

Pre-Trained Convolutional Neural Networks as Feature Extractors Towards Improved Malaria Parasite Detection In Thin Blood Smear Images

1st P. Akshay Kumar

Roll Number : MT19094

Indraprastha Institute of Information Technology (IIIT-D)

Delhi, India

akshay19094@iiitd.ac.in

2nd Gaurav Lodhi

Roll Number : MT19063

Indraprastha Institute of Information Technology (IIIT-D)

Delhi, India

gaurav19063@iiitd.ac.in

3rd Ritesh Singh

Roll Number : MT19044

Indraprastha Institute of Information Technology (IIIT-D)

Delhi, India

ritesh19044@iiitd.ac.in

4th Kaushal Sanadhya

Roll Number : MT19133

Indraprastha Institute of Information Technology (IIIT-D)

Delhi, India

Kaushal19133@iiitd.ac.in

Abstract—Malaria is a life-threatening disease caused by parasites called Plasmodium, which is transmitted to people through the bites of infected female Anopheles mosquitoes. This disease is preventable and curable. In 2018, 228 million malaria cases were reported worldwide, with a total of 405,000 estimated deaths. The standard method for diagnosing malaria is Light microscopy of thick and thin stained blood smears. This process involves the collection of blood smears; it's staining with Romanowski stains and analyzing RBC for Plasmodium. It is a very tiresome and time-consuming process. State-of-the-art Support Vector Machine (SVM) with feature extraction techniques like Oriented FAST and Rotated BRIEF (OFRB) and watershed segmentation is used to diagnose the disease. The accuracy of such methods highly depends upon the nature of the feature selection algorithm. On the other hand, deep learning models like convolutional neural networks (CNN) can perform end-to-end feature extraction and classification with highly promising accuracy. But there is no such thing as free lunch. Tuning of hyperparameters like learning rate, the number of hidden layers, selection of loss function, etc in these deep learning models is a challenging task. Fastai is a deep learning library that provides high-level components that can quickly and easily provide state-of-the-art results in standard deep learning domains. It also enables researchers to customize all the deep learning components to build up new approaches. Fastai also includes GPU-optimized libraries for faster computation. We have used these Fastai libraries to develop deep learning models in our work. Well known CNN models, namely Dense Net, Alex Net, Res Net, Squeeze Net, and vgg11bn, are used on our dataset. These CNN models with pre-trained mode were used to train them on our dataset further. Moreover, fivefold cross-validation is performed to avoid overfitting. We have achieved accuracy up to 0.97, which is highest reported in the literature so far. An ensemble of these models to further improve the performance is our future work.

I. INTRODUCTION

Malaria is a mosquito-borne infectious disease that affects humans. It is caused by the Plasmodium parasites transmitted

through the bite of female Anopheles mosquito. Symptoms of malaria typically include fever, tiredness, vomiting, and headaches. In extreme cases, it can cause yellow skin, seizures, coma, or death. Experts usually examine thick and thin blood smears to detect and diagnose this disease and compute parasitemia. This process is highly time-consuming and requires a certain amount of expertise to classify and count parasitized and uninfected cells of the human body.

Therefore, there is a high need for other verticals to perform this burdensome work. Already existing Machine learning [1] techniques like state-of-the-art Support vector machine (SVM) and Gradient Boosting combined with popular feature extraction techniques like Speeded-Up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB) and KAZE are quite efficient to develop an efficient solution to this problem.

Although these methods are popular yet suffer from various problems, the results produced by these methods highly depends upon the selected feature extraction algorithm. Moreover, there is no such feature extraction technique which is universal in this field. Sometimes they also fail to retrieve features from the images. These issues positively identify the limitations of the existing models.

Therefore we moved on to deep convolutional neural networks (CNN) [2] due to their auto feature extraction properties. The features are learned while CNN is being trained on the images through hundreds of its hidden layers. We have used Tensorflow API to develop the baseline neural network. Further, we have used Fast AI to implement CNN models due to its efficient deep learning models, which can be used to solve image analysis problems quickly and efficiently. Fast AI automatically solves the problem of hitting local minima and provides advanced techniques like transfer learning and differential learning rates. We have used the below models

after establishing our own baseline model:

- Alex Net [3]
- Dense Net [4]
- Res Net [5]
- Squeeze Net [6]
- VGGBN 11 [7]

These CNN models perform exceptionally well on the well-known Image net dataset challenge; hence we used these models for our problem. Moreover, we have used pre-trained versions of these models and further trained them on our dataset to achieve high accuracy.

II. MATERIALS AND METHOD

A. Dataset Description

The dataset [8] consists of 27,558 images of human cells from the support of the National Library of Medicine (NIH). Parasitized and the uninfected set contains 13,779 images each. All the photos are in RGB format with size 139 X 142 X 3. The notable feature of this dataset is that the cell images are not pre-processed in terms of light intensity, color distribution, etc.

B. Category Distribution

The images are divided into two classes, namely, Parasitized and Uninfected. Sample images are shown in Fig 1 and 2.

Fig. 1. Sample Image of Uninfected Cell

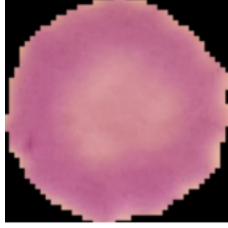


Fig. 2. Sample Image of Parasitized Cell



The dataset is balanced with the equal number of images in each category, as shown in the given pie chart. Since the images were captured with different light intensities, we have performed image normalization before feeding them into CNN models for training. The number of images is good enough to train an unbiased neural network. Therefore, there was no need for image augmentation or oversampling.

C. Feature Generation for Baseline Model

Before training the model we pre-processed all the images in Dataset by using the OpenCV library. It pre-processes the images in the following steps.

- Resized the image from its original size to 75x75 and standardized all the images to the same size for the entire process.
- performed smoothing which reduces noise of the images.
- Transformed the images from original RGB to another YUV colour space.
- Using histogram equalization we changes lower local contrast area to gain a higher contrast which makes images more sharpen.
- Finally, we changed the YUV colour space to RGB colour space back. These transformed images further used as input to the sequential model.
- We have shuffled all the images with their labels and splitted them in ratio of 70:30. Here 70% of the images used for train the model and remaining 30% of the images are used for validation purpose.

D. Implementation of Baseline Convolutional Neural Network

This is a layered sequential model that consists of 3 convolutional layers with two pooling layer in between(to reduce amount of parameters and speedup the computation), a flatten layer which is used for flattening the input in a one-dimensional vector, a dense layer with 64 neurons with RELU activation function. It gives an output of size 64x1 to the next layer. This layer is densely connected with the incoming input features. Then a dense layer, which is a softmax layer which takes an input of 64x1 dimensional vector and gives a binary output.

Although the architecture is simple but fine-tuning of a batch size of 50 and with epoch size of 20 it performs well. It classified 94.90% of the validation images correctly to their respective classes. This model uses batch size of 50 to resolve memory resource issue because training large number of images need huge amount of memory.

E. Proposed Architecture of Fast AI models

The flow chart of proposed architecture is shown in fig 4.

- Since the images were divided into two directories "uninfected" and "parasitized". We have merged them in a single directory that can be passed to the model. We have also renamed all the images as:
 - All the images under the "uninfected" directory were renamed as "0_XXX.jpg".
 - All the images under "parasitized" directory were renamed as "1_XXX.jpg".

Fig. 3. Class Distribution of Dataset

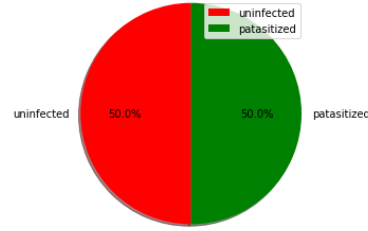
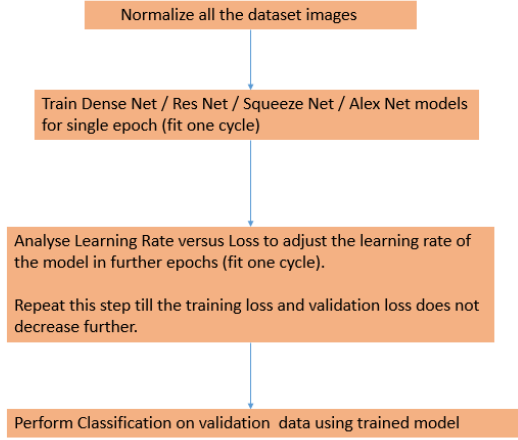
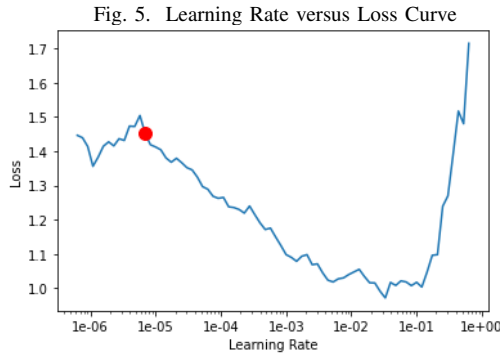


Fig. 4. Proposed Architecture



- The training images were divided into batches of size 20, each using ImageDataBunch provided by Fast.AI. The main reason behind creating the batches being the RAM constraints of the machine.
- The training and validation set were generated implicitly in the ratio of 80:20 using ImageDataBunch.
- All the images were normalized using data.normalize() method.
- Once the data is ready, the CNN model (Res Net, VGG BN 11, Dense Net, Alex Net, and Squeeze Net) is trained for a single epoch using fit_one_cycle() to find the model's optimal learning rate. It is estimated using loss versus learning rate curve. For example the suitable learning rate from the fig 5 is denoted by red dot.



- After estimating the optimal learning rate, train the model for a suitable number of epochs until training and validation loss does not decrease further.
- Perform prediction on the validation dataset.

F. Evaluation Parameters

- Accuracy : It is defined as the fraction of the number of correct predictions to the total number of predictions. This is one of the most essential evaluation metrics in the artificial intelligence field.

- Precision : It refers to the percentage of results which are relevant. The reason behind choosing this evaluation metrics is that classification of infected and uninfected malaria samples is a crucial task and the costs of having a misclassified actual positive (or false negative) is very high.
- Recall : It refers to the percentage of total relevant results correctly classified by an algorithm. Recall is the model metric which is used to select the best model when there is a high cost associated with false negative. If a parasitized cell sample undergoes the test and predicted as non-infected. The cost associated with false negative will be extremely high.
- Other than these three most important evaluation parameters we have also calculates F1-score, support, macro-average and weighted average.

G. Cross Validation

We evaluated the models through five-fold cross-validation. The cross-validation has been performed to ensure that the trained model is not overfitting. The dataset is equally divided in five folds (approximately) with an equal number of parasitized and uninfected samples in each fold.

III. RESULTS

A. Alex Net

Fig. 6. Confusion Matrix of Alex Net

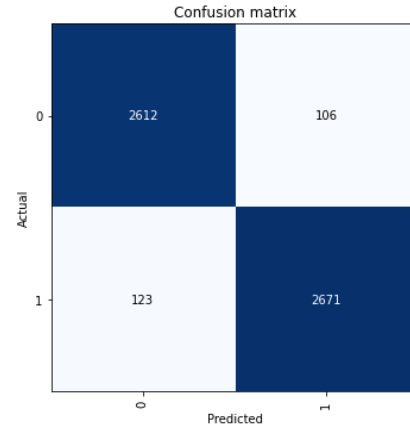


Fig. 7. Summary of Evaluation Parameters for Alex Net

	precision	recall	f1-score	support
1	0.95	0.97	0.96	2680
0	0.97	0.96	0.96	2832
accuracy			0.96	5512
macro avg	0.96	0.96	0.96	5512
weighted avg	0.96	0.96	0.96	5512

- Confusion matrix and values of evaluation parameters are summarized in Fig 6 and 7.
- Optimal learning rate to train the Alex net is 6.92 e-06.

- Precision for parasitised samples is same as that of baseline model but increased for uninfected samples.
- Total number of false positives and false negatives is 229.

B. Dense Net

Fig. 8. Confusion Matrix of Dense Net

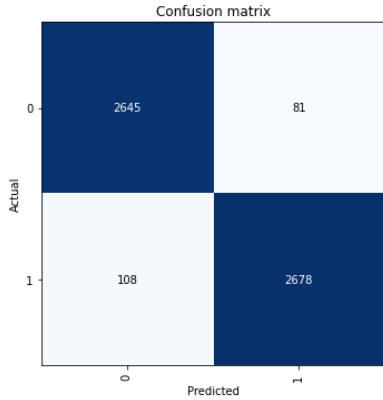


Fig. 9. Summary of Evaluation Parameters for Dense Net

	precision	recall	f1-score	support
1	0.96	0.97	0.97	2726
0	0.97	0.96	0.97	2786
accuracy			0.97	5512
macro avg	0.97	0.97	0.97	5512
weighted avg	0.97	0.97	0.97	5512

- Confusion matrix and values of evaluation parameters are summarized in Fig 8 and 9.
- Optimal learning rate to train the Dense net is between $1e-06$ and $1e-05$.
- Precision for parasitised and uninfected samples is improved over baseline model.
- Number of false positives and false negatives are least for the Dense Net model.
- Dense Net is the best implemented model so far among all models.

C. Res Net

- Confusion matrix and values of evaluation parameters are summarized in Fig 10 and 11.
- Optimal learning rate to train the Res net is between $1e-05$ and $1e-04$.
- Precision for parasitised and uninfected samples is improved over baseline model. Precision for parasitized samples is the highest among all the models implemented.
- Total number of wrongly classified samples are 196.

D. Squeeze Net

- Confusion matrix and values of evaluation parameters are summarized in Fig 12 and 13.

Fig. 10. Confusion Matrix of Res Net

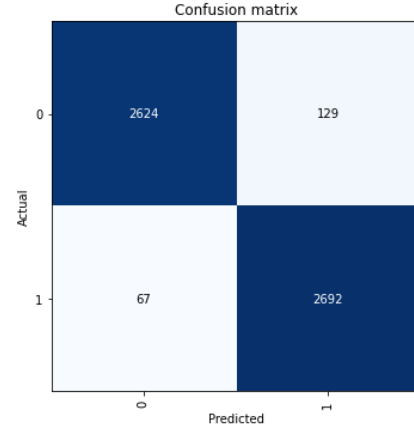


Fig. 11. Summary of Evaluation Parameters for Res Net

	precision	recall	f1-score	support
1	0.98	0.95	0.96	2753
0	0.95	0.98	0.96	2759
accuracy			0.96	5512
macro avg	0.96	0.96	0.96	5512
weighted avg	0.96	0.96	0.96	5512

Fig. 12. Confusion Matrix of Squeeze Net

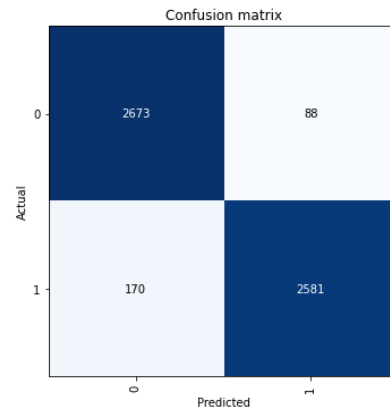


Fig. 13. Summary of Evaluation Parameters for Squeeze Net

	precision	recall	f1-score	support
1	0.94	0.97	0.95	2761
0	0.97	0.94	0.95	2751
accuracy			0.95	5512
macro avg	0.95	0.95	0.95	5512
weighted avg	0.95	0.95	0.95	5512

- Optimal learning rate to train the Squeeze net is between $1e-03$ and $1e-02$.
- Precision for uninfected samples is the highest among all the models implemented.
- Total number of wrongly classified samples are 258.

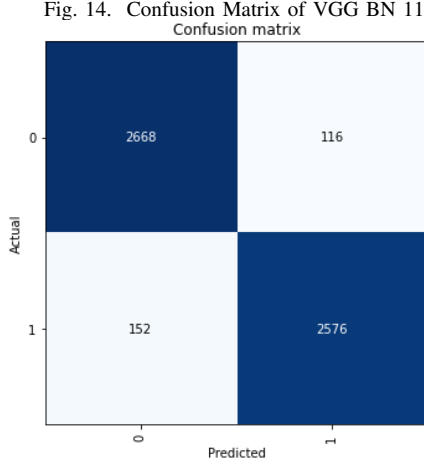


Fig. 15. Summary of Evaluation Parameters for VGG BN 11

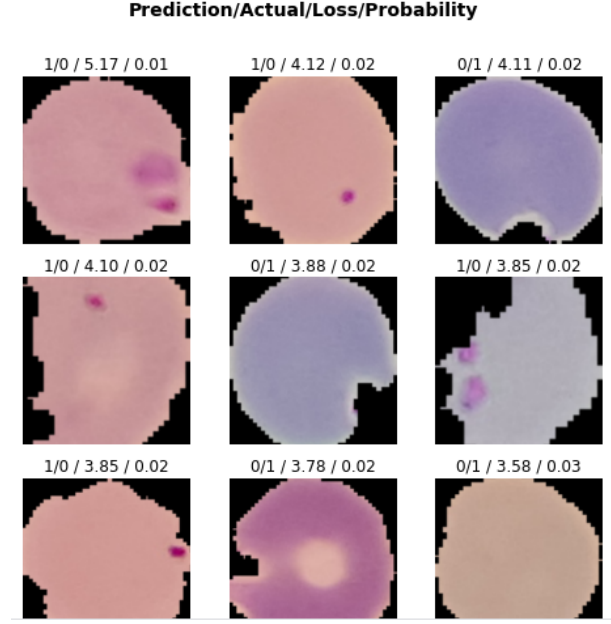
	precision	recall	f1-score	support
1	0.95	0.96	0.95	2784
0	0.96	0.94	0.95	2728
accuracy			0.95	5512
macro avg	0.95	0.95	0.95	5512
weighted avg	0.95	0.95	0.95	5512

- Confusion matrix and values of evaluation parameters are summarized in Fig 14 and 15.
- Optimal learning rate to train the Squeeze net is between $1e-06$ and $1e-02$.
- Accuracy of this model is somewhere near to the baseline model.
- Total number of wrongly classified samples are 268.

IV. DISCUSSION

The Dense Net 121 [4] outperformed the other trained models in terms of overall precision, recall and F1-score metrics. The possible reason for it could be that the densenet classifier extracts features of all complexity levels unlike other classifiers which tends to give more smooth decision boundaries. Here the only feature which any neural net has to identify is the small clot inside the cellular sample. The sample misclassified cell images associated with the top loss for Dense Net are shown in fig 14. Microscopists identify a small clot inside the cellular image for an infected sample to distinguish between parasitized and uninfected samples. It is a general pattern for correctly classifying cellular samples. Moreover, this is one pattern recognized by the neural networks during their training as well. But for wrongly classified images, this pattern is not followed, as shown in fig 14. All the models implemented for these work are associated with least top-1 and top-5 error rates, which is another reason for selected these neural networks specifically.

Fig. 16. Sample Cells With Top Loss



Fast AI provides us the functionality to use pre-trained models, which can be trained on any image dataset through transfer learning. These pre-trained models are trained on the Image Net dataset set [9], which consists of 14 Million images belonging to 1000 different classes. This highly improves accuracy and improves the feature extraction process. The Fast AI models implemented provides functionality to determine the optimal learning rate for training to reach global minima. Therefore, the selection of the learning rate for deep learning models is not empirical anymore. Moreover, we have passed an interval for a suitable learning rate instead of a single value, which further helps to reach global minima. The size of the dataset is enormous. Hence it was impossible to allocate all the images into RAM at a time while training. Therefore, we have divided the training images into batches of size 10-20, each which overcome the CUDA memory errors due to RAM shortage.

From the literature studies, we observed that the FAST AI had not been used much in this field; therefore, our work counts towards this domain's contribution through our project. We are currently working on the ensemble of these deep CNN models for further improvement of evaluation metrics. We are also working to train another state-of-the-art Efficient neural network [10]. We observed that the proposed models occupied only 68 GB of disk space and 12.5 GB of RAM along with GPU from our studies. Thus the memory allocation is not an issue. The training time of the model could be up to 15 hours due to the enormous dataset size. The deployed model could serve as a triage tool and minimize the time between disease diagnosis and treatment with minimal settings.

V. INDIVIDUAL CONTRIBUTION

All four group members collaborated equally in all the tasks since the project was implementation based. All the work, including literature reading, documentation, pre-processing of the dataset, training the models etc. was equally divided among team members.

REFERENCES

- [1] Das DK, Ghosh M, Pal M, Maiti AK, Chakraborty C. 2013. Machine learning approach for automated screening of malaria parasite using light microscopic images. *Micron* 45:97106 DOI 10.1016/j.micron.2012.11.002.
- [2] Rajaraman S, Antani SK, Poostchi M, et al. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ*. 2018;6:e4568. Published 2018 Apr 16. doi:10.7717/peerj.4568
- [3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12)*. Curran Associates Inc., Red Hook, NY, USA, 1097–1105.
- [4] Huang, Gao et al. "Densely Connected Convolutional Networks." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017): 2261-2269.
- [5] He, Kaiming Zhang, Xiangyu Ren, Shaoqing Sun, Jian. (2016). Deep Residual Learning for Image Recognition. 770-778. 10.1109/CVPR.2016.90.
- [6] A. G. Santos, C. O. de Souza, C. Zanchettin, D. Macedo, A. L. I. Oliveira and T. Ludermit, "Reducing SqueezeNet Storage Size with Depthwise Separable Convolutions," 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, 2018, pp. 1-6, doi: 10.1109/IJCNN.2018.8489442.
- [7] Martin, Charles Mahoney, Michael. (2019). Heavy-Tailed Universality Predicts Trends in Test Accuracies for Very Large Pre-Trained Deep Neural Networks.
- [8] Dataset Resource : <https://ceb.nlm.nih.gov/repositories/malaria-datasets/>
- [9] Olga Russakovsky*, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (* = equal contribution) ImageNet Large Scale Visual Recognition Challenge. *IJCV*, 2015
- [10] Tan, Mingxing Le, Quoc. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.
- [11] H. A. Nugroho, S. A. Akbar and E. E. H. Murhandarwati, "Feature extraction and classification for detection malaria parasites in thin blood smear," 2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), Semarang, 2015, pp. 197-201.
- [12] P. A. Pattanaik, T. Swarnkar and D. Sheet, "Object detection technique for malaria parasite in thin blood smear images," 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, 2017, pp. 2120-2123.
- [13] I. Suwalka, A. Sanadhya, A. Mathur and M. S. Chouhan, "Identify malaria parasite using pattern recognition technique," 2012 International Conference on Computing, Communication and Applications, Dindigul, Tamilnadu, 2012, pp. 1-4.
- [14] S. Nayak, S. Kumar and M. Jangid, "Malaria Detection Using Multiple Deep Learning Approaches," 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), Jaipur, India, 2019, pp. 292-297.
- [15] H. A. Mohammed and I. A. M. Abdelrahman, "Detection and classification of Malaria in thin blood slide images," 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, 2017.
- [16] Jing Sun, Xibiao Cai, Fuming Sun and J. Zhang, "Scene image classification method based on Alex-Net model," 2016 3rd International Conference on Informative and Cybernetics for Computational Social Systems (ICSS), Jinzhou, 2016, pp. 363-367.
- [17] Q. Zhao, N. Raoof, S. Lyu, B. Zhang and W. Feng, "RSNet: A Compact Relative Squeezing Net for Image Recognition," 2019 IEEE Visual Communications and Image Processing (VCIP), Sydney, Australia, 2019, pp. 1-4.
- [18] X. Wu and L. Zhao, "Study on Iris Segmentation Algorithm Based on Dense U-Net," in *IEEE Access*, vol. 7, pp. 123959-123968, 2019.
- [19] Gaili Yue, Lei Lu, "Face Recognition Based on Histogram Equalization and Convolution Neural Network", 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC).
- [20] Yamei Zhang, Liting Jing, "Convolution Neural Network Application for Glioma Image Processing", *IEEE Advanced Information Management,Communicates,Electronic and Automation Control Conference (IMCEC)*, May 2018.
- [21] Y. Zhang, J. Wan, A. Yang, C. Ji and J. Chu, "Research on Recommendation Method Based on Multi-Convolution Neural Network," 2019 International Conference on Communications, Information System and Computer Engineering (CISCE), Haikou, China, 2019, pp. 420-424.
- [22] S. Woo and C. L. Lee, "Decision Boundary Formation of Deep Convolution Networks with ReLU," 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress, Athens, 2018, pp. 885-888.
- [23] Xu Xu, Hao Ge, Shenghong Li, "An improvement on recurrent neural network by combining convolution neural network and a simple initialization of the weights", 2016 IEEE International Conference of Online Analysis and Computing Science (ICOACS), Chongqing, China, May 2016.
- [24] Jiawen Liu, Huimei Yuan, Mingyang Li, "Hidden information recognition based on multitask convolution neural network", 2017 6th International Conference on Computer Science and Network Technology (ICCSNT), October 2017.
- [25] L. Su, L. Ma, N. Qin, D. Huang and A. H. Kemp, "Fault Diagnosis of High-Speed Train Bogie by Residual-Squeeze Net," in *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 3856-3863.
- [26] Ling Li, Yuqi Tong, Hangyu Zhang, Dayu Wan, "Memory Saving Method for Enhanced Convolution of Deep Neural Network", 11th International Symposium on Computational Intelligence and Design (ISCID), December 2018.
- [27] D. Lin, G. Xu, Y. Wang, X. Sun and K. Fu, "Dense-Add Net: An Novel Convolutional Neural Network for Remote Sensing Image Inpainting," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 4985-4988.
- [28] A. Bashir, Z. A. Mustafa, I. Abdelhameid and R. Ibrahim, "Detection of malaria parasites using digital image processing," 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, 2017, pp. 1-5.
- [29] P. Samudre, P. Shende and V. Jaiswal, "Optimizing Performance of Convolutional Neural Network Using Computing Technique," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-4.
- [30] A. B. Abdul Qayyum, A. Arefeen and C. Shahnaz, "Convolutional Neural Network (CNN) Based Speech-Emotion Recognition," 2019 IEEE International Conference on Signal Processing, Information, Communication Systems (SPICSCON), Dhaka, Bangladesh, 2019, pp. 122-125.
- [31] H. Ninomiya and A. Sasaki, "A study on generalization ability of 3-layer recurrent neural networks," *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290)*, Honolulu, HI, USA, 2002, pp. 1063-1068.
- [32] S. Batalama, A. Koyiantis, D. Kazakos and P. Papantoni-Kazakos, "Fundamental design and learning concepts in robust recurrent neural networks," [Proceedings 1992] *IJCNN International Joint Conference on Neural Networks*, Baltimore, MD, USA, 1992, pp. 127-132.