ also jth clan $j = 1, 2, \ldots 1$ is represented by $G_{i,j}$.

2.1.1 Movement update of fittest elephant of each clan

In elephants that move away from a group is utilized for performance validation. A clan with elephants of negative evaluations is converted into a novel position.

2.1.2 Separating the worst elephants in the clan

The ME resides individually by leaving the family. In this perception, bits are changed randomly as multiplication has been performed under the application of random measure. If the possibility of randomness is offered, then value has to be estimated. where, $G_{newli,j}$ implies the maximized position, $G_{li,j}$ denotes the former position, and $G_{bestli,j}$ refers to the location of best in the clan. Followed by, features are determined with '1', so-called as subset of better features. These best features are fed into a classifier which improves the precision and performance efficiency of the newly developed approach.

Algorithm 1 EHO Algorithm

(i) Initialization process

Select the feature attributes from the given database, EH parameters are scale factor and random values.

(ii) Fitness evaluation

Best FS task considers the accuracy as FF

(iii) Location of a novel elephant

$$G_{new,l_{i,j=}} G_{ci,j} + \alpha\left(G_{best,li,j} - G_{li,j}\right) \times r$$

(iv) Movement upgrade

A position enhancement to best fit in a clan is depicted by

$$G_{new,t_{i,j}} = \beta \times G_{center,l_j} \text{ And } G_{center,l_j} = \sum_{i=j}^{n} G_{t_{i,j}} / nl$$

(v) Separate worst position

The inferior position is upgraded as

$$G_{worst,l_{ij}} = G_{\min} + (G\max - G_{\min} + 1) \times r$$

(vi) Memorize the better result

(vii) Exit if

The best solution with the best accuracy is accomplished.

else

Iteration =Iteration +1.

α and β, And $r\varepsilon$ 0to 1, $nl \rightarrow$ Entire count of elephants in a clan,

$G_{worst,ci,j} \rightarrow$Poor ME in the clan

$G\max$ and $G_{\min} \rightarrow$ Maximal and minimal adaptable boundary limits.

2.2 Feature extraction

Once the input image is segmented, the feature extraction process begins. Here, the Inception v4 optimized by Adagrad (AG) optimizer is used for extracting desired feature vectors in order to perform ICH classification. This can be accomplished by Inception v4 and Adagrad optimizer, as defined in the following:

2.2.1 Overview of CNN

CNN model is referred to as supervised learning with major advantages and the least count of parameters [12]. It is also composed of effective training speed when compared with deep ANN. It is clear that massive benefits are projected in image segmentation, forecasting, and classification. A feature map of the initial layer is achieved by convolving the input, and convolution kernels have been employed. In all convolutions, kernel contains a size of 5×5 and stride as 1. The size of a feature map is attained using the provided function:

$$n_f = \frac{n_i + 2p - f}{s} \tag{3}$$

where n_f showcases the size of a feature map; n_i implies the input size; p means a padding value; f denotes the kernel size and s depicts the stride value. The basic function of a convolution operation is given below:

$$a^l = \delta\left(W^l a^{l-1} + b^l\right) \tag{4}$$

where a^1 refers to the result of lth convolution layer; w^l implies convolution kernel of lth convolution layer; a^{l-1} denotes the outcomes of $l-1$ th convolution layer; b^l signifies bias of lth convolution layer; δ displays the activation function of lth convolution layer. For sub-sampling, pooling process was utilized. In particular, activation functions like sigmoid, tanh, and ReLU have been employed prominently. The features are gained by applying convolution layer and send it to FC layer at the last layer. Followed by, CNN contains different layers with certain features such as convolution layers, activation layers, pooling layers, FC layers and SoftMax layers. Figure 2 shows a procedure on CNN.

Initially, the Convolution layer has a filter which convolves the width and height of an input data. Followed by, the simulation outcome of a convolution layer is achieved by a dot product among the filter weight content and exclusive position of an input image. Finally, it offers 2D activation map which offers a response of filter at the spatial position. Some of the parameters are number of filters, filter size, weight contents, stride, and padding. Actually, the pooling layer wishes to eliminate the overfitting and by executing nonlinear down-sampling on activation maps, the dimension and process difficulty has been reduced to improve the model's efficiency.

In different parameters have been applied and some of them are filter size and stride, but padding is not utilized in pooling, as the dimension reduction is not a major operation of this model. Moreover, it is applied to every input channel. Likewise, the number of output as well as input channels is symmetric. In addition, 2 different pooling layers have been applied like max and average pooling. The performance of the pooling layer and convolution layer is similar. The only difference is that, rather than using a dot product of input and a filter, the maximum neighboring value could be employed to accomplish an exclusive position in input image. During the Average pooling layer, average values are used from the location of the input image. Followed by, the FC layer is termed as hidden layer that has been used in normal NN. Before computing this, the input array has to be converted into a 1D vector utilizing a flattening layer. Thus, each node of an input is connected with one another. This type of activation function is utilized in a resultant layer of CNN for achieving a categorical distribution over labels that offers the possibilities of inputs.

2.2.2 Inception v4 model

In order to achieve a better performance, previous CNN modules such as LeNet, AlexNet, and VGGNet move deeper and deeper by piling the convolution layers. This method ensures a massive count of trainable parameters and results in over-fitting, in particular the datasets, in which trainable samples are minimum. Inception networks v1, v2–v3, and v4 are effective models used to object prediction as well as classification based operations. Such networks are applied extensively and encapsulate various kernel size filters for predicting the salient features from objects with major difference in size for salient portions. Figure 3 illustrates the inception modules and inception models with dimensionality reduction [13].

Inception-v4 network is an extended version of v1, v2, and v3 systems with respect to simplicity and effective module. Inception networks have various Inception modalities that are composed of clusters with factorized convolution layers, convolution layer, and pooling layers.

Factorized as well as convolution layer is advantageous for dimensionality reduction and enhances the processing effectiveness of Inception systems. Here, Inception-v4 has been applied for presented classification operation of ICH. Figure 4 depicts the network formation of Inception v4 method [14].

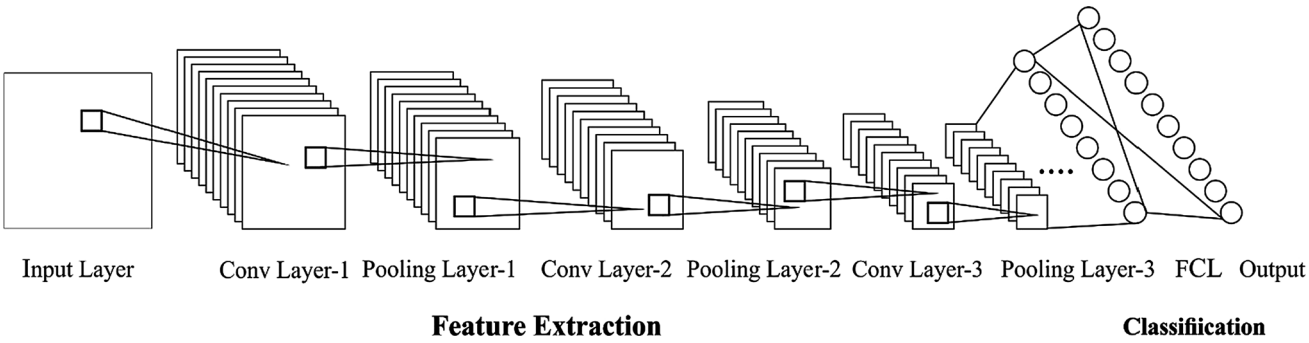


Fig. 2 Structure of CNN

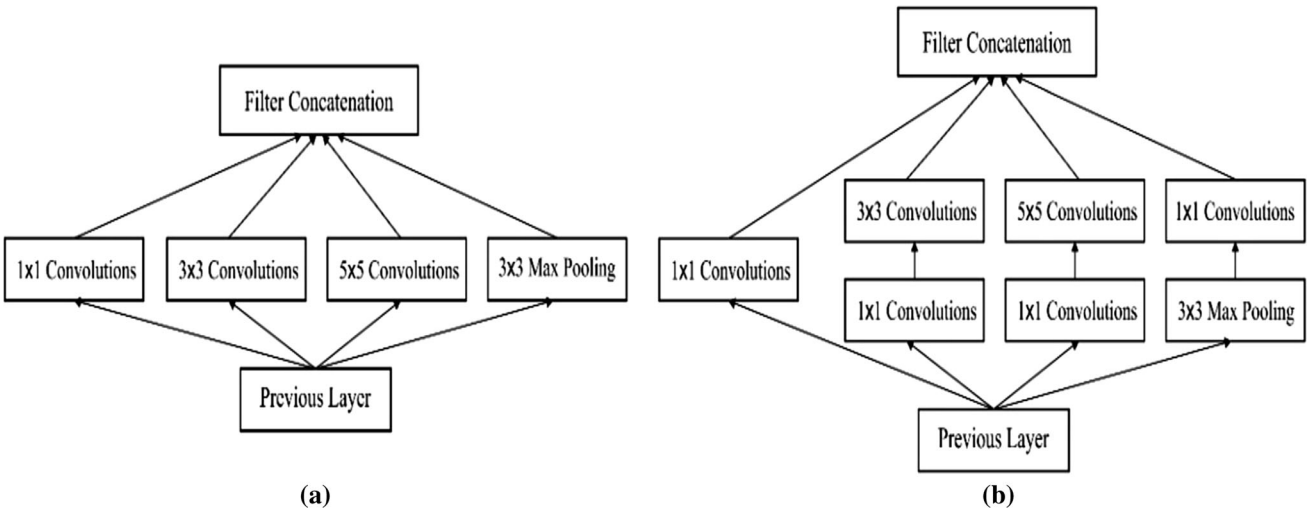


Fig. 3 Structure of inception network: **a** inception module and **b** inception module with dimensionality reduction

2.2.3 Hyperparameter optimization

Hyper-parameter is a kind of parameter used for manual selection and model training. The norm 'hyper-' is utilized for differentiating hyper-parameter to the parameter that is changed independently by applying optimization methods in training phase. The rate of learning implies the availability of parameter (θ) which has to be followed in opposite direction of a gradient estimate (g). Therefore, the hyper-parameter is highly complicated for fixation while the size is small and parameter update is minimum which consumes maximum time to accomplish moderate loss and if the fixed size is higher, then the parameter is shifted over the function and it does not achieve the considerable loss. But, higher-dimensional non-convex nature of NN optimization intends to make various sensitivity to every dimension. Next, a learning rate becomes lower for few dimensions and it becomes higher in additional dimension. Therefore, the classical model has been used to show the issues to DNN that is an AG technique [15].

AG is meant to be a gradient-based optimization that uses rate of learning to parameters, and computes the parameters with frequent features, while massive updates to parameters correlated with random features. To this purpose, it can be suitable for overcoming the sparse data. Moreover, it can be capable of enhancing the proficiency of SGD and applied for training NN. Thus, AG optimizer is named as gradient centric optimization model that is applicable to sparse gradients. Hence, a rate of learning was used for the parameters independently.

$$\Theta_{t+1} = \Theta_t - \frac{\alpha}{\sqrt{\varepsilon + \sum g_t^2}} \odot g_t \tag{5}$$

The basic function utilized for parameter improvement is expressed in Eq. (5) where θ_t implies the parameter at time t, α refers to the rate of learning, g_t denotes the gradient value, and $\odot$ represents the element wise multiplication.

2.3 MLP-based classification

MLP is composed of 3 layers such as input, hidden, and output layers. This module is applicable to capture numerous hidden layers by enabling the network for

Fig. 4 Network structure of inception v4

Input Image	Input Shape	(50,50, 1)
Stem	Input shape	(50,50, 1)
	Output shape	(40,40, 384)
Inception A	Input shape	(40,40, 384)
	Output shape	(40,40, 384)
Reduction A	Input shape	(40,40, 384)
	Output shape	(38,38, 1024)
Inception B	Input shape	(38,38, 1024)
	Output shape	(38,38, 1024)
Reduction B	Input shape	(38,38, 1024)
	Output shape	(36,36, 1536)
Inception C	Input shape	(36,36, 1536)
	Output shape	(36,36, 1536)
Average Pooling	Input shape	(36,36, 1536)
	Output shape	(4,4, 1536)
Dropout	Input shape	(4,4, 1536)
	Output shape	(4,4, 1536)
Flatten	Input shape	(4,4, 1536)
	Output shape	(24576)
MLP	Input shape	(24576)

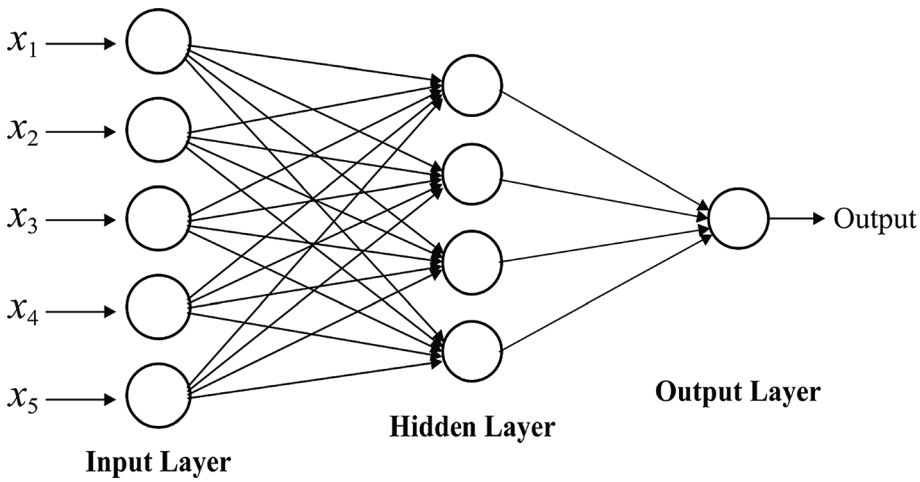
generating system outcomes. Figure 5 refers to an MLP network with a hidden layer, with limited weights that connect all layers.

The consequent result is estimated by the applied principles. At the initial stage, the inclusion of weights is evaluated by the given expression:

$$S_j = \sum_{i=1}^{n} w_{ij} x_i + \beta_i, \tag{6}$$

where x_i implies the input variable, w_{ij} describes the weight among input variable x_i and neuron j, and β_i illustrates the input variable's bias. Followed by, neurons' values of hidden layers are created from the values attained from weighted summation (Eq. (6)) by utilizing the activation function [16]. The popular option of such functions is

Fig. 5 Layers in MLP



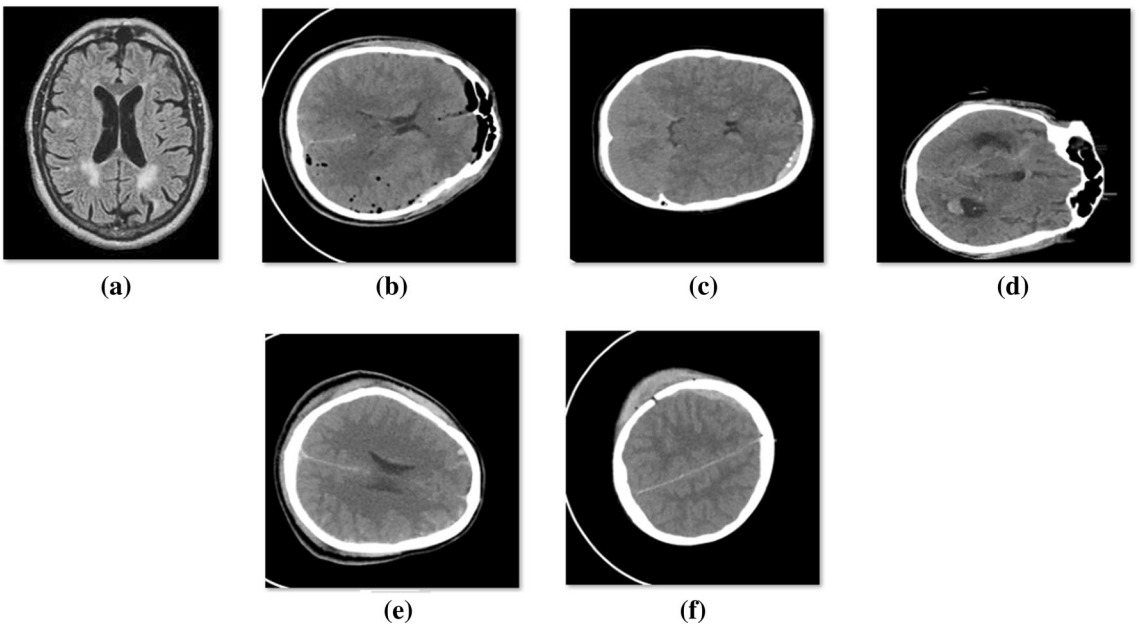


Fig. 6 Sample images **a** no hemorrhage, **b** epidural, **c** intraventricular, **d** intraparenchymal, **e** subdural and **f** subarachnoid

Table 1 Result analysis of proposed DL-ICH method on various epoch size

No. of epochs	Sensitivity	Specificity	Precision	Accuracy
Epoch-100	92.67	94.81	94.42	94.10
Epoch-200	93.71	95.72	95.31	95.22
Epoch-300	92.99	94.97	94.76	94.54
Epoch-400	93.89	95.98	95.69	95.39
Epoch-500	94.52	96.34	96.10	96.03

referred to be a sigmoid function as represented in the following:

$$f_j(x) = \frac{1}{1 + e^{-S_j}}, \tag{7}$$

where f_j denotes the sigmoid function to neurons j and S_j means the summation of weights. Eventually, the outcome of a neuron j is computed by the function:

$$y_j = \sum_{i=1}^{k} w_{ij} f_j + \beta_j, \tag{8}$$

where y_j implies the final outcome of a neuron j, w_{ij} represents the weight from resultant variable y_i and neuron j, f_j means the activation function to neuron j, and β_i illustrates the last variable's bias term.

3 Performance validation

The presented technique is implemented in a PC i5-8600 k processor, GeForce 1050Ti, 4 GB RAM, 16 GB OS Storage and 250 GB SSD File Storage. To simulate the presented technique, Python 3.6.5 tool is utilized with few packages namely tensorflow, keras, numpy, pickle, matplotlib, sklearn, pillow, and opencv-python. In order to estimate the performance of the presented model, an experimental analysis takes place using benchmark ICH dataset [17]. The dataset is gathered from the CT scans of 82 persons under the age group upto 72 years. In addition, the dataset involves images under six classes, namely Intraventricular with 24 slices, Intrarparenchymal with 73 slices, Subdural with 56 slices, Subarachnoid with 18 slices, and No Hemorrhage with 2173 slices. For experimental validation, tenfold cross validation process is used to split the dataset into training and test part. Figure 6 shows the sample set of images under different classes of ICH. The results are examined in terms of four measures, namely sensitivity, specificity, accuracy and precision. For comparison purposes, a set of methods used for comparison are U-Net, Watershed Algorithm with ANN (WA-ANN) [18], ResNexT [19], Window Estimator Module to a Deep Convolutional Neural Network (WEM-DCNN) [20], CNN and SVM models.

Table 1 and Fig. 7 illustrate the analysis of the results by the DL-ICH model under varying epoch count. By determining the classifier results under the epoch count of 100, the DL-ICH model has reached a higher sensitivity of

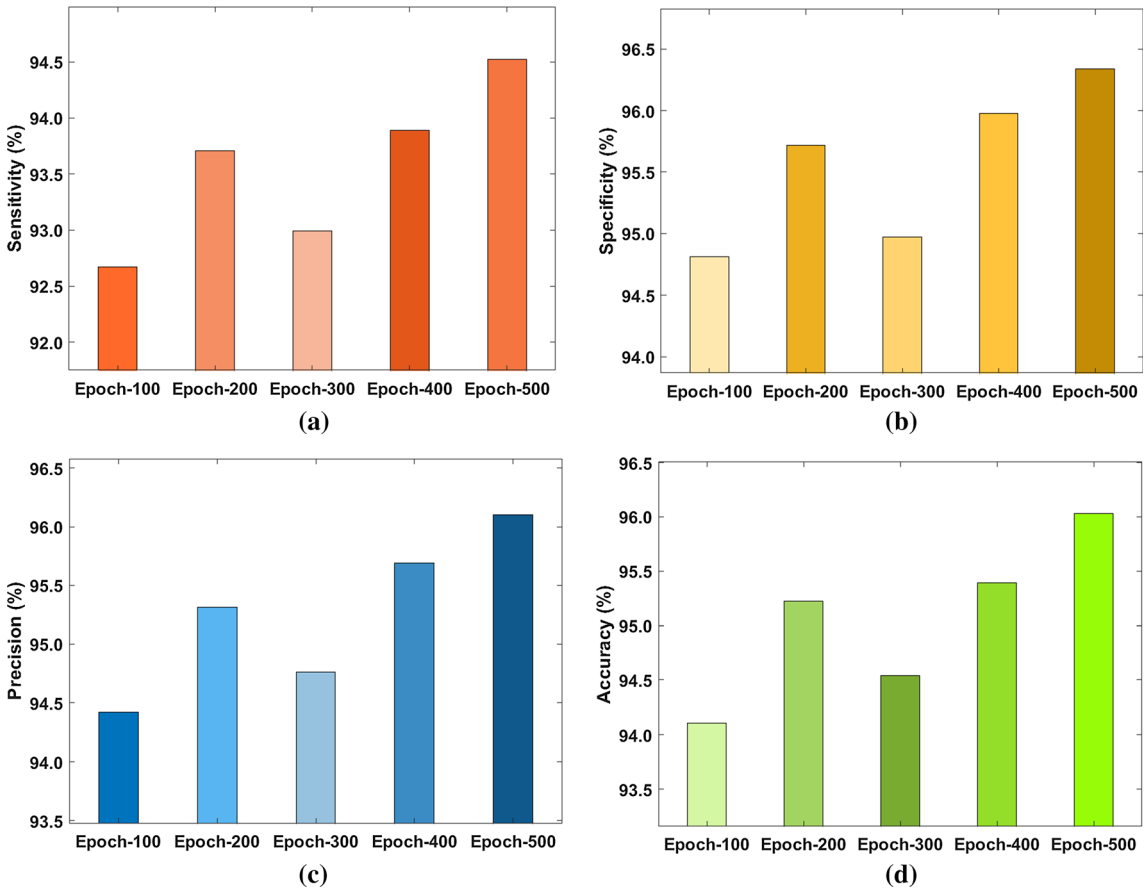


Fig. 7 Results analysis of DL-ICH model: **a** sensitivity, **b** specificity, **c** precision and **d** accuracy

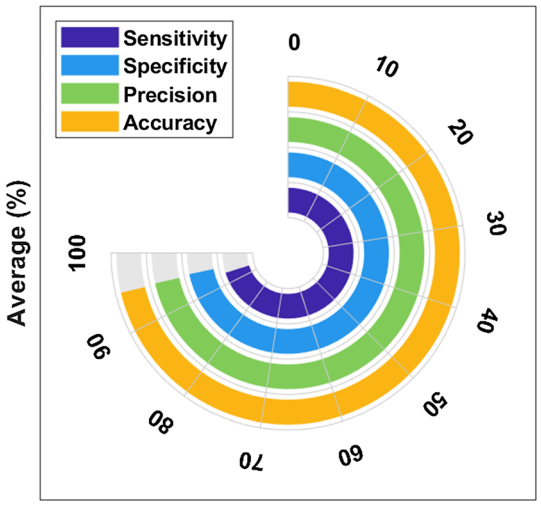


Fig. 8 Average results analysis of DL-ICH model

92.67%, specificity of 94.81%, precision of 94.42%, and accuracy of 94.10%.

At the same time, under the epoch count of 200, the DL-ICH model has achieved effective results with a sensitivity of 93.71%, specificity of 95.72%, precision of 95.31% and accuracy of 95.22%. Concurrently, under the epoch count of 300, the DL-ICH model has attained superior performance with a sensitivity of 92.99%, specificity of 94.97%, precision of 94.76%, and accuracy of 94.54%. Simultaneously, under the epoch count of 400, the DL-ICH model has exhibited proficient results with a sensitivity of 93.89%, specificity of 95.98%, precision of 95.69% and accuracy of 95.39%. Furthermore, the DL-ICH model has led to the optimal results with the sensitivity of 94.52%,

Table 2 Result analysis various measures on proposed DL-ICH with existing methods

Methods	Sensitivity	Specificity	Precision	Accuracy
DL-ICH	93.56	97.56	95.26	95.06
U-Net	63.10	88.60	88.19	87.00
WA-ANN	60.18	70.13	70.08	69.78
ResNexT	88.75	90.42	95.20	89.30
WEM-DCNN	83.33	97.48	89.90	88.35
CNN	87.06	88.18	87.98	87.56
SVM	76.38	79.41	77.53	77.32

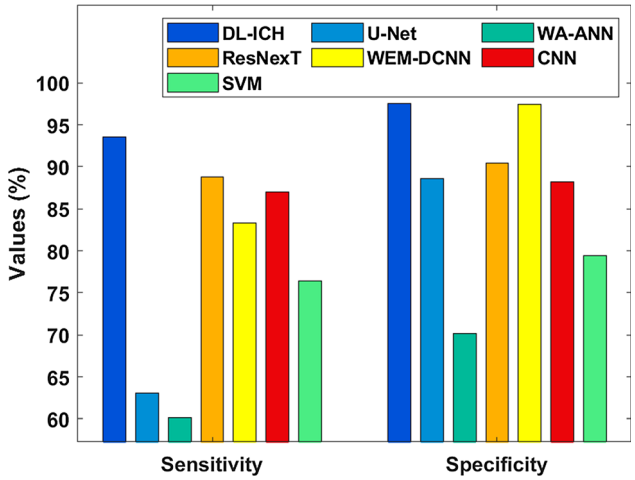


Fig. 9 Comparative analysis of DL-ICH method interms of sensitivity and specificity

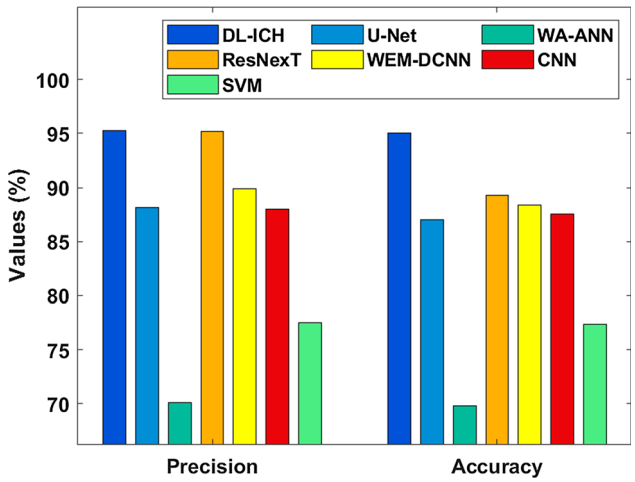


Fig. 10 Comparative analysis of DL-ICH method interms of precision and accuracy

specificity of 96.34%, precision of 96.10% and accuracy of 96.03%.

Figure 8 showcases the average accuracy analysis of the DL-ICH technique interms of different performance measures. The figure portrayed that the DL-ICH model has demonstrated optimal results with the average sensitivity of 93.56%, specificity of 95.56%, precision of 95.26% and accuracy of 95.06%, respectively.

A detailed comparative classification performance of the DL-ICH with existing methods is studied in Table 2. Figure 9 examines the sensitivity and specificity analysis of the DL-ICH with existing models on the applied set of images. The figure showcased that the WA-ANN model has failed to exhibit acceptable results with the sensitivity of 60.18% and specificity of 70.13%. In addition, the U-Net model showcased a somewhat higher sensitivity of 63.1% and specificity of 88.6%.

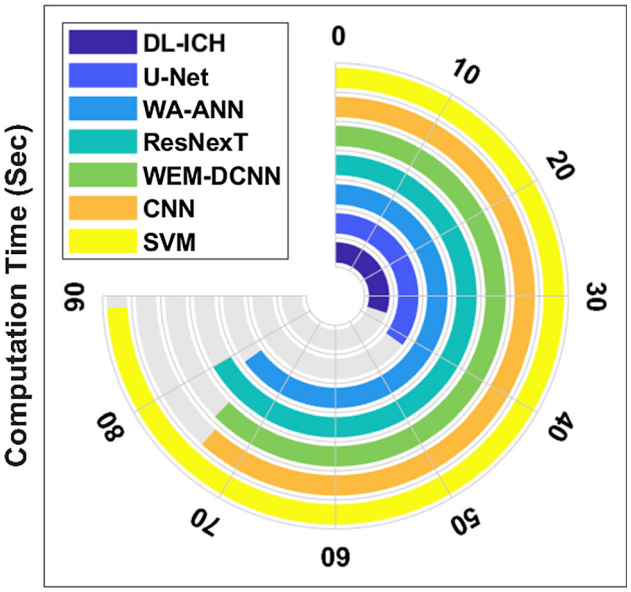


Fig. 11 CT analysis of DL-ICH with existing methods

Along with that, the SVM model has resulted to certainly better results with the sensitivity of 76.38% and specificity of 79.41%. Eventually, the WEM-DCNN model has tried to showcase moderate classification performance with the sensitivity of 83.33% and specificity of 97.48%. Concurrently, the CNN model has reached to manageable results with the sensitivity of 87.06% and 88.18%. Though the ResNexT model has demonstrated competitive performance with the sensitivity of 88.75% and specificity of 90.42%, the presented DL-ICH model has outperformed the previous models with the sensitivity of 93.56% and specificity of 97.56%.

Figure 10 investigates the precision and accuracy analysis of the DL-ICH with compared approaches on the sample test images. From the figure, it is noticed that the WA-ANN model has unsuccessful to display satisfactory performance with the precision of 70.08% and accuracy of 69.78%. Additionally, the SVM model showed rather higher precision of 77.53% and accuracy of 773.32%. Also, the CNN model has occasioned to certainly improved results with the precision of 87.98% and accuracy of 87.56%. Ultimately, the U-Net model has tried to show reasonable results with a precision of 88.19% and accuracy of 87. Simultaneously, the WEM-DCNN model has reached to controllable results with the precision of 89.9% and 88.35%. Though the ResNexT model has established competitive performance with the precision of 95.2% and accuracy of 89.3%, the presented DL-ICH model has outperformed the previous models with the precision of 95.26% and accuracy of 95.06%.

Table 2 and Fig. 11 examine the computation time (CT) analysis of the DL-ICH with a set of existing models. From

the figure, it is apparent that the ResNexT and SVM models have required a higher CT of 70 s and 89 s, respectively. Besides, the WA-ANN, WEM-DCNN and CNN models have needed slightly lower CT of 78%, 75% and 74%, respectively. Moreover, the U-Net model has tried to show near optimal CT of 42 s, whereas the presented DL-ICH model has reached a minimum CT of 36 s. From the above-mentioned experimental results, the DL-ICH model can be employed as an appropriate tool for ICH diagnosis. The superior performance is due to the following characteristics, optimal threshold selection using EHO algorithm, Inception v4 based feature extraction, and MLP based classification.

4 Conclusion

In this paper has developed a novel deep DL based ICH diagnosis and classification named DL-ICH model with optimal image segmentation with Inception Network. The proposed DL-ICH model comprises preprocessing, segmentation; feature extraction, and classification. The input data are preprocessed in the initial phase to convert the data into a compatible JPEG format. Besides, the KT-EHO algorithm is applied as an image segmentation technique to indicate the diseased portions. Finally, Inception v4 is applied to identify and classify the ICH. To evaluate the performance of the presented DL-ICH model, a comprehensive set of simulations takes place to indicate the superiority of the presented model. The obtained simulation outcome portrayed that the KT-EHO algorithm has showcased outstanding results with the sensitivity of 93.56%, specificity of 97.56%, precision of 95.25%, and accuracy of 95.06%. In future, the performance of the DL-ICH method is improvised using hyperparameter tuning approaches.

Funding The authors received no funding for this work.

Declarations

Conflict of interests The authors declare there are no conflict of interests.

References

1. Taylor CA, Bell JM, Breiding MJ, Xu L (2017) Traumatic brain injury-related emergency department visits, hospitalizations, and deaths-United States, 2007 and 2013. Morb Mortal Wkly Rep SurveillSumm 66:1–16
2. Yuh EL, Gean AD, Manley GT, Callen AL, Wintermark M (2008) Computer-aided assessment of head computed tomography (CT) studies in patients with suspected traumatic brain injury. J Neurotrauma 25:1163–1172
3. Li Y, Wu J, Li H, Li D, Du X, Chen Z, Jia F, Hu Q (2012) Automatic detection of the existence of subarachnoid hemorrhage from clinical CT images. J Med Syst 36:1259–1270
4. Kuo W, Häne C, Yuh E, Mukherjee P, Malik J (2018) Cost-sensitive active learning for intracranial hemorrhage detection. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G (eds) Medical image computing and computer assisted intervention—MICCAI 2018. Springer, Cham, pp 715–723
5. Prevedello LM, Erdal BS, Ryu JL, Little KJ, Demirer M, Qian S, White RD (2017) Automated critical test findings identification and online notification system using artificial intelligence in imaging. Radiology 285:923–931
6. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, Mahajan V, Rao P, Warier P (2018) Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet 392:2388–2396
7. Grewal M, Srivastava MM, Kumar P, Varadarajan S (2018) RADnet: radiologist level accuracy using deep learning for hemorrhage detection in CT scans. In: Proceedings of the 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018), Washington, DC, USA, 4–7 Apr 2018, pp 281–284
8. Ye H, Gao F, Yin Y, Guo D, Zhao P, Lu Y, Wang X, Bai J, Cao K, Song Q et al (2019) Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. EurRadiol 29:6191–6201
9. Jnawali K, Arbabshirani MR, Rao N, Patel AA (2018) Deep 3D convolution neural network for CT brain hemorrhage classification. In: Medical imaging 2018: computer-aided diagnosis; international society for optics and photonics: Washington, DC, USA, vol 10575, p 105751C
10. Chang P, Kuoy E, Grinband J, Weinberg B, Thompson M, Homo R, Chen J, Abcede H, Shafie M, Sugrue L et al (2018) Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT. Am J Neuroradiol 39:1609–1616
11. Rajinikanth V, Satapathy SC, Fernandes SL, Nachiappan S (2017) Entropy based segmentation of tumor from brain MR images—a study with teaching learning based optimization. Pattern RecognLett 94:87–95
12. Karolin M, Meyyappan T (2019) Secret multiple share creation with color images using visual cryptography. In: 2019 International conference on communication and signal processing (ICCSP). IEEE, pp 0058–0062
13. https://towardsdatascience.com/review-inception-v4-evolved-from-googlenet-merged-with-resnet-idea-image-classification-5e8c339d18bc
14. Tekchandani H, Verma S, Londhe N (2020) Performance improvement of mediastinal lymph node severity detection using GAN and Inception network. Comput Methods Programs Biomed 194:105478
15. Lydia A, Francis S (2019) Adagrad—an optimizer for stochastic gradient descent. Int J Inf Comput Sci 6(5)
16. Heidari AA, Faris H, Aljarah I, Mirjalili S (2019) An efficient hybrid multilayer perceptron neural network with grasshopper optimization. Soft Comput 23(17):7941–7958
17. Hssayeni MD, Croock MS, Salman AD, Al-khafaji HF, Yahya ZA, Ghoraani B (2020) Intracranial hemorrhage segmentation using a deep convolutional model. Data 5(1):14
18. Davis V, Devane S (2017) Diagnosis & classification of brain hemorrhage. In: 2017 International conference on advances in computing, communication and control (ICAC3). IEEE, pp 1–6
19. Danilov G, Kotik K, Negreeva A, Tsukanova T, Shifrin M, Zakharova N, Batalov A, Pronin I, Potapov A (2020)

Classification of intracranial hemorrhage subtypes using deep learning on CT scans. Stud Health TechnolInformat 272:370–373
20. Karki M, Cho J, Lee E, Hahm MH, Yoon SY, Kim M, Ahn JY, Son J, Park SH, Kim KH, Park S (2020) CT window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings. Artif Intell Med 101850.