



# Deep Multiple Instance Learning Approach for Classification in Clinical Decision Support Systems

Vani V. <sup>1</sup>, Piyush Kumar Pareek<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering Nitte Meenakshi Institute of Technology Bengaluru India

<sup>2</sup>Department of Artificial Intelligence and Machine Learning Nitte Meenakshi Institute of Technology Bengaluru India

Emails: [vani.v@nmit.ac.in](mailto:vani.v@nmit.ac.in); [piyush.kumar@nmit.ac.in](mailto:piyush.kumar@nmit.ac.in)

## Abstract

To get around the drawbacks of conventional classification algorithms that require manual feature extraction and the high computational cost of neural networks, this paper introduces a deep convolutional neural network with multiple instances learning approaches, namely dynamic max pooling and sparse representation. For the categorization of tuberculosis lung illness, this model combines deep convolutional neural networks and multiple instance learning. The design was composed of four phases: pre-processing, instance production, feature extraction, and classification. To perform feature extraction, a model based on a customized version of the VGG16 architecture was trained from scratch. Multiple instances learning techniques such as Diverse Density (DD) and the Maximum pattern bag formulation of the Support Vector Machine were used to evaluate how well the proposed classification algorithm performed in comparison (SVM). The numerical findings demonstrated that the new method offered a higher level of accuracy than the methods that had been used in the past. When evaluating the efficacy of the current method, accuracy, specificity, sensitivity, and error rate were all taken into consideration. The accuracy of the max-pooling based framework and the sparse representation framework was found to be greater than that of the other multiple instance strategies, coming in at 91.51% and 89.84%, respectively, when compared to that of the other methods. The improved accuracy of the present system that makes use of deep neural networks is mostly attributable to the contributions made by features such as transfer learning and automatic feature extraction.

**Keywords:** CAD; CNN; MIL; DCNN.

## 1. Introduction

While utilizing machine learning, one must nonetheless keep in mind the numerous technological, medical, and ethical considerations that must be taken into account [1]. The requirement for broad authentication of machine learning-based knowledge in real-world circumstances, objective valuation of aids and dangers, deterrence of scientific over-dependence, and the subsequent harm of scientific, moral, and communally pertinent administrative aptitudes are some of the influences elaborate here. Additional challenges include the obligation for comprehensive benchmarking and independent authentications, the evolution of end-user proficiency from computational specialists to field users, and the accountable allocation of code and data, which marks it conceivable to calculate pipelines in an open and translucent manner [2].

Over the past few decades, computational support has increasingly been included in clinical decision-making. Computer-aided technologies have been created to help with the medical diagnosis of illnesses like acute abdominal pain or the modeling of mortality in intensive-care units [3-4]. Using data-driven machine learning (ML)-based support systems and FDA-approved medical devices, the use of such knowledge-based systems in

hospitals and primary care facilities has increased over the past few years (FDA). ML algorithms are progressively being used in several studies in a wide range of fields [5].

The goal of ML (machine learning), an area of AI (artificial intelligence), is to build computer algorithms that automatically improve the accuracy of their output over time. To create predictions that will suit the sample data, machine-learning algorithms utilize sample data, sometimes referred to as a training dataset. Predictive Solutions, computer-aided detection systems, computer-aided diagnosis systems, recommendation systems, and early warning systems are some of the scenarios in which these predictions can inform healthcare decisions (EWS). PCSs are machine learning (ML) systems that help provide the foundation for downstream analytical techniques like marker-gene identification or drug development [6]. For instance, PCS systems could be used to validate questionnaires before evaluating them for clinical diagnosis. X-ray radiography, computed tomography, magnetic resonance imaging, and ultrasonography are examples of medical imaging that are analyzed using software-based analyses such as CADe and CADx. By spotting questionable details and drawing the clinician's attention to them, CADe helps doctors interpret medical imaging, whereas CADx offers a more structured diagnostic of a picture. For instance, CADe would identify worrisome polyps during a colonoscopy, but CADx would identify the pathogenic entity of such polyp [7]. RecSys, as its name suggests, uses one of the numerous user interfaces to offer dietary, medical, or pharmaceutical suggestions for a patient that are often personalized and tied to their health. EWS generates warning data that enable people, communities, and organizations to plan ahead and take proactive measures in the event of a particular risk, thereby offering extra knowledge and sparing valuable anticipation time [8].

Based on the data contained in the medical images produced by the aforementioned imaging modalities, doctors examine the images to assess the problems and suggest appropriate treatment measures. The development of CAD systems has significantly reduced the amount of human effort required for the diagnosis of problems due to recent advancements in the disciplines of artificial intelligence, machine learning, pattern recognition, medical imaging, radiography, computer hardware, and computer vision [9]. The CAD system examines medical photos as input, searches for any worrisome patterns, and alerts the doctors to the existence of ROI. Radiologists may notice a malignancy, a concussion, or a nerve blockage as the suspicious pattern or ROI. Additionally evaluates such ROI and offers the doctors a second assessment based on precedent-same cases [10].

Multiple instance learning (MIL) is a subset of supervised learning in machine learning. The learner receives a set of labeled bags, each holding multiple instances, as opposed to a set of instances that are individually labeled. A bag may be labeled negative in the straightforward situation of multiple-instance binary classification if every instance inside is negative [11]. On the other hand, a bag is classified as positive if it contains at least one positive occurrence. The learner attempts to either (i) infer a concept from a set of labeled bags that will enable accurate labeling of individual instances, or (ii) learn how to label bags without inferring the concept [12].

What follows is the outline for the rest of the paper. The related work is briefly described in part 2, and the methodology and the theoretical foundations of the methods used are described in section 3. The simulation results and analysis are presented in section 4. For the chapter's final section, "key findings" we summarize the most important results.

## **2. Related Work Done:**

In a study conducted by a small group of researchers, CNN and transfer learning were used to classify lung nodule malignancy. The three stages of this methodology were data processing, an implementation strategy, and evaluation. Since the majority of lung nodule classification methods do not offer reliable evaluation, deep characteristics were employed. The CT images were first processed, and after that, numerous CNNs were constructed [13]. Moreover, 10-fold cross-validations with the classifiers naive Bayes, multilayer perceptron, Support Vector Machine (SVM), K-Nearest Neighbours (KNN), and Random Forest were performed on each set of deep features. The ImageNet database was later used to train the model [14-15].

The network consists of three thick layers, an average pooling with a size equal to the size of the final feature maps, five convolutional layers with 2x2 kernels and leaky ReLU (Rectified Linear Unit) activations, and five convolutional layers [16]. In order to include more non-linear activations in this model, the kernel size was further decreased, but the overall receptive field was kept small enough (66) to only capture the pertinent local texture structure, which was absent from the previous models. Each layer contained a number of kernels corresponding to the receptive field of its neurons, allowing it to manage the aforementioned structures' increasing complexity. The dataset that was gathered was used to train the data. A comparison of a categorization model to earlier approaches demonstrated its efficacy [17].

Another team of researchers developed the MRCNN (Multi-scale Rotation invariant Convolutional Neural Network) model, which uses the Gabor-local binary pattern to categorize different types of lung tissue on High-Resolution Computed Tomography (HRCT) images (Gabor-LBP). In contrast to other efforts, this model addressed the issues brought on by an uneven distribution of samples among various classes by altering the overlapping size between the neighbouring patches [18]. An end-to-end system that used a hierarchical architecture for converting the low-level representations of input images into high-level structural data was used to research the Gabor-LBP. Then, on the basis of multi-scale images, complementary features were found that were anticipated to provide a rich and useful representation of lung tissue patterns [19]. The publicly accessible database of ILD cases, which includes HRCT pictures with a slice thickness of 1mm, served as the basis for the training on the data. In this technique, radiologists manually marked 2062 2-D ROI in 113 sets. This model's accuracy of 90.1% was determined to be higher than Gabor's accuracy of 89.7% and LBP's accuracy of 86.7% [20].

A model for categorizing lung nodules using three-dimensional deep convolutional neural networks was reported in a few other publications. A three-dimensional CNN with dense connections and short connections was employed in this model. These links made it possible to identify common and distinguishing characteristics of lung nodules. The preceding work's shallow three-dimensional CNN was unable to adequately capture the characteristics of spherical-shaped nodules. In order to capture 3D features, the three-dimensional CNN's dimensions were therefore expanded from two to three [21]. In addition, an alternative strategy was utilized to increase the performance of the classification model. Both models employed three-dimensional CNNs that had been trained to deal with the issue of vanishing gradient. Overfitting a model might result from training on an unbalanced dataset. Thus, a number of sampling and augmentation techniques were used to address the data skewness issue. It was claimed that this strategy outperformed current deep learning methods in terms of Competition Performance Measure (CPM) results [22].

The strategy used three techniques: pixel classification, texture feature extraction, and lung segmentation. Their research, which was based on other ML, Active Learning (AL), and 1-class classification algorithms, utilized MIL technology. The older models were thought to have suffered greatly from the uncertainty of the intrinsic properties [23]. As a result, the MIL technology was employed, and it was asserted that it was a unique learning prototype that resolved the issue of handling the uncertainty of instance labels. The categorization of images and texts was the focus of the AL. The lung likelihood map that was produced underwent post-processing to produce binary segmentation. In order to re-label the afflicted lesion, the MIL was combined with the AL model and employed as a re-ranking model. Using datasets with a 50/50 split between normal and atypical instances, this model was tested. This system's prediction rate (TB Diagnosis utilizing miSVM and PEDD) was discovered to be favorable with improved performance [24].

### **3. The Proposed Method:**

To get around the problems with manual feature extraction and the high computing cost of neural networks, this research paper introduces a deep convolutional neural network with several instance learning methodologies, including dynamic max pooling and sparse representation. Annotating each image's temporal context was unnecessary thanks to multiple instance learning.

In order to classify cases of tuberculous lung illness, the model proposed here combines deep convolutional neural networks with multiple instance learning, as illustrated in Figure 1. Pre-processing, Instance Generation, Feature Extraction, and Classification were the four main pillars of the design.

Noise cancellation was one of the first steps in the pre-processing phase. The photos were pre-processed, and then sent into an instance generating algorithm. A convolutional Neural Network was utilized to extract features from the examples. Pre-processing, instance generation, feature extraction, and multiple instance learning-based categorization were all part of the sparse representation of MIL. For feature extraction, a cost-effective and efficient neural network was used. As the total learnable parameter of VGG 16 was discovered to be greater, it was used for feature extraction. To make the CNN suitable for feature extraction, the final completely linked layer was deleted. The VGG 16 deep CNN's classification results were fed into two different multiple instance frameworks.

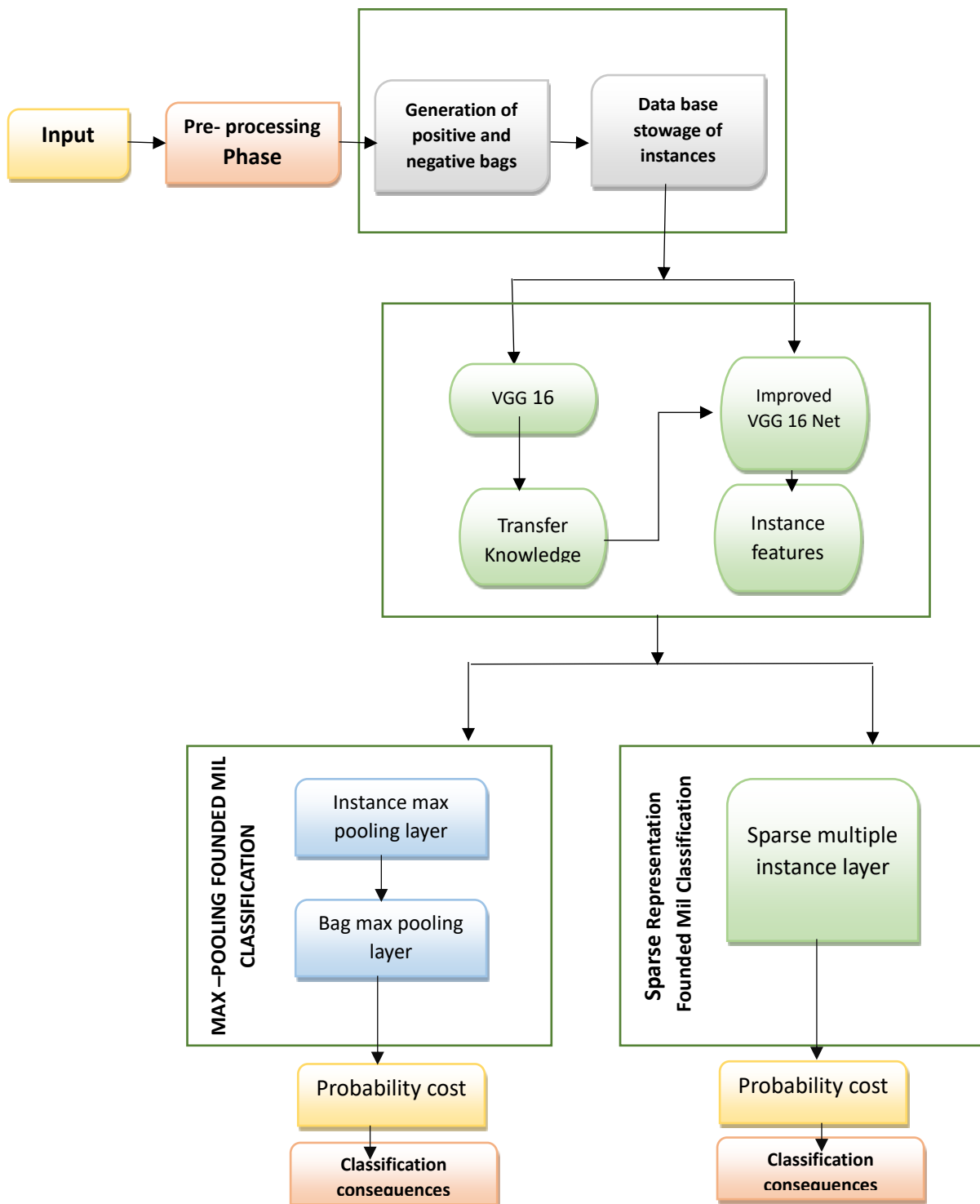


Figure 1: The Proposed Algorithm block Diagram.

4.1. Pre-processing measures were crucial in reducing the impact of the background on the inspection of abnormalities before the image was processed. The initial processing of the images involved localizing the lung diseases and removing background noise. A phase of pre-processing was run to restore image quality and make the input image processable. The process's primary aim was to enhance the image's overall quality. The preliminary processing of the lung CT images involved the employment of a Weiner filter. Noise suppression in images is the primary application of this non-linear filter.

#### 4.2. Instance Generation Phase:

Instead of focusing on single instances, Multiple Instance Learning (MIL) considered sets of bags that had been assigned positive and negative labels. The cases were label-blind in multiple-instance learning. Images were represented as collections of instances for the purpose of multiple instances learning for image classification.

#### 4.3. Feature Extraction Phase:

Traditional feature extraction algorithms make manual feature extraction a laborious and time-consuming task. In CNN, the processing was divided up into a number of "layers," each of which received its input from the layer before it. The feature map was the result of a series of computations carried out by each layer. The convolutional layer was responsible for performing the convolution at a specified kernel location. When it came to gathering relevant data, we turned to the VGG 16 convolutional neural network. To lower the error rate, VGG 16's transfer learning made adjustments to the weights via backpropagation. Part of the VGG 16 was used for feature extraction after the final completely linked layer was removed.

#### Algorithm: MaxPool based Multiple Instance Learning

Feature inputs from instances

1. The instance's properties were set up.
2. settings for the maximum pooling layer
3. The cross entropy was determined.
4. Adjusting the dynamic pooling layer's weight till the loss function was minimized.
5. Figuring out the likelihood of occurrences
6. A bag possibility was intended by totaling up all the separate occurrence probabilities.

The end result is an Image Class.

Multiple instances learning with a sparse representation was accomplished with sparse positive bags. In the lung image cases, Sparse MIL allocated varying amounts of weight to label-based assignments and label distributions. Because the masses are sparse, the cross-entropy function of the neural network was modified to include the sparsity factor. The genuine label format was unique to sparse categorical issues.

#### Algorithm: Sparse Representation based multiple instance learning

Characteristics of Instances as an Input

1. The characteristics of instances have been set.
2. The cross entropy was determined
3. Loss reduction was attained by regulating the layer weightiness of the vibrant pooling layer.
4. The likelihood of each occurrence was determined.
5. Probabilities at the instance level were pooled together to create the bag probability.

Ultimately, a classified image is the output.

#### 4. Result and Discussion:

The current technique was mathematically analysed by associating the consequences with those of prevalent learning approaches. Several instance methods were evaluated on a variety of criteria to determine which was best. Accuracy, specificity, sensitivity, and error rate were utilized as metrics to gauge the effectiveness of the suggested methods.

##### 4.1. Accuracy:

It is a typical metric for classifying test results numerically. Increased precision indicates a more efficient system. In Figure II, we see a comparison of the current feature selection model's accuracy to that of certain popular feature selection algorithms. The proposed method improved classification accuracy despite having fewer features than MOPSO alone by virtue of the incorporation of local Tabu search.

$$\text{Accuracy} = \frac{TN+TP}{\text{Total data Sample}} \times 100 \quad (1)$$

##### 4.2. Specificity:

Specificity was defined as the absence of incorrect data classification. True Negative Rate is another name for it (TNR). The figure displays the specificity of the current method in comparison to commonly utilized methods.

$$\text{Specificity} = \frac{TN}{TN+FP} \times 100 \quad (2)$$

##### 4.3. Precision:

The present work is believed to have the precision to provide useful outcomes. The value of precision indicates what fraction of valid affirmative identifications were made. The figure displays the results of a comparison between the current model and the commonly used PSO and BCO techniques in terms of accuracy. The present system's accuracy was determined by

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

##### 4.4. Sensitivity:

The accuracy with which the model places the test data into one of its classes constitutes the present method's sensitivity. How many true positives were successfully detected was the question it addressed. True Positive Rate is another name for it (TPR). Figure II displays the results of a comparison between the present strategy and the commonly adopted PSO and BCO approaches with regard to Recall.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100 \quad (4)$$

4.5. Error rate: It represents how many instances the classifier algorithm got wrong. The comparison with the existing approach is shown in Figure III.

$$\text{Error rate} = 1 - \text{Accuracy} \quad (5)$$

The assessment parameters are compared among several MIL methods in Table 1. The Diverse Density algorithm was discovered to have a significant loss and a large number of parameters. The current method required fewer iterations with minimal loss than miSVM and DD algorithms.

Table 1: Assessment of enactment processes

S. No	MIL Algorithms	Accuracy	Specificity	Sensitivity	Error Rate
1	Diverse Density	89.12	84.79	86.5	12.4
2	Max-pooling based MIL	90.37	88.46	89.8	1.04

3	Sparse representation of MIL	89.98	87.59	89.21	11.2
---	------------------------------	-------	-------	-------	------

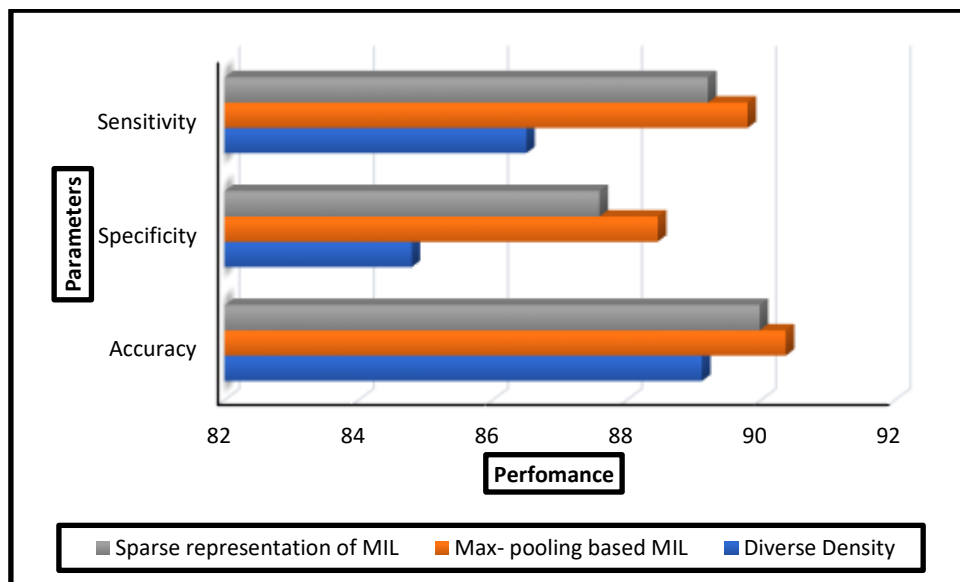


Figure 2: The performance comparison of different MIL algorithm.

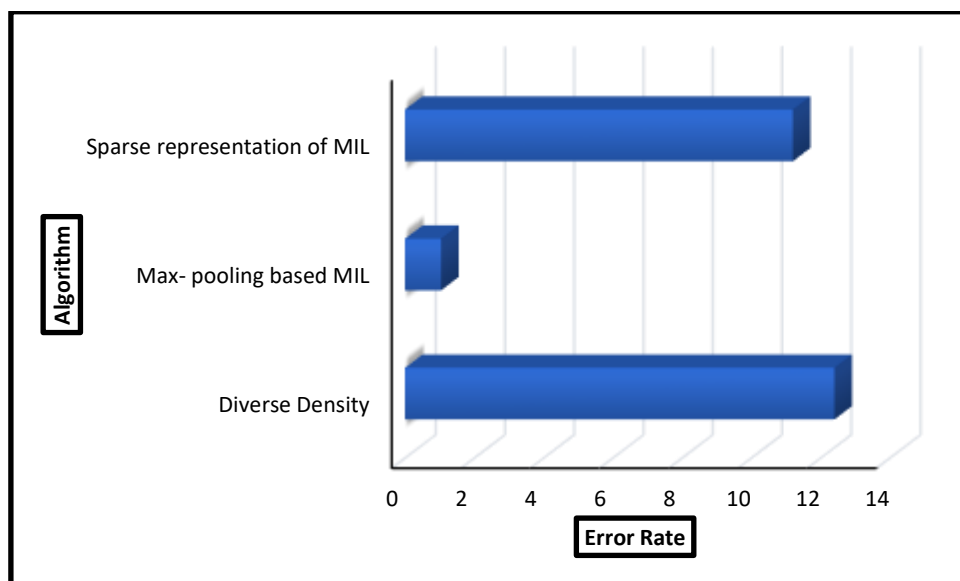


Figure 3: The error rate comparison of different MIL algorithms.

The values of the performance analysis are greater for the present technique, as can be shown in Table 1. This is in contrast to the results obtained by the other algorithms. Also, when compared to Diverse Density (DD) algorithms, the error rate that was achieved by the current method was lower.

Form figure I and II we can conclude that the numerical findings demonstrated that the new method offered a higher level of accuracy than the methods that had been used in the past. When evaluating the efficacy of the current method, accuracy, specificity, sensitivity, and error rate were all taken into consideration. The accuracy of the max-pooling based framework and the sparse representation framework was found to be greater than that of the other multiple instance strategies, coming in at 91.51% and 89.84%, respectively, when compared to that of

the other methods. The improved accuracy of the present system that makes use of deep neural networks is mostly attributable to the contributions made by features such as transfer learning and automatic feature extraction.

## 5. Conclusion and Future Scope:

The traditional CAD system was upgraded with the use of a multiple instance learning method. The Wiener filter was used to remove noise from the input photos before they were processed further. Bags of instances were created from the input dataset before feature extraction was performed. To perform feature extraction, a model based on a customized version of the VGG16 architecture was trained from scratch. Multiple instances learning techniques such as Diverse Density (DD) and the Maximum pattern bag formulation of the Support Vector Machine were used to evaluate how well the proposed classification algorithm performed in comparison (SVM). The numerical results showed that the accuracy of the new method was higher than that of the previous methods. Accuracy, specificity, sensitivity, and error rate were used to assess the current method's effectiveness. It was discovered that the accuracy of the max-pooling based framework and the sparse representation framework was higher than that of the other multiple instance techniques, coming in at 91.51% and 89.84%, respectively. Features like transfer learning and automated feature extraction contributed heavily to the present system's improved accuracy by employing deep neural networks.

According to the findings of this work, there is a need to make modifications to the sample line approach in order to extract lung coordinates from a variety of various orientations. These modifications are suggested in order to boost segmentation accuracy. It also provides some suggestions that the accuracy of CAD systems can be improved further by combining the results of clinical tests with the most advanced machine learning models.

## References

- [1] R.T. Sutton, D. Pincock, D.C. Baumgart, D.C. Sadowski, R.N. Fedorak, K.I. Kroeker An overview of clinical decision support systems: benefits, risks, and strategies for success *NPJ Digit. Med.*, 3 (2020), p. 17
- [2] L. Moja, K.H. Kwag, T. Lytras, L. Bertizzolo, L. Brandt, V. Pecoraro, G. Rigon, A. Vaona, F. Ruggiero, M. Mangia, et al. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis *Am. J. Public Health*, 104 (2014), pp. e12-e22
- [3] J. Varghese, M. Kleine, S.I. Gessner, S. Sandmann, M. Dugas Effects of computerized decision support system implementations on patient outcomes in inpatient care: a systematic review *J. Am. Med. Inform. Assoc.*, 25 (2018), pp. 593-602
- [4] S. Benjamens, P. Dhunoo, B. Meskó The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database *NPJ Digit. Med.*, 3 (2020), p. 118
- [5] J. Dong, Y. Geng, D. Lu, B. Li, L. Tian, D. Lin, Y. Zhang Clinical trials for artificial intelligence in cancer diagnosis: a cross-sectional study of registered trials in *Clinical Trials. Gov Front. Oncol.*, 10 (2020), p. 1629
- [6] A. Rajkomar, J. Dean, I. Kohane Machine learning in medicine *N. Engl. J. Med.*, 380 (2019), pp. 1347-1358
- [7] E.J. Topol High-performance medicine: the convergence of human and artificial intelligence *Nat. Med.*, 25 (2019), pp. 44-56
- [8] Masud, M.; Hossain, M.S.; Alhumyani, H.; Alshamrani, S.S.; Cheikhrouhou, O.; Ibrahim, S.; Muhammad, G.; Rashed, A.E.E.; Gupta, B.B. Pre-Trained Convolutional Neural Networks for Breast Cancer Detection Using Ultrasound Images. *ACM Trans. Internet Technol.* 2021, 21, 85.
- [9] Thigpen, D.; Kappler, A.; Brem, R. The Role of Ultrasound in Screening Dense Breasts—A Review of the Literature and Practical Solutions for Implementation. *Diagnostics* 2018, 8, 20.
- [10] Muhammad, M.; Zeebaree, D.; Brifceni, A.M.A.; Saeed, J.; Zebari, D.A. Region of interest segmentation based on clustering techniques for breast cancer ultrasound images: A review. *J. Appl. Sci. Technol. Trends* 2020, 1, 78–91.
- [11] Wang, N.; Bian, C.; Wang, Y.; Xu, M.; Qin, C.; Yang, X.; Wang, T.; Li, A.; Shen, D.; Ni, D. Densely Deep Supervised Networks with Threshold Loss for Cancer Detection in Automated Breast Ultrasound.

- In Proceedings of the International Conference on Medical Image Computing and Computer—Assisted Intervention, Granada, Spain, 16–20 September 2018; pp. 641–648.
- [12] Guo, R.; Lu, G.; Qin, B.; Fei, B. Ultrasound Imaging Technologies for Breast Cancer Detection and Management: A Review. *Ultrasound Med. Biol.* 2018, 44, 37–70.
- [13] Zhang, X.; Lin, X.; Tan, Y.; Zhu, Y.; Wang, H.; Feng, R.; Tang, G.; Zhou, X.; Li, A.; Qiao, Y. A multicenter hospital-based diagnosis study of automated breast ultrasound system in detecting breast cancer among Chinese women. *Chin. J. Cancer Res.* 2018, 30, 231.
- [14] Mohammed, M.A.; Al-Khateeb, B.; Rashid, A.N.; Ibrahim, D.A.; Abd Ghani, M.K.; Mostafa, S.A. Neural net-work and multi-fractal dimension features for breast cancer classification from ultrasound images. *Comput. Electr. Eng.* 2018, 70, 871–882.
- [15] Wang, Y.; Wang, N.; Xu, M.; Yu, J.; Qin, C.; Luo, X.; Yang, X.; Wang, T.; Li, A.; Ni, D. Deeply-supervised net-works with threshold loss for cancer detection in automated breast ultrasound. *IEEE Trans. Med. Imaging* 2019, 39, 866–876.
- [16] Shen, L.; Margolies, L.R.; Rothstein, J.H.; Fluder, E.; McBride, R.; Sieh, W. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Sci. Rep.* 2019, 9, 12495.
- [17] K. Kiruthika, S. Gayathri, R. Hemalatha, P. Menaga, Design and Development of Mobile Healthcare Application for “Ayurvedic” based Clinical Documents, *Journal of Cognitive Human-Computer Interaction*, Vol. 1, No. 1, (2021) : 18-27 (Doi : <https://doi.org/10.54216/JCHCI.010103>)
- [18] Badawy, S.M.; Mohamed, A.E.N.A.; Hefnawy, A.A.; Zidan, H.E.; GadAllah, M.T.; El-Banby, G.M. Automatic semantic segmentation of breast tumors in ultrasound images based on combining fuzzy logic and deep learning—A feasibility study. *PLoS ONE* 2021, 16, e0251899.
- [19] Almajalid, R.; Shan, J.; Du, Y.; Zhang, M. Development of a deep-learning-based method for breast ultrasound image segmentation. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 1103–1108.
- [20] Kalafi, E.Y.; Jodeiri, A.; Setarehdan, S.K.; Lin, N.W.; Rahmat, K.; Taib, N.A.; Ganggayah, M.D.; Dhillon, S.K. Classification of Breast Cancer Lesions in Ultrasound Images by Using Attention Layer and Loss Ensemble in Deep Convolutional Neural Networks. *Diagnostics* 2021, 11, 1859.
- [21] Cao, Z.; Duan, L.; Yang, G.; Yue, T.; Chen, Q. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Med. Imaging* 2019, 19, 51
- [22] Roy V., Shukla S. (2013) Image Denoising by Data Adaptive and Non-Data Adaptive Transform Domain Denoising Method Using EEG Signal. In: Kumar V., Bhatele M. (eds) Proceedings of All India Seminar on Biomedical Engineering 2012 (AISOB 2012). Lecture Notes in Bioengineering. Springer, India. [https://doi.org/10.1007/978-81-322-0970-6\\_2](https://doi.org/10.1007/978-81-322-0970-6_2).
- [23] Qi, X.; Yi, F.; Zhang, L.; Chen, Y.; Pi, Y.; Chen, Y.; Guo, J.; Wang, J.; Guo, Q.; Li, J.; et al. Computer-aided Diagnosis of Breast Cancer in Ultrasonography Images by Deep learning. *Neuro-computing* 2021, 472, 152–165.
- [24] Shankar, K.; Perumal, E.; Tiwari, P.; Shorfuzzaman, M.; Gupta, D. Deep learning and evolutionary intelligence with fusion-based feature extraction for detection of COVID-19 from chest X-ray images. *Multimedia Syst.* 2021, 2021, 1–13.