

# Spatial Feature Fusion for Biomedical Image Classification based on Ensemble Deep CNN and Transfer Learning

Sanskruti Patel, Rachana Patel, Nilay Ganatra, Atul Patel  
Smt. Chandaben Mohanbhai Patel Institute of Computer Applications  
Charotar University of Science and Technology, Changa, India

**Abstract**—Biomedical imaging is a rapidly evolving field that covers different types of imaging techniques which are used for diagnostic and therapeutic purposes. It plays a vital role in diagnosis and treating health conditions of human body. Classification of different imaging modalities plays a vital role in terms of providing better care and treatment options to the patients. Advancements in technology open up the new doors for medical professionals and this involves deep learning methods for automatic image classification. Convolutional neural network (CNN) is a special class of deep learning that is applied to visual imagery. In this paper, a novel spatial feature fusion based deep CNN is proposed for classification of microscopic peripheral blood cell images. In this work, multiple transfer learning features are extracted through four pre-trained CNN architectures namely VGG19, ResNet50, MobileNetV2 and DenseNet169. These features are fused into a generalized feature space that increases the classification accuracy. The dataset considered for the experiment contains 17902 microscopic images that are categorized into 8 distinct classes. The result shows that the proposed CNN model with fusion of multiple transfer learning features outperforms the individual pre-trained CNN model. The proposed model achieved 96.10% accuracy, 96.55% F1-score, 96.40% Precision and 96.70% Recall values.

**Keywords**—Biomedical images; convolutional neural network; ensemble deep learning; feature fusion

## I. INTRODUCTION

Biomedical imaging refers to capturing of an organ or tissue for diagnostic purpose. The field is very broad and rapidly evolving that covers different types of imaging modalities like ultrasound, magnetic resonance imaging (MRI), computerized tomography (CT), positron emission tomography (PET), etc. [1]. Biomedical and medical imaging plays a significant role in diagnosis and treating health conditions of human body. It helps to identify problematic health conditions in their early stages that certainly lead for providing better treatment to the patients [2]. The structural and functional changes in biological tissues of the human body normally cause the possible health problems. Biomedical imaging provides a way to view inside the human body that helps to reveal such changes [3].

Classification of different imaging modalities plays a vital role in terms of providing better care and treatment options to the patients. The traditional way to classify these different modalities of images is the naked eye classification that is

performed by medical professional or subject expert. This is the cumbersome and sometimes time-consuming method. Advancements in technology open up the new doors for medical professionals that involve computer-aided diagnosis (CAD) methods for automatic image classification [4]. The increasing advancement in the field of medical imaging technology, medical research and diagnosis become easy. Various types of imaging modalities and procedures are included in medical imaging technology that helps in diagnosis and treatment of the patients. Hence, it plays a dominant role in deciding the actions for the benefit of the patient's health.

In past few years, artificial intelligence (AI) brings a new way for analysing and interpreting data that also called predictive analysis that helps to identify the early signs of any of the health conditions [5]. Deep learning, an approach of AI, emerged as an outperformer in interpreting and analysing image data. Significant advancement has been made in the field of medical image diagnosis that improves disease diagnosis process significantly. Deep learning uses the architecture of artificial neural network that mimics the working of human brain. The complex computer vision tasks are effectively solved by the deep learning algorithms such as image recognition, classification and segmentation. The special class of deep learning algorithms is known as a convolutional neural network (CNN) is widely used to solve image classification problems and achieved a significant performance on benchmark datasets. The reason behind popularity of CNN is the large availabilities of datasets and support of powerful Graphics Processing Units (GPU) that makes the integration of deep learning methods with computer vision popular [6].

There are several distinct imaging modalities in which biomedical images are generated. They are different in shape and types. Due to the diverse data distribution patterns, it may happen that the same CNN model may show different performance on different datasets [7]. CNN models are sensitive to the particulars of the training data. This makes possible that each time they are trained; they may find a different set of weights. These different predictions generate high variance [8]. Moreover, these deep features face the problem of small intra-class variance and large inter-class variance [9].

To address the above issue, a novel spatial feature fusion-based approach for biomedical image classification based on ensemble Deep CNN and transfer learning is proposed. We

have used stacked generalization that is an ensemble learning method. We have used four benchmarked pre-trained CNN models as a deep feature extractor. Extracted features from different networks are merged using spatial feature fusion method. Finally, two FC (fully connected) layers with ReLU are used along with one FC layer with softmax activation function.

The major contributions of the paper include: (a) To propose an ensemble learning framework for creating a deep feature extractor that combines transfer learning features from more than one pre-trained CNN models (b) To apply a spatial feature fusion technique that creates a generalized feature space from extracted features (c) To proposed a deep CNN model that can be used for biomedical image classification with increased prediction accuracy.

## II. RELATED WORK

Health care is one of the fastest growing sectors that is delivered by health experts. It is sometimes difficult to recognize disease patterns from huge number of medical images. The special classes of Artificial Intelligence (AI) are machine learning and deep learning algorithms. These algorithms give impressive results for classification of biomedical images. Image classification is considered as one of the computer visions tasks. Many researchers have been worked in biomedical classification that resulting several robust methods that can be categorized into two types; traditional digital image processing techniques and deep learning models. The traditional image processing methods involves manual feature extractions methods whereas deep learning models perform the feature extraction without manual intervention.

N. Sharma et al [10] used texture primitive features to perform segmentation along with classification of medical images. They have used artificial neural network (ANN) for designing the algorithm that performs segmentation and classification. The algorithm has been used for CT and MRI images of distinctive body parts like brain and liver. Ahmed M. Sayed [11] applied machine learning techniques for diagnosis of breast cancer from MRI images. The highest classification accuracy is 94/6% given by KNN. M. I. Daoud et al [12] proposed a fusion approach based on multiple ROI for classification of breast ultrasound image. GLCM texture features are extracted in each ROI that are further classified using binary SVM classifier. The dataset considered for the experiment contains 64 benign and 46 malignant tumor images. The proposed approach provides very promising results with 98.2%, 98.4%, and 97.8% values for accuracy, specificity, and sensitivity respectively. P. Chak et al [13] proposed an Artificial Neural Network (ANN) and SVM based approach to classify the kidney stone images. To extract the features from the CT images, GLCM method was used. The ANN approach gives 95% accuracy, whereas SVM approach gives 99% accuracy. P. Nanglia et al [14] proposed a hybrid algorithm for classification of lung cancer images. The dataset considered for

an experiment contains 500 images and the overall accuracy of the approach is 98.08%.

Deep learning is effectively applied for various domains, including satellite imaging, observation frameworks, mechanical and medical procedures, and precision agriculture. Several researchers have worked upon applying deep convolutional neural network for different medical applications. H. T. Nguyen et al [15] improved the prediction of disease using shallow CNN. They have applied data visualization techniques on Metagenomic data and achieved promising results. S. M. Anwar et al [16] proposed deep transfer learning and LDA (Linear Discriminant Analysis) based approach for classification of medical image modality. They have considered pre-trained ResNet-50 model to implement transfer learning approach along with LDA approach. For experiment, a benchmark ImageCLEF-2012 dataset was considered. The classification accuracy obtained is 87.91% that is significantly better as compared to the state-of-the-art approaches. C.-H. Chiang et al [17] applied CNN for automatic classification of medical image modality. They have considered multiple image modalities that include CT images of abdomen & brain and MRI images of brain and spine. The accuracy achieved for validation and test sets are greater than 99.5%. Moreover, the F1-score for each of the category is greater than 99%.

B. P. Battula and D. Balaganesh [18] propose a hybrid model for medical image classification based on CNN and Encoder. HIS2828 and ISIC2017 are the datasets considered for the experiment. The results show that the accuracy of the proposed model is better than the existing models. A. A. Gomaa et al [19] discussed about how CNN and GANs are used to improve early prediction of plant disease. They have considered a tomato plant images that are infected with Tomato Mosaic Virus for conducting an experiment. The proposed CNN provides 97% accuracy. S. Patel [20] classified bacterial colony images using Atrous convolution using transfer learning approach. The dataset considered for the experiment contains 660 bacterial colony images classified into 33 distinct classes. The proposed model replaces the standard convolution layer of the VGG-16 pre-trained model with the Atrous convolution. The training and validation accuracy obtained from the experiment are 95.06% and 93.38% respectively.

## III. PROPOSED METHOD

In this research, we have proposed a novel spatial feature fusion based Deep CNN model for biomedical image classification. We have developed a feature fusion network using Ensemble deep CNN that leverages the power of state-of-the-art pre-trained CNNs. Ensemble learning is the technique that combines different individual CNN models in order to increase prediction accuracy through generalization [21]. As shown in Fig. 1, the proposed model has three stages, (a) deep feature extraction and feature maps (b) spatial feature fusion of deep transfer learning features and (c) classification.

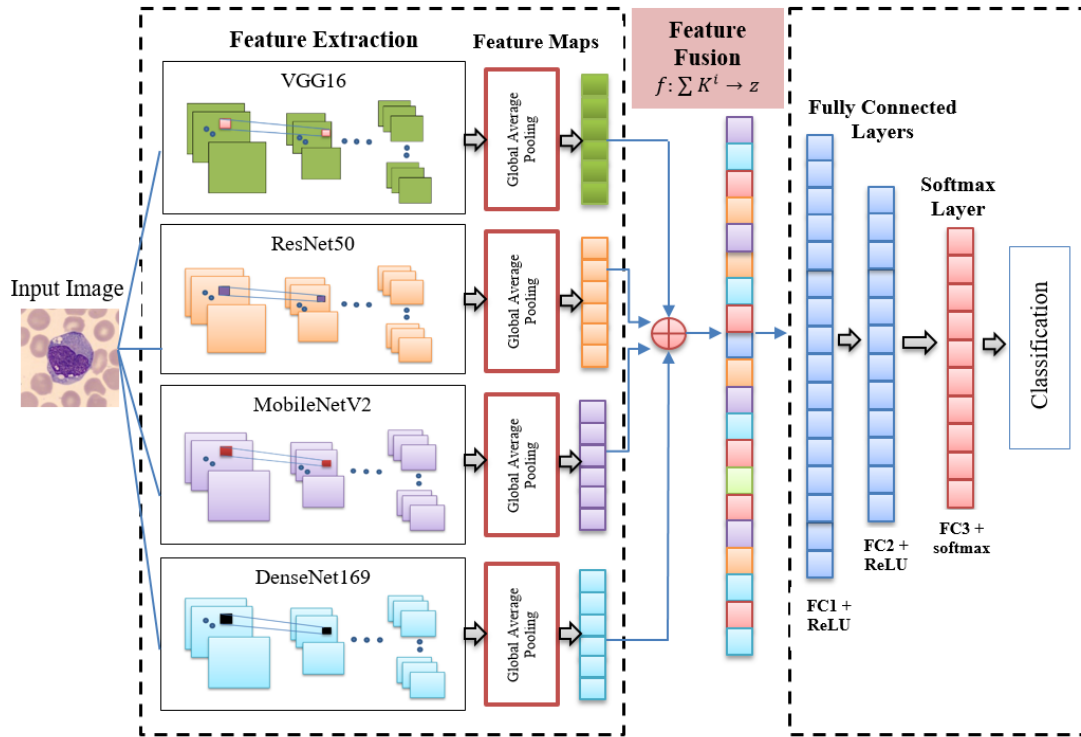


Fig. 1. Spatial Feature Fusion based Deep CNN Model using Ensemble Deep CNN and Transfer Learning.

### A. Deep Feature Extraction and Feature Maps

Feature extraction and feature map generation is the crucial step for classification of images using deep learning networks. A deep learning network normally requires a large amount of resources and dataset to be trained for precise feature extraction. The common practice followed is to use the pre-trained model instead of building and training the CNN model scratch. There are several pre-trained CNN models already exist which are developed to solve a similar problem [22]. All modern pre-trained CNN models have emerged from the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) challenge [23]. The challenge is organized yearly from 2010 to 2017 that provides a platform for researchers to present their algorithms related to object localization and image classification tasks. The pre-trained models belong to ILSVRC are trained on ImageNet dataset that consists about 14 million images categorized in several classes. The pre-trained models used the concept of transfer learning, i.e. knowledge gained by solving one problem can be applied to solve similar related problems by using an already trained model to learn a different set of features. It eliminates the need of large amount of data with significant better performance along with reduced convergence time [24]. The commonly used pre-trained models are VGG16, VGG19, MobileNet, InceptionNet, ResNet, DenseNet, etc.

We have followed stacked generalization ensemble [25] approach to ensemble four pre-trained CNN models. For that, as the preliminary step, four benchmarked CNN models i.e. VGG19, ResNET50, MobileNet and DenseNet are considered as base models for the proposed novel model. We have reused these pre-trained networks with the parameter they have

learned on ImageNet dataset. Also, we have removed their final softmax classification layer as it contains 1000 neurons. We have frozen all the intermediate layers of base models to keep their original trained weights. After that, the global average pooling is applied for dimensionality reduction and an output feature vector is created for each of the base model.

### B. Spatial Feature Fusion of Deep Transfer Learning Features

Fusing the features obtained from various pre-trained CNN models supported many applications to achieve better performance than the conventional approach of utilizing the single deep CNN network for the task of classification [26]. The classification approach presented in this paper maps the feature space obtained by various pre-trained CNN models like VGG16, ResNet50, MobileNetV2 and DenseNet169 to a generalized feature space [26]. The feature space generated by each feature extractor with different dimension is represented as:

$$f_i = \{F_i^{(m)}\}_{m=1}^j \quad (1)$$

Where  $i=1, 2, \dots, N$  represents the number of images in the dataset and  $m=1, 2, \dots, j$  represents the number of pre-trained CNN models used as feature extractors. The cumulative feature set for a particular image sample is defined as:

$$f_k = \{f_k^1, f_k^2, f_k^3, \dots, f_k^l\} \quad (2)$$

Where,  $f_k^i$  is representation of  $i$ th dimension feature in the given  $k$ th sample and  $i, i \in [1, D]$  represents the individual feature dimension. Thus, features obtained by the particular

CNN model forms individual feature space [27]. This, individual feature space is defined as:

$$\psi = \{f_1, f_2, \dots, f_n\} \quad (3)$$

Where  $f_i, i \in [1, n]$  is the feature set of individual sample and  $n$  represents the total number of samples in the dataset.

Feature fusion technique combines heterogeneous features obtained from various CNN models and utilize combined features for comprehensive processing for the cumulative decision-making. The feature fusion technique combines feature spaces obtained from individual CNNs and provide feature subspace, which is generalized than the original feature space. The feature fusion technique to create a generalized feature space is represented as:

$$f: (\psi_p, \psi_q, \psi_r, \psi_s) \rightarrow \psi \quad (4)$$

Where,  $f$  is future fusion function and  $\psi_p, \psi_q, \psi_r$  and  $\psi_s$  represents the feature spaces obtained by individual CNN model. The data imbalance and noise are the major limitations of the features obtained using individual CNNs. However, obtaining generalize feature vector by combining different feature spaces helps in selecting important features captured by different CNN models which leads to more accurate classification accuracy.

### C. Spatial Feature Fusion

In this research, the spatial feature fusion algorithm fuses the feature maps obtained by four different deep pre-trained CNN models [28]. Hence, four pre-trained CNN models are connected with each other using feature fusion techniques and point of connection between four models is known as fusion point. The training of softmax classifier is the next step after the fusion point in order to achieve the result of the classification as represented in Fig. 1. The spatial feature fusion function represented as:

$$f: K^p + K^q + K^r + K^s \rightarrow z \quad (5)$$

Where,  $K^p, K^q, K^r$  and  $K^s$  presents the set of features extracted by feature extractor P, Q, R and S respectively. Here,  $z$  denotes the fusion space features' set and  $K^p, K^q, K^r, K^s, z \in \psi^{LWD}$  where L states the length, W states the width and D states the channels of the feature set, respectively [28].

### D. Classification

After feature extraction, it is required to classify the data into various classes. This can be achieved using fully connected layers. In the proposed model, the third stage is performing classification. We have obtained a final feature vector after performing spatial feature fusion. For classification, two FC layers with ReLU and one FC layer with softmax are added to the proposed network as depicted in Fig. 1. A Rectified Linear Unit (ReLU) is used as an activation function for the first two fully connected layers. As the network classifies the input image into eight distinct classes, the last fully connected layer is applied with the softmax activation function.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

The details regarding dataset, experimental setup and results are discussed in detail in this section. We have implemented the proposed ensemble CNN model using spatial feature fusion along with single model of VGG16, ResNet50, MobileNetV2 and DenseNet169.

### A. Dataset

The experiment was performed on the microscopic images dataset that represents peripheral blood cell images [29]. The dataset contains 17902 microscopic jpg images with size of 360 X 363 pixels. The images are captured in the Core Laboratory at the Hospital Clinic of Barcelona. The images are taken using the analyzer CellaVision DM96. These all images are annotated by the clinical expert pathologists. The dataset contains very high-quality labelled images that can be used to train any algorithms belonging to machine learning and deep learning. The dataset categorized into eight distinct groups that are platelet, neutrophil, monocyte, lymphocyte, immature granulocytes, erythroblast, eosinophil and basophil. The dataset is specifically used for hematological diagnosis using computational tools. All the images of the dataset are acquired from the year 2015 to 2019 [30]. Fig. 2 represents the sample images from platelet, neutrophil and monocyte class.

The proportion of images in the dataset is represented in Fig. 3. As mentioned earlier there are 17902 microscopic images that categorized into 8 distinct classes.

From the dataset, 80% data are used for training and 20% data are used for validation.

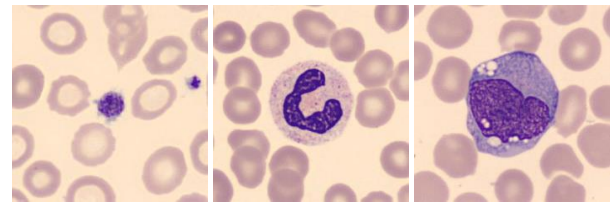


Fig. 2. Sample Images of Platelet, Neutrophil and Monocyte Classes.

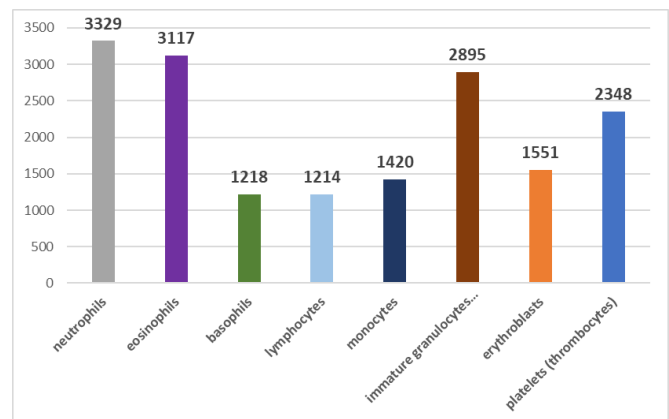


Fig. 3. Distribution of Microscopic Peripheral Blood Cell Dataset.

### B. Environment Setup

The experiment is carried out on a workstation configured with an Intel® Core™ i7 8th generation processor, 32GB RAM, and NVIDIA Titan XP GPU with 64-bit Windows 10 operating system. To build the model for microscopic peripheral blood cell images classification, TensorFlow and Keras [31] API are used. TensorFlow is an open-source library and it provides many of the high-level and low-level APIs. Keras is one of the high-level API that built upon the TensorFlow library. It is a user-friendly and extensible library that makes the building of deep neural network fast. Keras offers various building blocks like activation functions, layers, optimizers, batch normalization etc. that are required to build a convolutional neural network. It provides the functionalities to run the CNN on a graphical processing unit (GPU) or tensor processing unit (TPU).

### C. Network Training and Hyperparameter Tuning

One of the pre-requisite for any of the CNN model deployment is Training. One of the good practices is to use the pre-trained model instead of building and training the CNN model from scratch. Transfer learning is about re-using the pre-trained CNN model for solving a new problem [32]. All benchmark pre-trained CNN models are trained on the large dataset like ImageNet, PASCAL, COCO, etc. In this research, the weights along with learning of pre-trained model are transfer in the CNN model. Also, a fine-tuning process is performed by unfreezing the few of top layers to retrain them for classification purpose [33]. The dataset considered in the experiment is a small dataset as it contains 17902 images categorized into 8 distinct classes. Transfer learning and fine-tuning is performed by removing the last layer and introducing fully-connected softmax layer for the purpose of classification.

The basic building block of CNN model is artificial neural network. Thus, CNN models possess self-learning capability. For that, during training process, the weight of the network is adjusted in order to minimize the loss function [34]. There are set of parameters which controls the entire training process [35][36]. In this experiment, model specific hyperparameters are used. Table I summarizes the hyperparameters along with their optimized value.

TABLE I. OPTIMIZED VALUES FOR HYPERPARAMETERS

Hyperparameter (s)	Values
Number of Epochs	20
Learning rate	0.001
Batch size	32
Optimizer	SGD
Activation Function	ReLU / softmax
Dropout Rate	0.5

### D. Evaluation Metrics

The benchmark evaluation metrics such as Precision, Recall, and F1-score are used to evaluate the performance of the proposed model. The equations used for the evaluation metrics are as per following.

$$\text{Precision}(r) = \frac{TP(r)}{TP(r)+FP(r)} \quad (6)$$

$$\text{Recall}(r) = \frac{TP(r)}{TP(r)+FN(r)} \quad (7)$$

$$F1 - \text{Score}(r) = \frac{2}{\frac{1}{\text{Recall}(r)} + \frac{1}{\text{Precision}(r)}} \quad (8)$$

Here, TP and FP are the true positive and true negatives, respectively. TN and FN are the true negatives and false negatives, respectively. To get the values of TP, TN, FP and FN, N X N confusion matrix is used, where N represents the number of classes available in the dataset [38]. The confusion matrix compares the actual values with the predicted values predicted by CNN model. It represents the values for TP, TN, FP and FN [37].

## V. RESULT AND DISCUSSION

In this section, the performance of the proposed ensemble CNN model along with the performance of each individual CNN model i.e. VGG16, ResNet50, MobileNetV2 and DenseNet169 is discussed. To evaluate the performance of the model, two metric namely accuracy and loss are considered. The average accuracy and average loss for training and validation are measured. Accuracy is a metric used to measure the performance of a model. The accuracy is defined as:

$$\text{Accuracy} = \frac{TP+FN}{\text{Total no of samples}} \quad (9)$$

Where TP is the true positive and FN indicates false negatives [38].

Loss is defined as an error happened in prediction by a model. It is calculated by the difference between the value predicted by a CNN model and an actual value present in the dataset. While training the CNN model, the aim is to decrease the loss by optimizing the weights. Usually, two functions are used to measure a loss namely mean square error and cross-entropy. The cross-entropy is used to measure the loss and the formula is defined by the following equation [39].

$$\text{Cross - entropy} = - \sum_{i=1}^n \sum_{j=1}^n y_{i,j} \log(p_{i,j}) \quad (10)$$

Table II shows the values obtained for training and validation accuracy and loss respectively. It can be observed that the proposed CNN model achieved the highest accuracy i.e. 96.45% during training and 96.10% during validation. Moreover, it has the lowest loss i.e. 0.028 during training and 0.024 during validation. It shows that the proposed CNN model learns well as compare to the other models.

TABLE II. ACCURACY AND LOSS VALUES

Deep Learning Model	Training		Validation	
	Accuracy (%)	Loss	Accuracy (%)	Loss
Proposed CNN Model	96.45	0.028	96.10	0.024
DenseNet169	93.10	0.036	92.69	0.039
MobileNetV2	92.78	0.045	92.91	0.042
ResNet50	92.10	0.059	91.75	0.063
VGG16	87.56	0.097	86.80	0.089

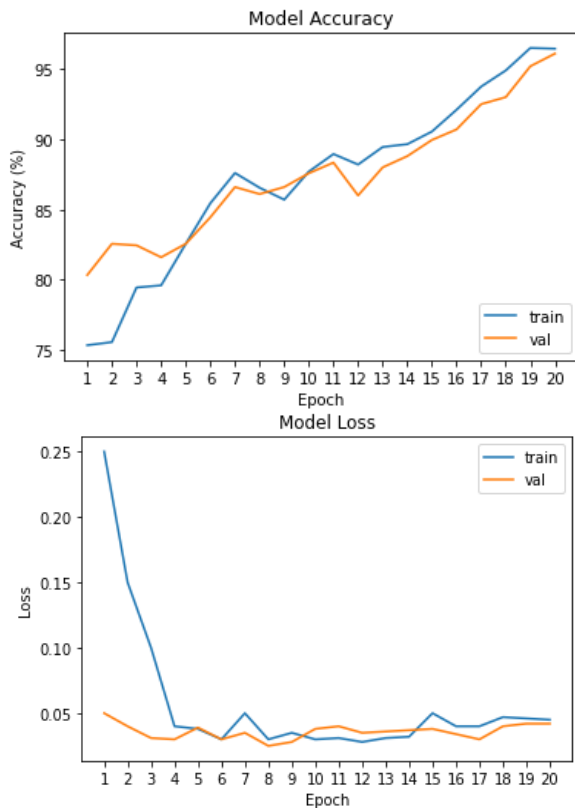


Fig. 4. Accuracy and Loss for the Proposed CNN Model.

The proposed CNN model also converges adequately and it can be shown from Fig. 4. The figure represents the accuracy and loss curves obtained during training and validation. Fig. 4 shows that the curves for both; training and validation for accuracy are going parallel and close to each other. Moreover, the loss curves are also going parallel and close to each other. This shows that model is trained adequately without facing the issues of underfitting and overfitting.

Table III represents the values for the precision, recall and F1-score for the different models obtained during experiment.

The F1-score, Precision and Recall values for VGG16 are 87.30%, 87.45% and 87.15 respectively. For ResNet50, the values are 92.17%, 92.25% and 92.10% respectively. MobileNetV2 provides 92.76%, 92.75% and 92.82% for F1-score, Precision and Recall. DenseNet169 gives 93.02%, 92.90% and 93.15% for F1-score, Precision and Recall. However, the proposed CNN model outperforms all other individual pre-trained CNN models and provides the highest values i.e. 96.55%, 96.40% and 96.70% for F1-score, Precision and Recall respectively.

From the results, it can be concluded that the spatial feature fusion approach for classification of microscopic blood cell images provides better results with enhancement in learning ability. The generalized transfer learning feature space developed by deep feature extractor reduces the problem of bias and variance while increase the performance. Hence, the proposed CNN model has better performance than the individual pre-trained CNN models.

TABLE III. PRECISION, RECALL AND F1-SCORE VALUES

Deep Learning Model	F1-Score (%)	Precision (%)	Recall (%)
<b>Proposed CNN Model</b>	96.55	96.40	96.70
<b>DenseNet169</b>	93.02	92.90	93.15
<b>MobileNetV2</b>	92.76	92.75	92.82
<b>ResNet50</b>	92.17	92.25	92.10
<b>VGG16</b>	87.30	87.45	87.15

## VI. CONCLUSION

Deep learning models provide promising results in biomedical image analysis. It brings a new way for analysing and interpreting data that helps to identify the early signs of health conditions. A spatial feature fusion approach for biomedical image classification based on ensemble deep CNN and transfer learning is proposed in this paper. For that, the generalized transfer learning feature space is developed and a spatial fusion is applied to merge the learned features of different pre-trained CNNs. The paper covers the details of implementation and evaluation of the most proposed CNN model for classifying microscopic peripheral blood cell images. The dataset contains 17902 images of blood cell that are categorized into 8 class labels. The experiment shows that the proposed ensemble CNN model outperforms individual pre-trained CNN model and provides better precision, recall and F1-score values.

## REFERENCES

- [1] "Biomedical imaging & image processing," *Embs.org*. [Online]. Available: <https://www.embs.org/about-biomedical-engineering/our-areas-of-research/biomedical-imaging-image-processing>. [Accessed: 19-Jan-2022].
- [2] "The important role medical imaging plays in diagnosis and treatment :: PBMC health," *Pbmchealth.org*. [Online]. Available: <https://www.pbmchealth.org/news-events/blog/important-role-medical-imaging-plays-diagnosis-and-treatment>. [Accessed: 19-Jan-2022].
- [3] U. F. O. Themes, "The role of imaging in medicine," *Radiology Key*, 09-Aug-2020. [Online]. Available: <https://radiologykey.com/the-role-of-imaging-in-medicine>. [Accessed: 19-Jan-2022].
- [4] Z. Lai and H. Deng, "Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron," *Comput. Intell. Neurosci.*, vol. 2018, p. 2061516, 2018.
- [5] F. Pesapane, M. Codari, and F. Sardanelli, "Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine," *Eur. Radiol. Exp*, vol. 2, no. 1, 2018.
- [6] I. Patel and S. Patel, "An optimized deep learning model for flower classification using NAS-FPN and Faster R-CNN," *Ijstr.org*. [Online]. Available: <http://www.ijstr.org/final-print/mar2020/An-Optimized-Deep-Learning-Model-For-Flower-Classification-Using-Nas-fpn-And-Faster-R-cnn.pdf>. [Accessed: 19-Jan-2022].
- [7] H.-H. Zhao and H. Liu, "Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition," *Granul. Comput.*, vol. 5, no. 3, pp. 411–418, 2020.
- [8] M. Hassaballah, A. A. Abdelmgeid, and H. A. Alshazly, "Image features detection, description and matching," in *Image Feature Detectors and Descriptors*. Cham: Springer International Publishing, 2016, pp. 11–45.
- [9] S. Ge, C. Bai, Y. Liu, Y. Liu, and T. Zhao, "Deep and discriminative feature learning for fingerprint classification," in *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, 2017, pp. 1942–1946.
- [10] N. Sharma, A. K. Ray, S. Sharma, K. K. Shukla, S. Pradhan, and L. M. Aggarwal, "Segmentation and classification of medical images using

- texture-primitive features: Application of BAM-type artificial neural network,” *J. Med. Phys.*, vol. 33, no. 3, pp. 119–126, 2008.
- [11] A. M. Sayed, “Machine learning augmented breast tumors classification using magnetic resonance imaging histograms,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 12, 2021.
- [12] M. I. Daoud, T. M. Bdair, M. Al-Najar, and R. Alazrai, “A fusion-based approach for breast ultrasound image classification using multiple-ROI texture and morphological analyses,” *Comput. Math. Methods Med.*, vol. 2016, p. 6740956, 2016.
- [13] P. Chak, P. Navadiya, B. Parikh, and K. C. Pathak, “Neural network and SVM based kidney stone based medical image classification,” in *Communications in Computer and Information Science*, Singapore: Springer Singapore, 2020, pp. 158–173.
- [14] P. Nanglia, S. Kumar, A. N. Mahajan, P. Singh, and D. Rathee, “A hybrid algorithm for lung cancer classification using SVM and Neural Networks,” *ICT Express*, vol. 7, no. 3, pp. 335–341, 2021.
- [15] H. T. Nguyen, T. Bao, H. Hoang, T. Phuoc, and Nghi, “Improving disease prediction using shallow convolutional neural networks on metagenomic data visualizations based on mean-shift clustering algorithm,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 6, 2020.
- [16] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, “Medical image analysis using convolutional neural networks: A review,” *J. Med. Syst.*, vol. 42, no. 11, p. 226, 2018.
- [17] C.-H. Chiang, C.-L. Weng, and H.-W. Chiu, “Automatic classification of medical image modality and anatomical location using convolutional neural network,” *PLoS One*, vol. 16, no. 6, p. e0253205, 2021.
- [18] B. P. Battula and D. Balaganesh, “Medical image data classification using deep learning based hybrid model with CNN and encoder,” *Rev. d intell. artif.*, vol. 34, no. 5, pp. 645–652, 2020.
- [19] A. A. Goma and Y. M. A. El-Latif, “Early Prediction of Plant Diseases using CNN and GANs,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 5, 2021.
- [20] S. Patel, “Bacterial colony classification using atrous convolution with transfer learning,” *Annals of RSCB*, pp. 1428–1441, 2021.
- [21] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, “Ensemble deep learning: A review,” *arXiv [cs.LG]*, 2021.
- [22] S. Patel and R. Patel, “A comprehensive analysis of A Comprehensive Study of Applying Convolutional Neural Network for Computer Vision,” *Int. j. adv. sci. technol.*, vol. 29, no. 6s, pp. 2161–2174, 2020.
- [23] O. Russakovsky *et al.*, “ImageNet large scale visual recognition challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.
- [24] N. Ganatra and A. Patel, “Deep learning methods and applications for precision agriculture,” in *Machine Learning for Predictive Analysis*, Singapore: Springer Singapore, 2021, pp. 515–527.
- [25] J. Brownlee, “Stacking ensemble for deep learning neural networks in python,” *Machine Learning Mastery*, 30-Dec-2018. [Online]. Available: <https://machinelearningmastery.com/stacking-ensemble-for-deep-learning-neural-networks/>. [Accessed: 21-Mar-2022].
- [26] T. Akilan, Q. M. J. Wu, Y. Yang, and A. Safaei, “Fusion of transfer learning features and its application in image classification,” in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2017, pp. 1–5.
- [27] Y. Lavinia, H. H. Vo, and A. Verma, “Fusion based deep CNN for improved large-scale image action recognition,” in *2016 IEEE International Symposium on Multimedia (ISM)*, 2016, pp. 609–614.
- [28] S. Zhou, S. Cai, C. Zeng, and Z. Wang, “A cotton and flax fiber classification model based on transfer learning and spatial fusion of deep features,” in *Advances on P2P, Parallel, Grid, Cloud and Internet Computing*, Cham: Springer International Publishing, 2020, pp. 152–162.
- [29] A. Acevedo, S. Alférez, A. Merino, L. Puigví, and J. Rodellar, “Recognition of peripheral blood cell images using convolutional neural networks,” *Comput. Methods Programs Biomed.*, vol. 180, no. 105020, p. 105020, 2019.
- [30] A. Acevedo, A. Merino, S. Alférez, Á. Molina, L. Boldú, and J. Rodellar, “A dataset of microscopic peripheral blood cell images for development of automatic recognition systems,” *Data Brief*, vol. 30, no. 105474, p. 105474, 2020.
- [31] Keras Team, “Keras: the Python deep learning API,” *Keras.io*. [Online]. Available: <https://keras.io>. [Accessed: 19-Jan-2022].
- [32] A. P. Ganatra, “Performance analysis of fine-tuned convolutional neural network models for plant disease classification,” *Int. J. Contr. Autom.*, vol. 13, no. 03, pp. 293–305, 2020.
- [33] J. Browarczyk, A. Kurowski, and B. Kostek, “Analyzing the effectiveness of the brain-computer interface for task discerning based on machine learning,” *Sensors (Basel)*, vol. 20, no. 8, p. 2403, 2020.
- [34] A. Brodzicki, J. Jaworek-Korjakowska, P. Kleczek, M. Garland, and M. Bogyo, “Pre-trained deep convolutional neural network for *Clostridioides difficile* bacteria cytotoxicity classification based on fluorescence images,” *Sensors (Basel)*, vol. 20, no. 23, p. 6713, 2020.
- [35] K. Nongthombam and D. Sharma, “Data Analysis using Python,” *Int. J. Eng. Res. Technol. (Ahmedabad)*, vol. 10, no. 7.
- [36] R. Susmaga, “Confusion Matrix Visualization,” in *Intelligent Information Processing and Web Mining*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 107–116.
- [37] L. Cavallaro, O. Bagdasar, P. De Meo, G. Fiumara, and A. Liotta, “Artificial neural networks training acceleration through network science strategies,” *Soft Comput.*, vol. 24, no. 23, pp. 17787–17795, 2020.
- [38] I. Patel and S. Patel, “Flower Identification and Classification using Computer Vision and Machine Learning Techniques”, *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8, no. 6, pp. 277–285, 2019.
- [39] E. Gordon-Rodríguez, G. Loaiza-Ganem, G. Pleiss, and J. P. Cunningham, “Uses and abuses of the cross-entropy loss: Case studies in modern deep learning,” *arXiv [stat.ML]*, pp. 1–10, 2020.