

## Medical Image Data Classification Using Deep Learning Based Hybrid Model with CNN and Encoder



Bhanu Prakash Battula<sup>1\*</sup>, Duraisamy Balaganesh<sup>2</sup>

<sup>1</sup> Lincoln University College, Wisma Lincoln, No. 12-18, Jalan SS 6/12, Petaling Jaya 47301, Selangor Darul Ehsan, Malaysia

<sup>2</sup> Faculty of Computer Science and Multimedia, Lincoln University College, Wisma Lincoln, No. 12-18, Jalan SS 6/12, Petaling Jaya 47301, Selangor Darul Ehsan, Malaysia

Corresponding Author Email: [prakashbattula33@gmail.com](mailto:prakashbattula33@gmail.com)

<https://doi.org/10.18280/ria.340516>

### ABSTRACT

**Received:** 20 July 2020

**Accepted:** 6 October 2020

#### Keywords:

*CNN, encoder, medical images, classification*

Healthcare sector is one of the prime and different from other trade. Society expects high priority and highest level of services and care irrespective of money. Presently medical field suffers from accurate diagnosis of diseases and it create huge loss to society. The prime factor for this is due to the nature of medical data, it is a combination of all varieties of data. Medical image analysis is a key method of Computer-Aided Diagnosis (CAD) frameworks. Customary strategies depend predominantly on the shape, shading, and additionally surface highlights just as their mixes, a large portion of which are issue explicit and have demonstrated to be integral in medical images, which prompts a framework that does not have the capacity to make portrayals of significant level issue area ideas and that has poor model speculation capacity. In this paper we are attempting a medical image data classification technique using hybrid deep learning technique based on Convolutional Neural Network (CNN) and encodes. What's more, we assess the proposed approach on two benchmark clinical picture datasets: HIS2828 and ISIC2017. The proposed algorithm is applied on the considered 2 datasets for performing data classification using deep learning based CNN and encoders. The proposed model is compared with the traditional methods and the results show that proposed model classification accuracy is better than the existing models.

## 1. INTRODUCTION

The World Health Organization (WHO) estimated that around 34% of all global deaths in 2018 were due to misjudgement of patient medical record. Naming it the world's number one killer because it causes huge portion. Therefore, it is crucial and urgent to improve and accelerate all parts of clinical diagnosis, with the goal of widespread and early diagnosis of correct abnormalities across the population [1]. It is also a goal to make all diagnostic methods as little invasive as possible. Medical imaging has a key role in this framework, because, in order to uncover Cardiovascular Diseases (CVDs), medical practice makes heavy use of imaging scans to quantify cardiac function through segmentation. There are many medical imaging modalities, such as magnetic resonance imaging (MRI), computed tomography (CT), ultrasound (UT) and late gadolinium enhancement (LGE).

Computer-assisted diagnostics (CADE), also referred to as computer-assisted diagnostics (CADx), are devices that assist doctors in the analysis of medical pictures [2]. X-ray, MRI and ultrasound diagnostic imaging methods offer a range of knowledge to be examined and properly assessed by the radiologist or other expert in a limited period of time. Digital images are processed by CAD systems for typical appearances and highlighting notable sections, such as disease, so as to provide feedback to support professional decision-making. CAD is an interdisciplinary technology which combines elements of artificial intelligence and computer vision with

processing radiology and pathology [3].

LGE is the current standard for fibrosis detection, however the diagnosis depends on gadolinium-based contrast which comes at a risk [4]. This is because patients suffering from CVDs often have other diseases, such as severe renal dysfunction, that impair other organs. These dysfunctions can lead to potentially lethal adverse reactions to gadolinium-based contrast agent. It follows, that (native) T1 mapping in cardiovascular MRI has become a rapidly emerging modality which is increasingly popular as an alternative to LGE, due to it being contrast agent-free and considered low-risk [5]. However, post-processing of T1 mapping images is a complex multi-step process, affected by a range of sources of variability that directly affect the outcome of the predicted T1 value. This is because the estimation of T1 values is based on the delineation, or segmentation, of regions of interest (ROIs), and so far, most segmentations of T1 maps have been manual or semi-automatic.

There are efforts in the field of deep learning directed towards making the segmentation step fully automatic for such imaging scans [6]. They aim to develop accurate, automatic segmentation methods, in order to improve consistency and reproducibility in medical image analysis. Meanwhile, assessing the accuracy of the automatic segmentation methods already created can be difficult given the lack of a gold standard or ground truth to how the segmentation should be done [7]. However, this does not mean that there is no consensus in the field [8]. There are several guidelines to

ensure as accurate and similar segmentation as possible worldwide. This consensus is the base of all teaching and clinical practice. On the contrary, when it comes to the training of medical analyst's trainees, training is carried out at the level of a single center or medical unit, with little agreement across centers [9]. To put it differently, this means that there is some form of consistency across people undergoing training by the same expert; meanwhile, there is little agreement on the training process among experts [10]. Furthermore, the training process can also be a tedious and repetitive process where the instructor may have to sit with each trainee individually, repeating steps and correcting errors as they go along. Due to this, the standardization and digitization of the training process becomes an equally important issue as the automatic segmentation.

In the last few decades, the field of computer training has evolved tremendously and its area of use has developed from online to self-driving vehicles [11]. Time efficiency, an indigenous property of machine learning calculations, is one reason behind the broad area of operation [12]. One potential enthusiast of machine learning techniques is the social insurance market, where machines have proved unmistakably faster than human beings to process tolerant information [13].

One of the key benefits of machine learning is the ability to identify objects in images accurately and make them suitable for object detection and grading. Machine learning is divided into numerous branches, one of which is profound learning. Deep education approaches use artificial neural networks inspired by the human brain neural network [14]. A network can process several hundred pictures in seconds and thus become a valuable tool for long-term tasks. For example, a network in health care can be used to assess a diagnosis quicker than any human [15]. A deep learning solution could also be a tool for smaller health centres which lack clinical specialists [16].

The focus for this paper has been to identify kind of medical abnormalities in images. These abnormalities are located in various regions of human body [17]. Many deceases are diagnosed using medical images like x-ray, scans etc. so this system helps a doctor to diagnosis accurate issue of the patient which was held in patient [18].

The rest of the paper is organized as follows section-2 describes the state-of the art mechanisms, next section presents proposed methodology of medical image classifications, next section makes experimental evaluation and last section concludes the paper.

## 2. RELATED WORK

In a past work, Cruz-Roa et al. [1] suggested the segmentation of chest X-ray using convolutional neural network. In their work, they introduced image segmentation into bone tissue and non-bone tissue. The aim of their work was to develop an automatic or an intelligent segmentation system for chest X-rays. The system was established to have the capability to segment bone tissues from the rest of the image. This model considers the limited features only for the classification process.

They were able to achieve the aim of the research by using a convolutional neural network, which was tasked with examining raw image pixels and hence classifying them into "bone tissue" or "non-bone tissue". The convolutional neural networks were trained on the image patches collected from the chest X-ray images.

It was recorded in their work that the automatic segmentation of chest X-rays using the convolutional neural networks, and approaches suggested in their research produced plausible performance.

In another recent research, Patil and Kuchanur [19] presented the application of some image processing techniques in the classification of patients chest X-rays into whether cancer is present or not (benign or malignant). In this work, it was shown that by extracting some geometric features that are essential to the classification of the images such area, perimeter, diameter, and irregularity; an automatic classification system was developed.

Furthermore, in the same research, texture features were considered for a parallel comparison of results on the classification accuracy. The texture features used in the work are average gray level, standard deviation, smoothness, third moment, uniformity, and entropy. The back propagation neural network was used as the classifier, and an accuracy of 83% was recorded. The model considers only limited features and the dataset size is also less.

In this thesis the classification of chest X-ray radiographs into two classes has been achieved using artificial neural networks. The two classes are the normal one which has no disease or conditions, and the abnormal one which may have any types of diseases that may encounter the chest organs including heart, lungs. Chest etc. A deep network called Stacked Auto-encoder is used for (SAE), which relies on a supervised and unsupervised learning algorithm was used to train the network on the images collected for the research.

Zhang et al. [12] suggested methods for feature selection. The methods of feature selection are divided into three types: filter, wrapper and hybrid approach. Filter methods implement an independent test without a learning algorithm, while the wrapper approach utilises a pre-determined learning algorithm. Both approaches have their advantages and drawbacks for filter and wrapper. The filter approach is low, but less accurate, while the wrapper method has better accuracy with higher computation costs. The hybrid process incorporates the benefits of the first two processes.

Li et al. [14] suggest an overview of the dimension reduction based on data. An adaptive classification method is used to check the input data level. Eigen matrix and Eigen vector for dimensional reduction are found in the proposed PCA-based method. Due to the methodology suggested, redundant information from the original data input reduces the size of the data collection.

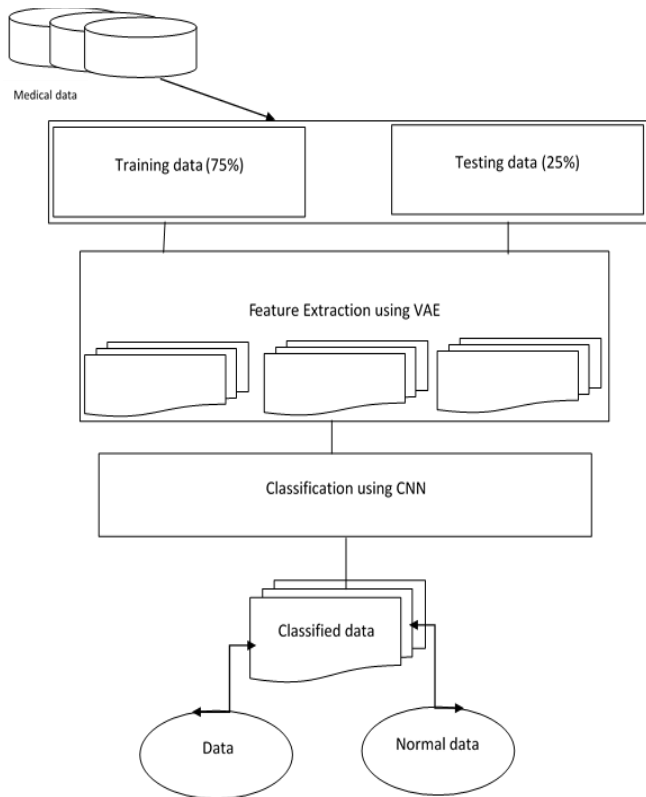
Bar et al. have [15] proposed an algorithm for the principal component analysis which was used to reduce the dimensions and assess the characteristics of stroke disease prediction and predict whether patients are affected or not by the stroke. Stroke was a significant cause of long-term, serious disability in the United States.

Shin et al. [16] suggested a hybrid approach to KNN and genetic algorithms to boost the precision of the classification of the data set on heart disease. The proposed method used genetic search to prune redundant or obsolete attributes and to find classification attributes. The least graded attributes are excluded and classification is conducted according to high classification.

## 3. PROPOSED WORK

Here mainly concentrate on medical image data

classification based on deep learning based mechanisms. Here we use CNN and Encoders in a combination to work with medical images to make better classification of medical image data. Below Figure 1 explains the proposed mechanism architecture.



**Figure 1.** Propose mechanism architecture

The representative capacity of functions extracted by a CNN depends in the time and variability captured in the training data set on the complexity of the detection problem. In real world applications, the presence of a target object is subject to major changes due to view angles, lighting, background disruptions and occlusions. Handcrafted features built to recognize the domain have also proved in certain instances to be more distinctive and reliable [20]. A smart fusion of the two feature modalities is expected to boost the detection efficiency. In this work we propose an improved CNN object detection system by combining handmade features with deep features in the embedding space. In conjunction with gradient and orientation histograms, we use fundamental colour canals: RGB, HSV and LBP to improve deeper features in overcoming prediction errors due to variations in appearance [21]. The proposed model performs feature extraction from the dataset and the classification technique using CNN is applied by selecting hidden layers for classification of data [22]. The encoders are used for the accuracy improvement and selecting only features that are relevant for accurate classification [23].

### 3.1 Features extraction in medicine

Pattern recognition is the process of developing systems that have the capability to identify patterns; while patterns can be seen as a collection of descriptive attributes that distinguishes one pattern or object from the other. It is the study of how machines perceive their environment, and

therefore capable of making logical decisions through learning or experience. During the development of pattern recognition systems, we are interested in the manner in which patterns are modeled and hence knowledge represented in such systems. Several advances in machine vision have helped revamp the field of pattern recognition by suggesting novel and more sophisticated approaches to representing knowledge in recognition systems; building on more appreciable understanding of pattern recognition as achieved in the human visual processing.

Derived from the important results of the above-mentioned methods, comprehensive research has specifically established deep-fusion networks, where intermediate multi-level layers are fused by new side branches. This helps us to combine the strengths of individual layers and produce superior prediction through a profoundly fused representation. While these high-performance fusion networks also need a large number of additional parameters to produce the side branches. Moreover, their fusion modules (for example, sum pooling) do not take into account the value of various side branches.

Typical pattern recognition as the following important phases for the realization of its purpose for decision making or identification.

- **Data acquisition:** This is the stage in which the data relevant to the recognition task are collected.
  - **Pre-processing:** It is at this stage that the data received in the data acquisition stage is manipulated into a form suitable for the next phase of the system. Also, noise is removed in this stage, and pattern segmentation may be carried out.
  - **Feature extraction/selection:** This stage is where the system designer determines which features are significant and therefore important to the learning of the classification task.
  - **Features:** The attributes which describe the patterns.
  - **Model learning estimation:** This is the phase where the appropriate model for the recognition problem is determined based on the nature of the application. The selected model learns the mapping of pattern features to their corresponding classes.
  - **Model:** This is the particular selected model for learning the problem, the model is tuned using the features extracted from the preceding phase.
  - **Classification:** This is the phase where the developed model is simulated with patterns for decision making. The performance parameters used for accessing such models include recognition rate, specificity, accuracy, and achieved mean squared error (MSE).
  - **Post-processing:** The outputs of the model are sometimes required to be processed into a form suitable for the decision making phase stage. Confidence in decision can be evaluated at this stage, and performance augmentation may be achieved.
  - **Decision:** This is the stage in which the system supplies the identification predicted by the developed model.
- There are many approaches to the pattern problem picture processing was commonly used in medicine. Improving the picture is often the most common process in this area. A multi-part medial picture produces a lot of noise. This makes it very difficult for doctors to identify it correctly. Image processing can be useful in this case because it helps detect and improve the images, since all the parts of the image, including the noise, vary in brightness and intensity from each other. This work

therefore uses image processing techniques to improve X-ray images in the chest and minimise the noise present therein. This is achieved using various techniques such as filtering, histogram balance and intensity correction for image enhancement. Figure 2 gives an example of the working theory of the proposed algorithm.

Many filters, such as median, mean, Gaussian filter, may be used for filtering. The images are screened for median filter, as some of them have noise artefacts to be eliminated to increase the quality of images. Median filter is good noise reduction technique as it gives good rejection of the noise of salt and pepper found in some medical photos.

Moreover, image intensities adjustment can be also used for enhancing the quality of images. This technique involves the mapping of the pixels intensity distribution from one level to another level. To highlight the images more and more, the intensities of pixels are increased by mapping them into other values. This ended up with brighter images where the cells are clearer; including the cancerous cells.

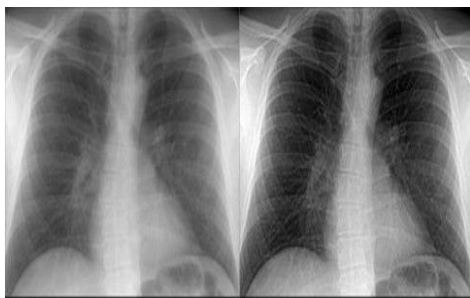


Figure 2. Medical image enhancement

### 3.2 Deep learning

Deep Learning is another and progressed documented of Machine Learning. It has been produced and enhanced keeping in mind the end goal to move the moving Machine Learning to be nearer its primary and unique objective Artificial Intelligence.

Deep Learning is called "Deep" because of its structure of the neural systems. Prior, neural systems used to have two layers deep on the grounds that it was not computationally plausible to assemble bigger systems. These days, a neural system with in excess of 10 layers and considerably more layers are being started and constructed. These sorts of systems are called deep neural systems. Figure 3 shows engineering of a deep neural system. It demonstrates that the system comprises of numerous layers which makes it deep.

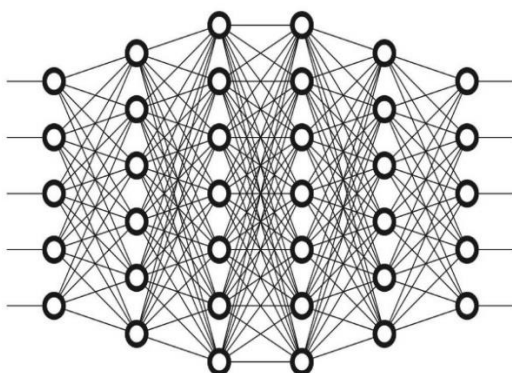


Figure 3. Deep network structure

Stacked auto-encoder is one sort of deep systems that is prepared utilizing another calculation called covetous layer wise training. The eager layer wise approach for pre-training a deep system works via training each layer thus. In this page, you will discover how auto-encoders can be "stacked" in an insatiable layer wise design for pre-training (instating) the weights of a deep system.

It is a neural system comprising of different layers of meager auto-encoders in which the outputs of each layer are wired to the inputs of the progressive layer. Formally, consider a stacked auto-encoder with n layers. Utilizing documentation from the auto-encoder area, let  $W1(k,1), W1(k,2), b1(k,1), b1(k,2)$  signify the parameters  $W1(1), W1(2), b1(1), b1(2)$  for kth auto-encoder. At that point the encoding venture for the stacked auto-encoder is given by running the encoding advance of each layer in forward request:

$$a^{(l)} = f(z^{(l)})$$

$$z^{(l+1)} = W^{(l,1)}a^{(l)} + b^{(l,1)}$$

The disentangling step is given by running the translating pile of every auto-encoder in switch request:

$$a^{(n+l)} = f(z^{(n+l)})$$

$$z^{(n+l+1)} = W^{(n-l,2)}a^{(n+l)} + b^{(n-l,2)}$$

The data of intrigue is contained inside  $a(n)$ , which is the actuation of the layer of hidden units. This vector gives us a representation of the input as far as higher-arrange highlights. The highlights from the stacked auto-encoder can be utilized for classification issues by nourishing  $a(n)$  to a Softmax classifier.

### 3.3 Training

A decent method to acquire great parameters for a stacked auto-encoder is to utilize ravenous layer-wise training. To do this, first prepare the primary layer on crude input to get parameters  $W1(1,1), W1(1,2), b1(1,1), b1(1,2)$ . Utilize the primary layer to change the crude input into a vector comprising of enactment of the hidden units, A. Prepare the second layer on this vector to acquire parameters  $W1(2,1), W1(2,2), b1(2,1), b1(2,2)$ . Rehash for ensuing layers, utilizing the output of each layer as input for the resulting layer.

In a grading problem, the aim is to learn about the decision surface, which accurately maps an input space into a class mark output space [1]. Various computer researchers have been trying to use various techniques in the field of medicine to improve the accuracy of data classification for provided data, classification techniques which better provide sufficient information in order to identify potential patients and thus improve the accuracy of diagnostic measures.

This technique exclusively prepares the parameters for each layer and solidifies parameters for the rest of the model. To create better results after this training time, fine tuning with backpropagation can be used to improve the results by changing the parameters of all the layers.

Should it be difficult to detect the "disentangling" layers of the stacked auto-encoder and to connect the last secret layer  $a(n)$  with the Softmax classifier, the usual procedure would be to detect the fine tuning for the classification reasons. The angles from the classification error are then scattered back to the

encoding layers.

### 3.3.1 Convolutional neural network (CNN)

It similar to artificial neural network, as both of them are made up of self-optimized neurons, which are imported by inputs and perform a non- linear transformation [15]. Compared with the artificial neural network, convolution neural network is widely used in pattern recognition on images, for it encodes image specific features into the network architecture, making the network more suitable for image-based feature learning [17].

There are five basic elements within the convolution neural network: input layer, convolutional layer, non-linear layer, pooling layer and fully connected layer. The basic functionality of a convolution neural network can be put into five key points.

Point 1: The input layer holds the pixel values of the image.

Point 2: The convolutional layer carries the convolution calculation of the inputs.

Point 3: The non-linear layer is used to apply non-linear transformation of inputs produced by the previous layer.

Point 4: The pooling layer performs sampling of the given inputs, reducing the number of parameters involved in the activation.

Point 5: The fully connected layer shows the class scores from the activations and they are used for classification.

### 3.3.2 Hybrid algorithm for medical image data classification

#### Pseudo code for Medical data classification using VAE and CNN

Input: Medical data

Output: Classified images i.e. disease and not disease;

Step-1: Take the input data set and split it into training and testing data

Step-2: while training

For each image in the Training Data:

    If it exists in the model:

        For each pixel in the Content pixels:

            Calculate the correlation between the pixel vectors.

Parent → set of pixels with Maximum Correlation

    Add the pixels to the Content Tree as the child of the

parent

Feature map → {Set of nodes in the content tree}

Apply CNN on the feature map

Return classified data

}

The above algorithm explains how medical image data classifies the medical images into diseased and normal images. Here mainly it focuses on extracting the features, major features are extracted by Encoders and minor features are extracted using CNN.

## 4. RESULTS AND DISCUSSIONS

We executed the coding system to separate the elevated level highlights in python Numpy, which is a matlab tool stash that actualizes convolutional neural systems, just as removed the conventional highlights dependent on shading minute and surface. We have planned a progression of tests to confirm the viability of our technique on two benchmark clinical picture datasets.

One is the HIS2828 dataset [15], and the other is the

ISIC2017 dataset [18]. We led the entirety of our analyses on a PC with i5-6500 3.2 GHz CPU, 32G principle memory, and GTX1060 GPU.

HIS2828 dataset is made out of 4 classes of major tissue pictures that are illustrative of various tissue types. Each picture is a RGB picture of size 720 \* 480. 2is dataset contains 2828 pictures, which can be recorded as follows: 1026 sensory tissue pictures, 484 connective tissue pictures, 804 epithelial tissue pictures, and 514 strong tissue pictures, in which we use 1, 2, 3, and 4 to speak to the marks. Table 1 shows the creation of the HIS2828 dataset.

**Table 1.** HIS2828 dataset

| Image Category    | Number of Images | Label |
|-------------------|------------------|-------|
| Nervous Tissue    | 1026             | 1     |
| Connective Tissue | 484              | 2     |
| Epithelial Tissue | 804              | 3     |
| Muscular Tissue   | 514              | 4     |

ISIC2017 is a dataset of skin injuries that is given by 2e International Skin Imaging Collaboration (ISIC). It incorporates 2000 pictures; 374 of them are threatening skin tumours alluded to as "Melanoma" and 1626 of them are favourable skin tumours alluded to as "Nevus of Seborrheic Keratosis". 2us, it is a twofold picture arrangement task that recognizes (a) Melanoma and (b) Nevus and Seborrheic Keratosis. Each picture right now an alternate goal, which we should address.

**Table 2.** ISIC2017 dataset

| Image Category      | Number of Images | Label |
|---------------------|------------------|-------|
| Melanoma            | 374              | 1     |
| Nevus of seborrheic | 1626             | 2     |

Table 2 shows the piece of the ISIC2017 dataset. So as to assess our trials, we utilized the accompanying arrangement. In the first place, each dataset was partitioned into a preparation set, an approval set, and a test set.

Right now, will give a progression of trials on the exactness and calculation running time on two genuine clinical picture datasets. i.e. exactness here is characterized as the level of effectively ordered clinical pictures. To all the more likely look at the calculations, we utilize the disarray network and collector working trademark (ROC) bend to additionally assess the model. 2e disarray grid is a table format that can depict the quantity of genuine positive, bogus negative, genuine negative, and bogus negative in an assessment of a multiclass picture characterization calculation. ROC bend is a realistic plot that is acquired through processing the genuine positive rate (TPR) against the bogus positive rate (FPR) by setting various limits, where the meaning of TPR and FPR are as per the following:

Figure 4 represents the Precession of two data sets and different mechanisms with respect to augmentation time. And here it represents the different models comparison with two different data sets. At each epoch proposed model represents better precession than SVM, NB and only CNN. In our proposed mechanism combination of CNN and auto encoder is used. Basically, encoder takes the major features from the medical images and CNN captured the in deep features and make the patterns in a novel way. So, it can made better precession results than all other classification mechanisms.

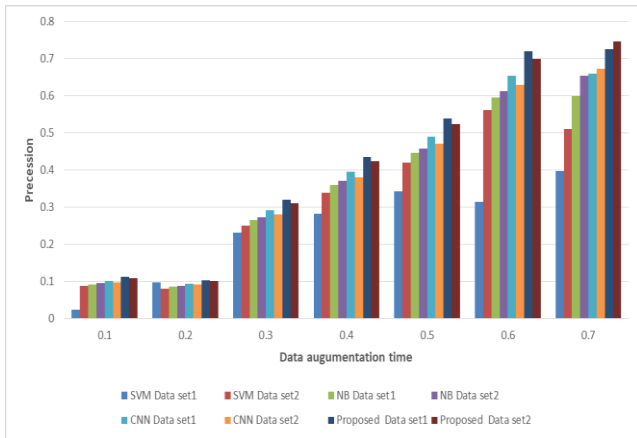


Figure 4. Precision

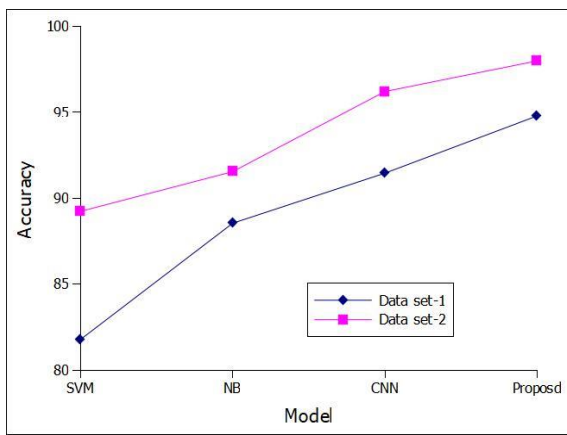


Figure 5. Accuracy

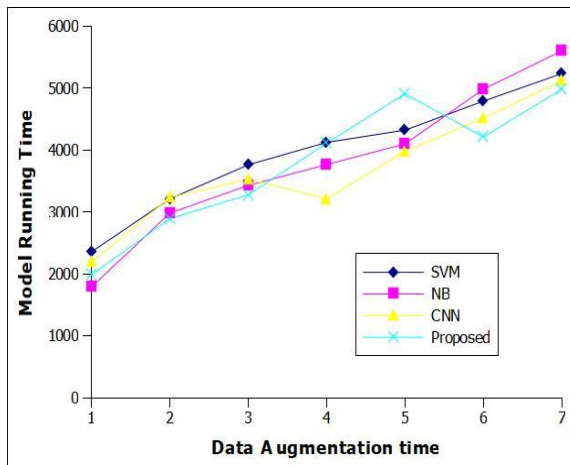


Figure 6. Model run time

Figure 5 represents the Accuracy of different models with two different medical data sets and. Proposed model gives more accurate than SVM, NB and only CNN. In our proposed mechanism combination of CNN and auto encoder is used. Basically, encoder takes the major features from the medical images and CNN captured the in deep features and make the patterns in a novel way. So, it can made accurate results than all other classification mechanisms.

Figure 6 represents model run time, generally run time means time taken for a model to execute on a particular data set. But here we made data set seven different into epochs. Each ephod contains set of images. Here we made comparison

with different existing models of SVM, NB and CNN. Proposed model is slightly cope with the CNN. but it produces better outcome than SVM and NB.

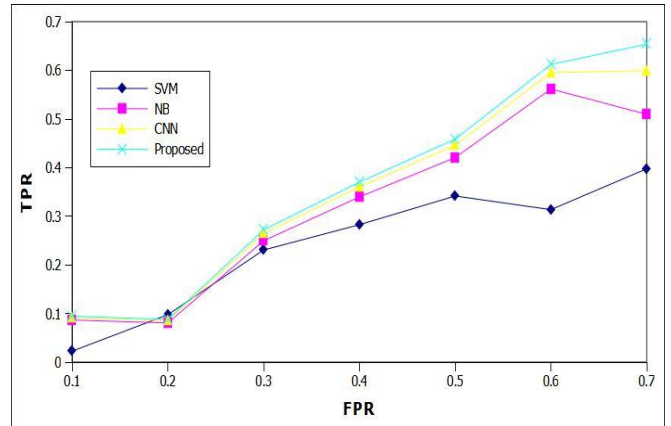


Figure 7. Area under curve

The area under curve AUC bend is a realistic plot that is acquired through figuring the TPR and FPR by setting various edges. It is valuable to assess the exhibition of the double class picture order calculation since the parallel dataset contains an example imbalanced issue. Figure 7 looks at the ROC bend of various calculations on the ISIC2017 dataset. The closer the bend is to the upper left corner of the tomahawks, the better the exhibition. It is anything but difficult to see that our model gets the best execution and that R include combination has the second best presentation. The classification performance levels are depicted in Figure 8. The proposed model has better classification accuracy levels.

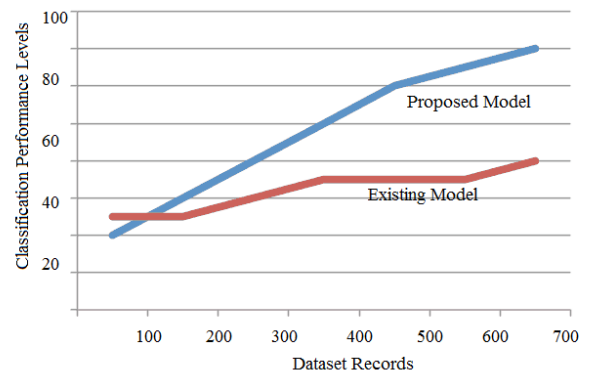


Figure 8. Classification accuracy levels

The number of nodes and number of layers are depicted in Figure 9.

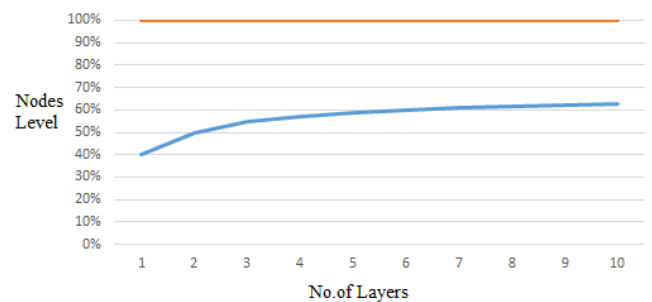


Figure 9. No. of nodes and layers levels

## 5. CONCLUSION

During the ongoing scarcely any years, deep learning has increased a focal situation toward the mechanization of our everyday life and conveyed impressive upgrades when contrasted with conventional AI calculations. In our proposed mechanism combination of CNN and auto encoder is used. Basically encoder takes the major features from the medical images and CNN captured the in deep features and make the patterns in a novel way. So it can made accurate results than all other classification mechanisms. And we evaluate the proposed approach on two benchmark medical image datasets: HIS2828 and ISIC2017. The proposed mechanism performs better with respect to existing SVM, NB and CNN.

## REFERENCES

- [1] Cruz-Roa, A., Caicedo, J.C., González, F.A. (2011). Visual pattern mining in histology image collections using bag of features. *Artificial Intelligence in Medicine*, 52(2): 91-106. <https://doi.org/10.1016/j.artmed.2011.04.010>
- [2] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639): 115-118. <https://doi.org/10.1038/nature21056>
- [3] Zare, M.R., Mueen, A., Awedh, M., Seng, W.C. (2013). Automatic classification of medical X-ray images: Hybrid generative-discriminative approach. *IET Image Processing*, 7(5): 523-532. <https://doi.org/10.1049/iet-ipr.2013.0049>
- [4] Hinton, G.E., Osindero, S., Teh, Y.W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7): 1527-1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
- [5] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6): 1097-1105. <https://doi.org/10.1145/3065386>
- [6] Simonyan, K., Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. <http://arxiv.org/abs/1409.1556>.
- [7] Barata, C., Ruela, M., Francisco, M., Mendonça, T., Marques, J.S. (2014). Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Systems Journal*, 8(3): 965-979. <https://doi.org/10.1109/JSYST.2013.2271540>
- [8] Iyatomi, H., Oka, H., Celebi, M.E., Hashimoto, M., Hagiwara, M., Tanaka, M., Ogawa, K. (2008). An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. *Computerized Medical Imaging and Graphics*, 32(7): 566-579. <https://doi.org/10.1016/j.compmedimag.2008.06.005>
- [9] Stoecker, W.V., Wronkiewicz, M., Chowdhury, R., Stanley, R.J., Bangert, A., Shrestha, B., Calcara, D.A., Rabinovitz, H.S., Oliviero, M., Ahmed, F., Perry, L.A., Drugge, R. (2011). Detection of granularity in dermoscopy images of malignant melanoma using color and texture features. *Computerized Medical Imaging and Graphics*, 35(2): 144-147. <https://doi.org/10.1016/j.compmedimag.2010.09.005>
- [10] Riaz, F., Hassan, A., Nisar, R., Dinis-Ribeiro, M., Coimbra, M. (2017). Content-adaptive region-based color texture descriptors for medical images. *IEEE Journal of Biomedical and Health Informatics*, 21(1): 162-171. <https://doi.org/10.1109/JBHI.2015.2492464>
- [11] Ramirez, J., Gorriz, J.M., Segovia, F., Chaves, R., Salas-Gonzalez, D.S., López, M., Alvarez, I., Padilla, P. (2010). Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification. *Neuroscience Letters*, 472(2): 99-103. <https://doi.org/10.1016/j.neulet.2010.01.056>
- [12] Zhang, Y., Chen, S., Wang, S., Yang, J., Phillips, P. (2015). Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine. *International Journal of Imaging Systems and Technology*, 25(4): 317-327. <https://doi.org/10.1002/ima.22144>
- [13] Zhang, Y., Dong, Z., Liu, A., Wang, S., Ji, G., Zhang, Z., Yang, J. (2015). Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine. *Journal of Medical Imaging and Health Informatics*, 5(7): 1395-1403. <https://doi.org/10.1166/jmih.2015.1542>
- [14] Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D.D., Chen, M. (2014). Medical image classification with convolutional neural network. *13th International Conference on Control Automation Robotics & Vision (ICARCV)*, Singapore, pp. 844-848. <https://doi.org/10.1109/ICARCV.2014.7064414>
- [15] Bar, Y., Diamant, I., Wolf, L., Greenspan, H. (2015). Deep learning with non-medical training used for chest pathology identification. *Proceedings of SPIE Medical Imaging, International Society for Optics and Photonics, Orlando, FL, USA, February 2015*, p. 94140V. <https://doi.org/10.1117/12.2083124>
- [16] Shin, H.C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5): 1285-1298. <https://doi.org/10.1109/TMI.2016.2528162>
- [17] Ahn, E., Kumar, A., Kim, J., Li, C., Feng, D., Fulham, M. (2016). X-ray image classification using domain transferred convolutional neural networks and local sparse spatial pyramid. *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, Prague, pp. 855-858. <https://doi.org/10.1109/ISBI.2016.7493400>
- [18] Ouyang, W., Luo, P., Zeng, X., Qiu, S., Tan, Y., Li, H., Yang, S., Wang, Z., Xiong, Y., Qian, C., Zhu, Z., Wang, R., Loy, C., Wang, X., Tang, X. (2014). DeepID-net: multi-stage and deformable deep convolutional neural networks for object detection. <http://arxiv.org/abs/1409.3505>.
- [19] Patil, S., Kuchanur, M. (2012). Lung cancer classification using image processing. *International Journal of Engineering and Innovative Technology(IJEIT)*, 2(3): 37-42
- [20] Fernando, B., Fromont, E., Muselet, D., Sebban, M. (2012). Discriminative feature fusion for image classification. *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI*, pp. 3434-3441.

- <https://doi.org/10.1109/CVPR.2012.6248084>
- [21] Chan, T., Jia, K., Gao, S., Lu, J., Zeng, Z., Ma, Y. (2015). PCANet: A simple deep learning baseline for image classification? *IEEE Transactions on Image Processing*, 24(12): 5017-5032. <https://doi.org/10.1109/TIP.2015.2475625>
- [22] Zeng, R., Wu, J., Shao, Z., Chen, Y., Chen, B., Senhadji, L., Shu, H. (2016). Color image classification via quaternion principal component analysis network. *Neurocomputing*, 216: 416-428. <https://doi.org/10.1016/j.neucom.2016.08.006>
- [23] Rakotomamonjy, A., Petitjean, C., Salaun, M., Thiberville, L. (2014). Scattering features for lung cancer detection in fibered confocal fluorescence microscopy images. *Artificial Intelligence in Medicine*, 61(2): 105-118. <https://doi.org/10.1016/j.artmed.2014.05.003>