IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 2, July 2023

Transfer Learning and Tuning of Deep Pre-trained Architecture for Face Recognition

Shem L. Gonzales

Faculty, CEIT, Surigao del Norte State University, Surigao City, Philippines

slgonzales@ssct.edu.ph

Abstract: *Automatic Image Identification is one of the interests of software developers with the application of machine and deep learning methods. With the incorporation of Transfer Learning and Tuning in pre-trained architecture, a substantial increase in the model's performance is evident. This paper performs face recognition using an image identification and recognition approach. Feature extraction was performed using ResNet50 pre-trained architecture with Support Vector Machine as a classifier. Initial evaluation was made to generate a precision of 62.50%, recall of 65.55%, and f1-score of 63.99%. With this poor performance of ResNet50, the hyperparameters were tuned using transfer learning and tuning. After several times of manual experiments, a significant increase in precision is 93.75%, recall is 94.36%, and f1-score is 94.05%. Based on the remarkable yield of 35.25% for accuracy, 38.79% for recall, and an f1-score of 30.06%, it is advisable to apply the model for image identification and recognition*

Keywords: CNN Architectures, Transfer Learning, Image Identification, ResNet50

I. INTRODUCTION

Deep Learning in image identification and recognition took advantage of traditional methods. Deeper layers generate important information that can be used for solving other related problems [1]. This information is termed as features from images that represent information as to what the image is all about. Transfer Learning used the concept to minimize the hardware specification requirement cost and the time consume for the training experiment[2].

Face Recognition is one of the challenging biometric applications in computer vision. It has evolved into a distinctive technology in a variety of fields, including security, surveillance, and authentication [3].Its performance is accelerated with the application of state-of-the-art Deep Learning using Transfer Learning and Fine Tuning. Utilizing the pre-trained model of Deep CNN architectures involved taking a model on a large dataset and fine-tuning it on a problem with a lesser dataset[4]. These strategies have been seen in various applications of the classification problem. However, achieving an accurate and reliable face recognition model is still a challenge with factors including lighting conditions, expressions, and the quality of the image.

As mentioned, Transfer Learning is one of the approaches to reinforce the traditional method. This learning approach has been seen in different research applying Deep CNN's architectures adapting the task of facial feature extraction and recognition [5]. On the other hand in the Transfer Learning design, fine-tuning entails the configuration of the architecture's hyperparameters to a specific value for recognition purposes. The last layers of the neural network are adjusted to facilitate better the unseen features of faces for better recognition performance [6].

This paper experimented ResNet50 architecture of Deep CNN to create a model for face recognition. Eight classes with 30 images class were used as datasets. Diverse datasets were considered as to the impact of pre-trained models. The initial process was done through feature extraction using Support Vector Machine as classifier. The comparison to performance metrics was evaluated with the application of Transfer Learning and Fine Tuning. With the results of the experiments, this paper seel to contribute to the development of a strong face recognition framework

II. METHODOLOGY

2.1 Face Image Collection

Images were collected from eight classes of local personalities. Images were collected with 30 facial photos for every class representation from photo album collections. The images were converted to jpeg file format using a scanner. These images were cropped into 224x224 pixel dimensions as the basic requirement for the ResNet50 input layer. Fig. 1shows the sample of images used in this study. The distribution of images per class is presented in Table 1.

Fig. 1. Sample of the images used in this experiment

TABLE I: image Distribution across classes

Classes	Number of Images
Person 1	30
Person 2	30
Person 3	30
Person 4	30
Person 5	30
Person 6	30
Person 7	30
Person 8	30

2.2 Image Preprocessing

Due to the limited number of images in the dataset, each image is augmented. Python programming language was used to increase the sample size of the image. Image Augmentation was already proven to help increase the model's performance for the training process [7][8].

2.3 Feature Extraction

An initial experiment was conducted on ResNet50 using an SVM classifier [9]. Table II presents the layer's location where feature extraction was performed. After the process, evaluation metrics were performed using precision, recall, and f1-score [10]. The following equation were used:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, July 2023

$$F1 - score = 2\frac{P - R}{P + R}$$

TABLE II: image Distribution across classes

Architecture	Layer	Name	Type
ResNet50	174	avg_pool	Global Average Pooling

2.4 Transfer Learning and Fine Tuning Configuration

After the execution of feature extraction, the model evaluation was performed. Although Deep CNN architectures were proven to have an outstanding performance, the results from the feature extraction method in ResNet50 with SVM cannot be used as a model. Transfer learning frequently entails taking the first layers' pre-trained weights, which are frequently applied to a wide range of datasets, initializing the final layers at random, and then training them for classification purposes. In machine learning and deep learning, tuning is an optimization solution that can increase the performance of the classification model (Wani& Afzal, 2018; Ipopolito, P.P., 2019; Abu et al., 2022). The hyperparameter of the chosen architecture should be tuned to make it applicable to the face recognition problem. The target is to find the right combination of hyperparameter values such as learning rate, batch size, epoch, and optimizers that will obtain the best observation. Table III shows the tuned configuration of ResNet50 with corresponding values.

TABLE IIII: tuned configuration of resnet50

Hyperparameter	Value
Wieght Learn Factor	20
Bias Learn Rate Factor	20
Mini BAtch Size	5
Epochs	30
Initial Learn Rate	.0001
Validation Frequency	10
Optimizer	Adam

III. RESULTS AND DISCUSSION

3.1 Learning and Loss Curves in Transfer Learning and Fine Tuning

Fig. 1 shows the learning and loss curves after the training process. The process generated a training accuracy of 100% and a validation accuracy of 100%. The curves validated a good fit as it goes to the number of epochs declared in Table III as it stops at 2400 iterations at 23 minutes and 33 seconds. In this case, overfitting did not occur.

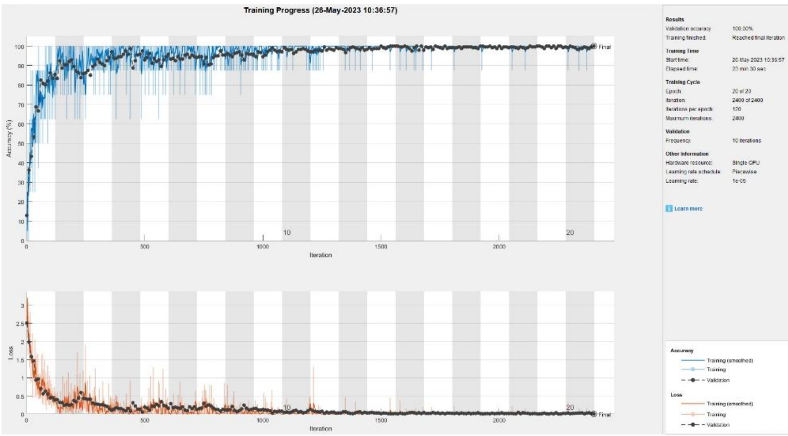

Fig. 2. Learning and Loss Curves

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, July 2023

3.2 Performance Evaluation Comparison

Fig.2 presents the performance of ResNet50 using feature extraction and transfer learning and hyperparameter tuning. A comparison of the methods shows that transfer learning and hyperparameter tuning achieved better performance than feature extraction as to precision, recall, and f1-score. Transfer learning yielded a 35.25% precision, recall of 28.79%, and f1-score of 30.16% which outperformed the features extraction method with the SVM classifier.

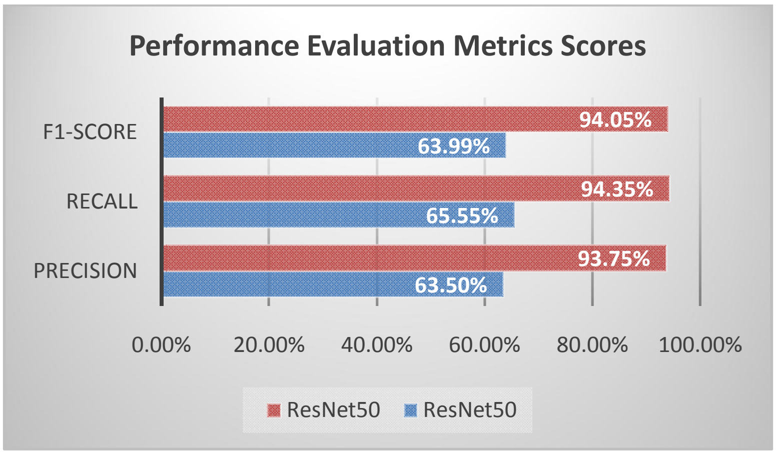

Fig. 3. ResNet50 comparison as to feature extraction and transfer learning and fine-tuning

IV. CONCLUSION

ResNet50 was experimented with for face recognition purposes. Feature extraction with an SVM classifier and transfer learning and fine-tuning were the two methods used. The process was evaluated based on the scores of precision, recall, and f1-score. Feature extraction shows poor performance compared with transfer learning and fine-tuning. The evaluation rating shows that ResNet50 performed well in transfer learning and fine-tuning with a mean precision of 94.05%, mean recall of 94.35% and f1-score of 93.75%. these scores result in a yield of 35.25% for precision, 28.79% for recall, and 30.16% for f1-score. This implies that ResNet50 pre-trained model can be a good architecture for an outstanding face recognition framework.

REFERENCES

[1]. L. Alzubaide, et al., Review of Deep Learning: Concepts, CNN Architectures Challenges, Applications, Future Directions. Journal of Big Data, 2021, vol. 8

[2]. H. Würschinger, M. Mühlbauer, M. Winter, M. Engelbrecht, and N. Hanenkamp, Implementation, and Potentials of a Machine Vision System in a Series Production using Deep Learning and Low-cost Hardware. Procedia CIRP, 2020, vol. 90.

[3]. H. Imaoka, et al., The future of biometrics technology: From face recognition to related applications. APSIPA Transactions on Signal and Information Processing, 2021, vol. 10.

[4]. G. Vrbrancic& V. Podgorelec, Transfer Learning with Adaptive Fine-tuning. Digital Object Identifier, 2020, vol. 8.

[5]. H. Li, et al.,UFaceNet: Research on Multi-Task Face Recognition Algorithm Based on CNN. Algorithms, 2021, vol. 14.

[6]. R. Yamashita, et al., Convolutional Neural Networks: An Overview and Application in Radiology. Insights Imaging, 2018, vol. 9.

[7]. M. Kiran, S. Mondal, & B. Nemade, A Review: Data Pre-processing and Data Augmentation Techniques, Global TransistionsProcessings, 2022, vol. 3.

[8]. K. Alomar, H.I. Aysel, & X. Cai, Data Augmentation in Classification and Segmentation: A Survey and New Strategies, Imaging, 2023, vol. 9.

[9]. W. Salama& M.H. Aly, Deep Learning Design for Benign and Malignant Classification of Skin Lesions: A New Approach. Multimedia Tools and Application, 2021, vol. 80.

[10]. I. Ahmad, et al., Optimizing Pretrained Convolutional Neural Networks for Tomato Leaf Disease Detection. Complexity, 2020, vol. 6.