Vol. XVII & Issue No. 12 December - 2024
INDUSTRIAL ENGINEERING JOURNAL

STRESS RECOGNITION FROM FACIAL EXPRESSION: PERFORMANCE ANALYSIS THROUGH A COMBINATION OF AI ALGORITHMS

Deepti Kulkarni
Research scholar, Department of Computer Science, Oriental University, Indore, India
Email: deeptikulkarni018@gmail.com

Rashmi Soni
Research Supervisor & Associate Professor CSE department Oriental University ,Indore, India
Email: drrashmicseofficial@gmail.com

Abstract

We as a society have discovered the sins of mental health in recent times and one who has experienced knows how difficult it can be. Stress causes a lot of mental fatigue in most age groups, with even teens suffering from it showing how important it is to be addressed as soon as possible. So, we came up with a monitoring system which looks at the basic symptoms and also recognizes them through facial expressions. With the FER 2013 dataset, we created two broad categories stressed, happy, and neutral. This paper shows how deep learning models and their combinations when supervised sincerely with good data can read symptoms and signs of stress easily. Based on the results we found we propose a solid method with accurate results and a better understanding of our regular stressed life.

Keywords: Stress monitor, Machine learning, Transfer Learning, FER 2013 Dataset.

1. INTRODUCTION

A person's behaviour and its expression tell a lot about their mood, health, social as well as economic conditions and more. This is been observed and proven how the behaviour of a person shows their mental status and working. We here with the help of both facial and behavioural patterns are detecting the obvious person to have stress [1,2,11,12]. Transfer Learning is also one of the great factors to our model combinations when add with pre-built models most of ANN, and CNN is completely transformed into a better model with better feature extraction and utilization of them. This helps our neural layers to learn more and faster with tested and reliable extraction done by the transfer learning models. Transfer Learning brings lots of customization and a lot of flexibility with the reliability of the model in it. [3,13,14,15] With Deep learning models like ANN (Artificial Neural Network), CNN (Convolution Neural Network), etc which are designed like a neural network which learns the characteristic of image with better accuracy and less time compared to Machine Learning basic models like Decision tree, SVM, etc. With ImageNet weights, we can module and customize our pre-trained backbone models to help us more in our dataset for better feature extraction [10,19,20]. This paper makes use of all the customization and concepts for the best results and accuracy with the least time in the most efficient way possible [9,21,22].

2. LITERATURE REVIEW

There are existing papers that suggest recognizing whether the person is stressed or not but most are based on the various hand gesture and basic body language in the compact environment [1] which is not easy to monitor in a work environment or in general while some track the continuous data of a person all time [2] with their camera open and model running all the time which is not viable to work with and also take a lot of time and memory in anyone's laptop and have a lot of loopholes. So, what we need is a one-shot recognition with either camera in real time for once or by image once to monitor it on regular basis, this helps because it will look into previous data and also notice changes in real-time. This can be used in an attendance system or any other kind of device so it can be looked into with existing devices. We are using Transfer learning methods [3] in deep learning for classification as seen in the reference paper while we compare different models and their compatibility with each other like ResNet50 Vs Resnet101 [4], ResNet50 Vs InceptionV3[5], and InceptionV2 vs InceptionV3 [6]. We use them on the dataset we have acquired from FER 2013 dataset [7] and modelled them to our needs in different categories. We used ImageNet weights to perfect our custom-build models [8] as shown in the reference paper. Then compare the results on base of accuracy and time and approach in various ways [9].

3. DATASET

With the FER 2013 dataset, we created three broad categories stressed, happy and neutral. We choose facial features as well as behaviours of the person in the dataset to separate them into these categories. We took Sad, fearful, disgusted and other feature data and handpicked the images with the following features of stress. We took hand behaviour as well as a fad and uneasy facial features to categorize them into stressed data. We took those hand gestures and behaviour into consideration and features because people in general exhibit this behaviour when tired and stressed and this led to fatigue. While facial expressions and facial features were chosen by looking into a different dataset and observing them. For another category, we randomly took 1000 of datasets from the existing FER 2013 dataset [7,18].

Image Sample of Neutral Faces

Fig 1. (Neutral Faces)

Image Sample of Happy Faces

Fig 2. (Happy Faces)

Image Sample of Stressed Faces

Fig 3. (Stressed Faces)

4. METHODOLOGY

With the Dataset good to go, we will start data pre-processing and will resize all the image data for consistency and better symmetry in further feature extraction processes, and will split data into training and testing data to train the models and test them in a similar way for non-biased results.

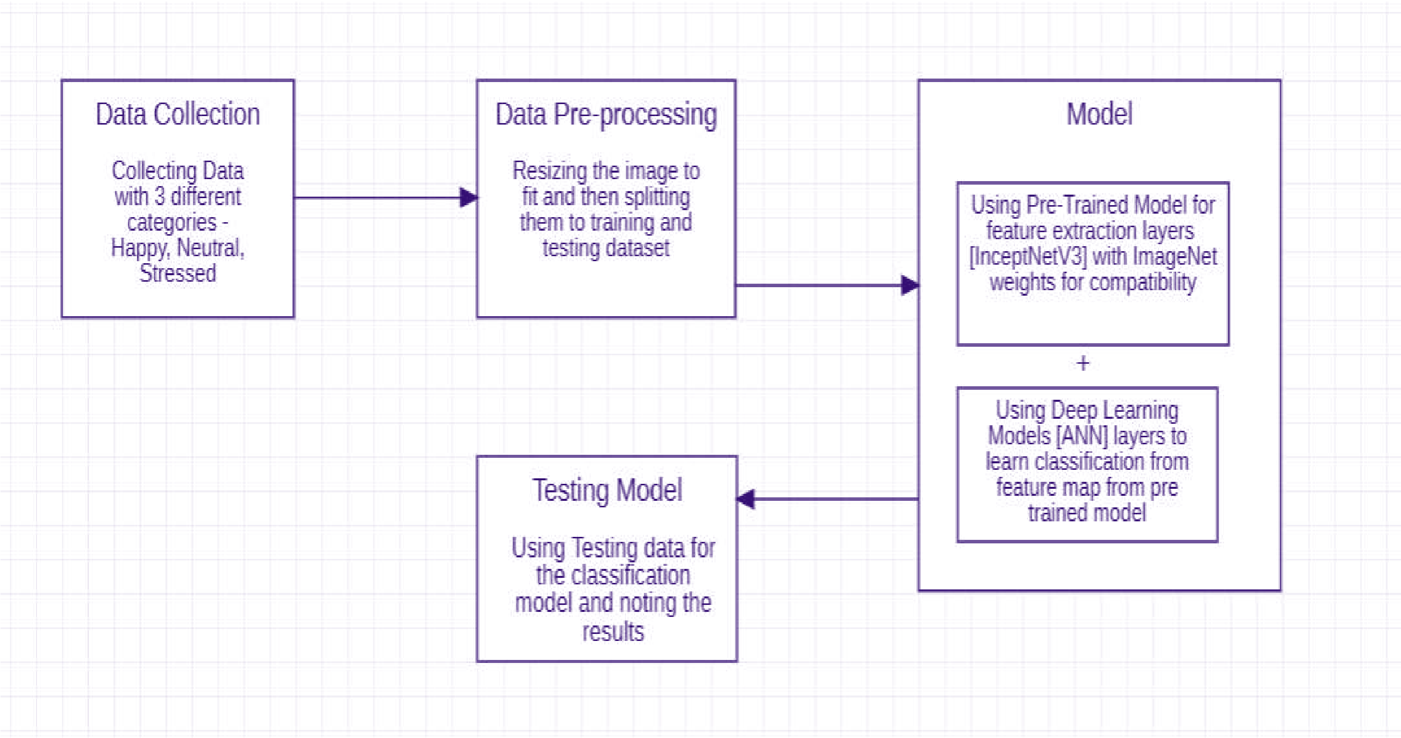
We will use the approach of Transfer Learning to make a highly customized model perfect for our dataset and usage for the best accuracy possible with the limitation of real-time analysis in our hands. Transfer learning will allow us to have backbone pre-trained models like (ResNet50, ResNet101, InceptionV2, InceptionV3, etc.) which will give us the feature extraction of the image which in every classification to help bring out the important characteristics of the image and form a relation or similarity between images of the same classification. Using the pre-trained models for extraction helps us get the proven layers in use without additional build and have a reliable extraction for further process of the model.

We then send those feature maps to our own custom layers of Artificial Neural Networks (ANN), and Convolution Neural Networks (CNN) which will learn the features and classification on the basis of extraction by transfer learning models. The ANN/CNN layers are here responsible only to learn from the extracted feature map from image sets, this will help our model to be light and more easily accessible as the distribution of work is very helpful in real-time prediction with less memory consumption and faster work with high accuracy.

With all the options of customization, we are also able to customize the weights of learning and look at whether it can make a difference in bettering the model. We have used ImageNet weights in our custom models to make them more accurate in some places and used to scope to grow our model in a better way. The ImageNet weights specify that the backbone model was pre-trained on the ImageNet dataset which is decided on the basis of our dataset which is similar to ImageNet data than coco which is also one of the most popular datasets.

With all the customization and different combination in hand, we compare the models on the training accuracy (when the model is tested on the image set from the training dataset for classification accuracy) and testing accuracy (when the model is tested on a separate image set then it is trained on for classification accuracy) to get the best outcome possible.

Fig 4. Architecture



5. RESULTS

Here we can see the results of various custom models with different layers and different feature extraction methods and how significant it can be in accuracy…

Table 1. Model Accuracy

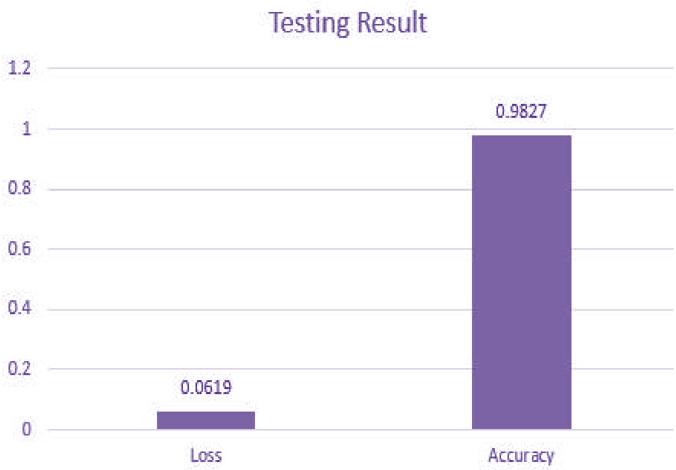
Model	Testing Accuracy	Training Accuracy
ANN + ResNet50	97.87	92.16
ANN + InceptNetV2	80.30	70.37
ANN + InceptNetV3	97.24	98.27
CNN + ResNet50 + ImageNet(weights)	93.00	72.00
CNN + ResNet101 + ImageNet(weights)	89.00	71.00
ANN + InceptNetv3+ ImageNet (weights)	**99.77**	**98.74**

- ANN + InceptNetv3 (training Results)

Fig 5. ANN + InceptNetv3 (training Results)

ANN + InceptNetv3 (testing Results)

Fig 6. ANN + InceptNetv3 (testing Results)

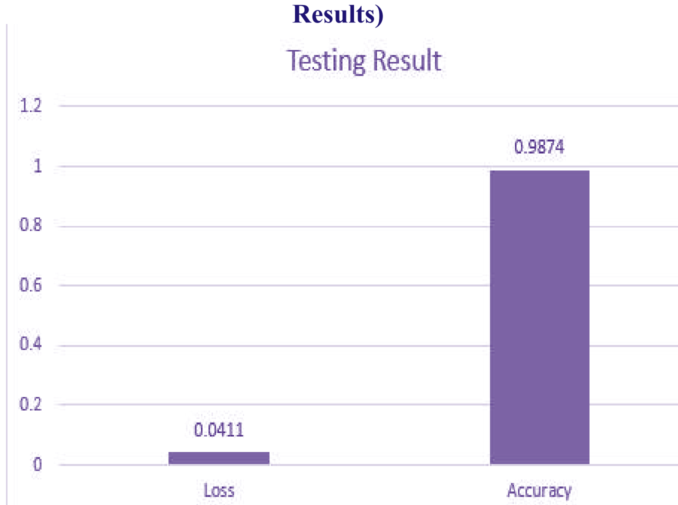


- ANN + InceptNetv3+ ImageNet (weights) (training Results)

Fig 7. ANN + InceptNetv3+ ImageNet (weights) (training Results)

- ANN + InceptNetv3+ ImageNet (weights) (testing Results)

Fig 8. ANN + InceptNetv3+ ImageNet (weights) (testing Results)



- CNN Custom model (training vs testing Results)

Fig 9. CNN Custom model1 (training vs testing Results)

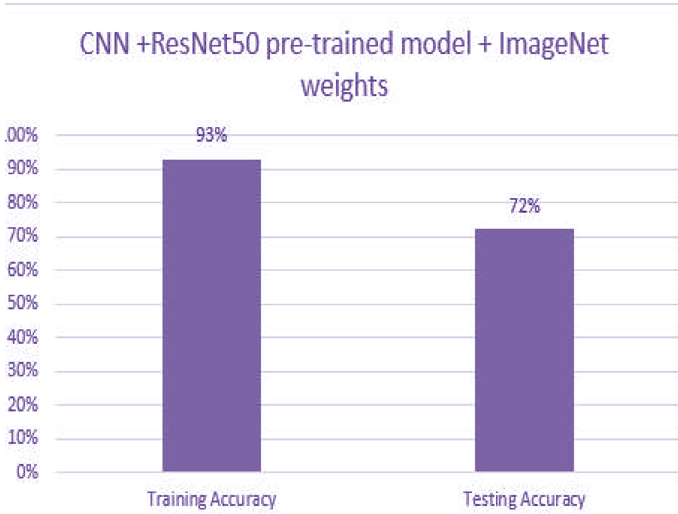
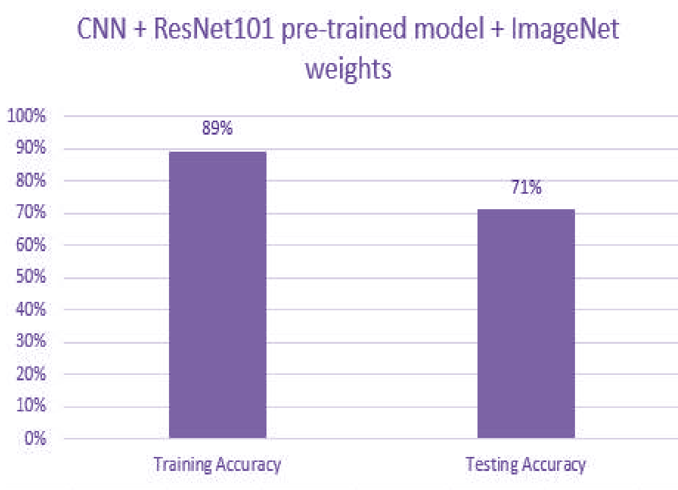


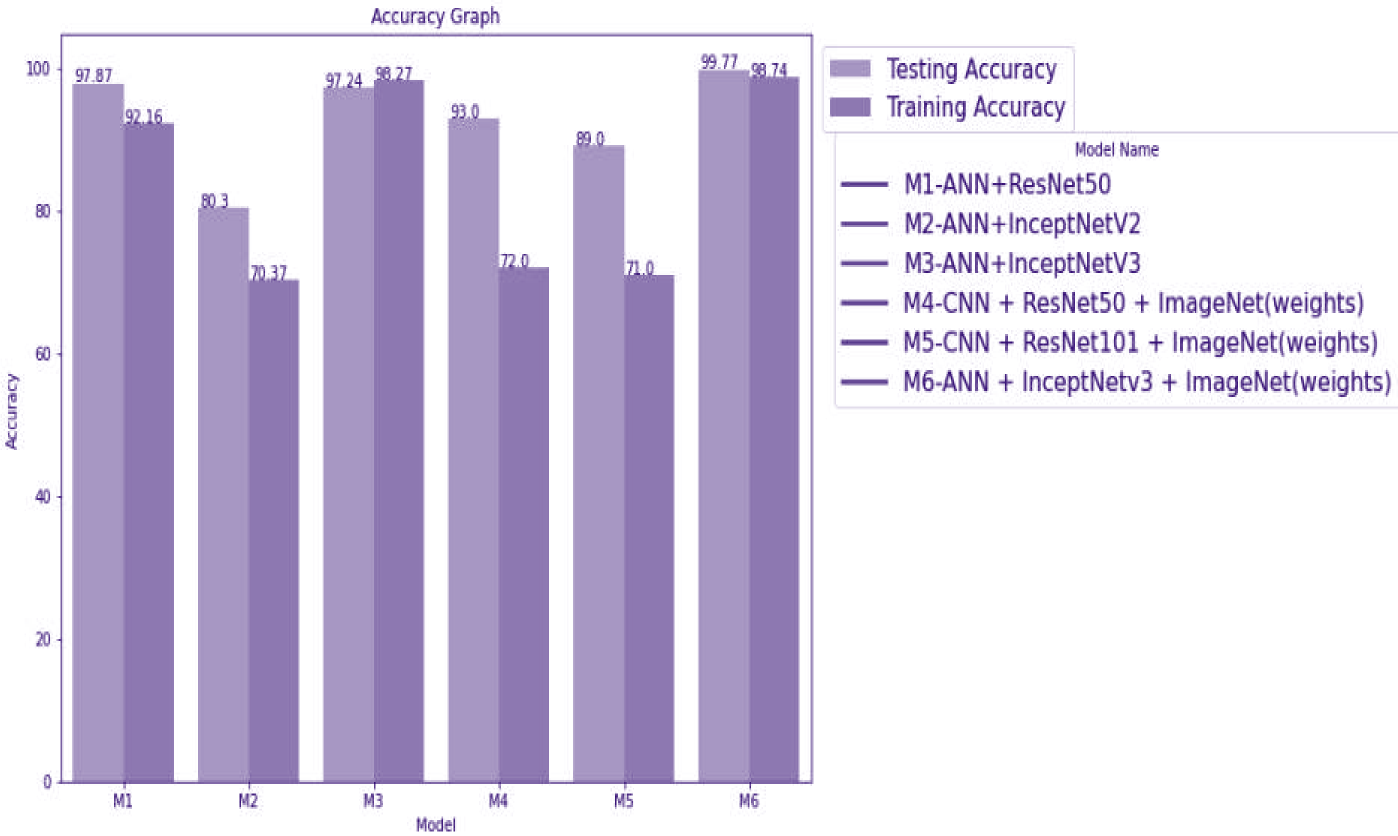
Fig 10. CNN Custom model 2 (training vs testing Results)



6. CONCLUSION

The results show us how a customization of a pre-trained model to different layers can either increase or decrease the accuracy on dataset, which shows us how compatibility of all layers is important and also, to maintain the accuracy in both training and testing dataset is crucial for our analysis of data.

Fig 11. Comparison of Results



We can see InceptNetV3 + ANN alone Is pretty good but when paired with layers of ImageNet it gives even better results and due to transfer learning, we can have results in the ANN custom model (ANN+ InceptNetV3+ImageNet) quicker which shows us that the InceptNetV3 + ANN model had scope to be better and in comparison, to every other model we have created even CNN and ResNet50/ResNet101. The ANN custom model is superior in classification. Hence, we make a case for this custom model to get the best classification for stress recognition in this paper.

7. CONFLICT OF INTEREST/ COMPETING INTEREST

We Confirm that this manuscript is original and has not been published anywhere. Both authors have approved the manuscript and agree with submission to “Industrial Engineering” journal. There is no conflict of interest.

8. FUNDING

No Funding from any source.

9. ACKNOWLEDGEMENT

I would like to acknowledge and give my warmest thanks to Dr. Sunanada Manke who made this work possible. Her guidance and advice carried me through all the stages of writing my research paper.

10. AUTHORS' CONTRIBUTIONS

The first author Deepti Kulkarni and Co-author Rashmi Soni contributed to paper writing and concepts implementation. There is no role of any organization & university in our manuscript writing, implementing results and concepts.

REFERENCES

[1] Jeon, T., Bae, H., Lee, Y., Jang, S., & Lee S. : Stress recognition using face images and facial landmarks. In 2020 International Conference on Electronics, Information, and Communication (ICEIC) 1-3,(2020).

[2] Metaxas, D., Venkataraman, S., & Vogler, C.: Image-based stress recognition using a model-based dynamic face tracking system. In Computational Science-ICCS 2004: 4th International Conference, Kraków, Poland, June 6-9, 2004, Proceedings, Part III 4 813-821 (2004).

[3] Rezende, E., Ruppert, G., Carvalho, T., Ramos, F., & De Geus, P.:Malicious software classification using transfer learning of resnet-50 deep neural network. In 2017 16th

IEEE International Conference on Machine Learning and Applications (ICMLA) 1011-1014 (2017).

[4] Reddy, A. S. B., & Juliet D. S.:Transfer learning with ResNet 50 for malaria cell-image classification. In 2019 International Conference on Communication and Signal Processing (ICCSP) 0945-0949 (2019).

[5] Lin C., Li L., Luo W., Wang K. C., & Guo J. 2019 Transfer learning based traffic sign recognition using inception-v3 model. Periodica Polytechnical Transportation Engineering, 47(3), 242-250 (2019).

[6] Hussain, M., Bird, J. J., & Faria, D. R. : A study on cnn transfer learning for image classification. In Advances in Computational Intelligence Systems: Contributions Presented at the 18th UK Workshop on Computational} Intelligence, September 5-7, 2018, Nottingham, UK 191 202 (2018).

[7] Zahara, L., Musa, P., Wibowo, E. P., Karim, I., & Musa, S. B. :The facial emotion recognition (FER-2013) dataset for prediction system of micro-expressions face using the convolutional neural network (CNN) algorithm based Raspberry Pi. In 2020 Fifth international conference on informatics and computing (ICIC) 1-9 (2013).

[8] Graziani, M., Lompech, T., Müller, H., Depeursinge, A., & Andrearczyk, V. : On the scale invariance in state of the art CNNs trained on ImageNet. Machine Learning and Knowledge Extraction, 3(2), 374-391 (2021).

[9] Hasan, M., Ullah S., Khan M. J., & Khurshid K. : Comparative analysis of SVM, ANN and CNN for classifying vegetation species using hyperspectral thermal infrared data. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 1861 1868 (2019).

[10] Deng, J., Dong ,W., Socher, R., Li, L, J., Li, K., & Fei-Fei, L. : Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition 248-255 (2009).

[11] Prasetio, B. H., Tamura, H., & Tanno, K. : The facial stress recognition based on multi-histogram features and convolutional neural network. In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC) 881 887 (2018) .

[12] Buenaposada, J. M., Munoz, E., & Baumela, L.: Recognising facial expressions in video sequences. Pattern Analysis and Applications, 11, 101-116 (2008).

[13] Rezende, E., Ruppert, G., Carvalho, T., Theophilo, A., Ramos, F., & Geus, P. D. : Malicious software classification using VGG16 deep neural network's bottleneck features. In Information Technology-New Generations: 15th International Conference on Information Technology (pp. 51-59) 2018.

[14] Rajaraman, S., Jaeger, S., & Antani, S. K. : Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. PeerJ, 7, e6977 (2019).

[15] Singh, I., Singh, S. K., Kumar, S., & Aggarwal, K. : Dropout VGG based convolutional neural network for traffic sign categorization. In Congress on Intelligent Systems: Proceedings of CIS 2021, Volume 1 247-261 (2022).

[16] He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., & Li, M. : Bag of tricks for image classification with convolutional neural networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 558 567 (2019).

[17] Bunrit, S., Kerdprasop, N., & Kerdprasop, K. : Improving the representation of cnn based features by autoencoder for a task of construction material image classification. Journal of Advances in Information Technology Vol, 11(4) (2020).

[18] John, A., Abhishek, M. C., Ajayan, A. S., Sanoop, S., & Kumar, V. R. : Real-time facial emotion recognition system with improved preprocessing and feature extraction. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) 1328-1333(2020).

[19] Belilovsky , E., Eickenberg, M., & Oyallon , E. : Greedy layerwise learning can scale to imagenet. In International conference on machine learning 583-593(2019).

[20] Yamada, Y., & Otani, M.: Does robustness on ImageNet transfer to downstream tasks?. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 9215-9224 (2022).

[21] Behera, S. K., Rath, A. K., & Sethy , P. K. : Maturity status classification of papaya fruits based on machine learning and transfer learning approach. Information Processing in Agriculture, 8(2), 244-250 (2021).

[22] Rehman, A., Naz, S., Razzak, M. I., Akram, F., & Imran, M.: A deep learning-based framework for automatic brain tumors classification using transfer learning. Circuits, Systems, and Signal Processing, 39, 757-775 (2020).