

Designing Effective Models for COVID-19 Diagnosis through Transfer Learning and Interlayer Visualization

Cuneyt Ozdemir¹

¹ Department of Computer Engineering University of Siirt University, Siirt, Turkey, (e-mail: cozdemir@siirt.edu.tr).

Manuscript received Jun 1, 2023; accepted August 6, 2023.

Abstract— Creating an optimal model for a specific dataset can be challenging and time-consuming. This article presents an innovative approach to model creation by leveraging transfer learning models and employing the interlayer visualization method. The study aims to overcome the complexities of designing a new model specifically for the COVID-19 dataset. Transfer learning models, which are pre-trained models, offer a practical solution due to their adaptability. However, not all layers in these models may be suitable for a given dataset, emphasizing the importance of identifying and removing unnecessary layers for successful model performance.

Experimental studies were conducted using transfer learning models, including Densenet201 and InceptionV3, on the COVID-19 dataset. The interlayer visualization method was utilized to identify irrelevant layers, resulting in the creation of new models. The evaluation metrics demonstrated that the derived models outperformed the original transfer learning models. The Mixed3 model derived from InceptionV3 achieved a higher accuracy of 98.6%, compared to the original model's accuracy of 98.3%. Similarly, the Pool4 model derived from DenseNet201 achieved an accuracy of 98.8%, surpassing the original model's accuracy of 98.19%. Moreover, the sensitivity for detecting images belonging to the Covid-19 class reached 98.9% and 99.7% for the Mixed3 and Pool4 models, respectively.

The new models obtained through the interlayer visualization method offer several advantages, including lightweight design, faster training processes, and improved performance. This study highlights the effectiveness of leveraging transfer learning models with the interlayer visualization method for creating robust models tailored to specific datasets. The proposed approach serves as a valuable solution to the challenges associated with model creation, particularly in the context of COVID-19 diagnosis using medical imaging data.

Index Terms—Covid-19, DenseNet201, InceptionV3, Interlayer visualization, Model pruning, Transfer learning models

I. INTRODUCTION

Transfer Learning is a method in machine learning, utilizes pre-trained models to perform new tasks. When designing a new deep neural network architecture, transfer learning models are commonly employed [1]. This approach allows the adaptation of a pre-trained model's general features to a new task, proving

particularly useful when data scarcity or large datasets hinder training a model from scratch. Transfer learning leverages features learned from one dataset for another, leading to superior results with reduced training data [2].

Transfer learning finds extensive applications in various domains, including agriculture, healthcare, image processing, natural language processing, and speech recognition [3-7]. However, not all transfer learning models are suitable for every dataset, and some models may contain unnecessary layers. Therefore, selecting the most appropriate transfer learning model for a specific dataset and identifying the optimal layers within that model are crucial for achieving successful outcomes.

The size and characteristics of the dataset influence the selection of an appropriate model. Larger datasets benefit from complex and deep models, while smaller datasets are better suited for simpler models. Hence, the choice of transfer learning models should align with the specific characteristics of the dataset.

Despite the advantages of transfer learning, there are certain drawbacks to consider. Firstly, not all layers in a pre-trained model are necessary for a new task, and some layers may introduce redundancy or contribute to overfitting depending on the dataset's characteristics [8]. Therefore, it is essential to remove or modify unnecessary layers to enhance the performance of the pre-trained model. Adapting pre-trained models to match the size and complexity of the dataset is crucial, and having a clear understanding of which layers to modify or remove is essential for optimizing model performance.

In some cases, employing all layers of a pre-trained model can result in longer training times or reduced model efficiency. Thus, pre-trained models should be adaptable to the dataset's size and complexity. Nevertheless, transfer learning models require less time and computational power compared to training a model from scratch for a new task. Moreover, pre-trained models typically capture more general features due to their training on larger datasets. These features can be fine-tuned for the new task, resulting in a more accurate and efficient model.

To address unnecessary or redundant information in transfer learning models, the inter-layer visualization technique, known as model pruning, is employed. Model pruning aims to eliminate unnecessary parameters or layers from a model. It

identifies redundant layers by analyzing the model's layers or filters, eliminating similar or weak activations. Model pruning reduces the model's size, shortens training time, and improves its generalization ability.

Zeiler et al. [9] utilized the AlexNet deep learning model to visualize how each layer learns its features and represents each feature map. This visualization process aimed to comprehend the combination of lower-level features learned by certain layers in higher-level layers, leading to the production of the final output. The study investigated the impact of removing specific layers on the model's performance, revealing a significant drop in performance when the first two layers were removed. This result demonstrated that the initial layers of the model learn lower-level features, while higher-level layers learn higher-level features through their integration.

Urban et al. [10] investigated the effect of removing layers in deep learning models on the model's performance. The study showed that, in certain cases, the number of layers in deep learning models could be reduced, and some models achieved high performance even with a few layers.

Bau et al. [11] introduced an approach to comprehend the features of layers within deep learning models. This approach introduced the concept of "neuron activation" to quantify the relevance of each feature to a specific object or concept.

Li et al. [12] developed a method for visualizing the loss landscape of various deep learning models. This method enables an understanding of the impact of removing or modifying layers on model performance by visually demonstrating the effects of different layers in the model.

This study aims to create new models and enhance their performance by conducting inter-layer visualization on transfer learning models. The inter-layer visualization method can identify layers that carry less information by visualizing the outputs of each layer in the model. When a model is provided with input, inter-layer visualization visualizes the activations generated by each layer of the model. These visualizations assist in determining the importance of each layer and which layers contribute more significantly to the model's performance.

COVID-19 X-ray images were employed to examine the performance of model pruning operations on transfer learning models. COVID-19, a viral disease caused by SARS-CoV-2, was initially identified in Wuhan, China. While the initial cases were reported in Wuhan, the precise origins and transmission of the virus remain subjects of ongoing scientific research and investigation.

COVID-19 primarily affects the respiratory system and belongs to the coronavirus family. It is a novel virus for humans, posing a significant risk, particularly to elderly individuals and those with chronic illnesses. COVID-19 can induce various symptoms, including fever, cough, fatigue, shortness of breath, muscle aches, headaches, and loss of taste or smell, among others. The outbreak of this disease rapidly escalated into a major worldwide health crisis, profoundly impacting healthcare systems and economies of numerous countries. Consequently, COVID-19 continues to be a globally-relevant research topic [13-14].

The health effects of COVID-19 are not limited to the respiratory system. The virus can cause brain inflammation, psychological disorders, cardiovascular diseases, kidney failure, and damage to other organs. PCR or antigen tests are commonly employed for COVID-19 diagnosis. With the COVID-19 pandemic becoming a global health issue, artificial intelligence techniques, such as deep learning, have started to play a crucial role in addressing this problem [15,16].

Ertugrul et al. [17] proposed a machine learning-based automatic diagnosis system that utilizes X-ray images to detect COVID-19-related diseases. The system employed robust texture features, such as Histogram of Oriented Gradients, Law's Tissue Energy Measure, Gabor Wavelet Transform, Gray Level Co-occurrence Matrix, and Local Binary Pattern, to train a random neural network. This approach facilitated a rapid and robust diagnostic process for COVID-19 by extracting identifying indicators from the two-dimensional space of X-ray images from virus patients.

Kaya et al. [18] presented a novel approach for COVID-19 detection using X-ray images. They introduced the Angle Transform (AT) method, which captures the angle information between pixels in the images. The AT method generated eight different images per dataset image, which were then used to train a hybrid deep learning model combining GoogleNet and LSTM. The proposed approach achieved a high classification accuracy of 98.97% and demonstrated success in COVID-19 detection using chest X-ray images.

This study involves the creation of new models through model pruning operations on transfer learning models using the COVID-19 dataset. Experimental studies have been conducted using these newly obtained models, and the performance of transfer learning models and the new models derived from them through model pruning methods has been observed.

II. DATASET AND PREPROCESSING

A. Dataset

A group of researchers from Doha, Qatar and Dhaka University, Bangladesh have shared a dataset that contains images. These images consist of chest X-rays for COVID-19 positive cases, Normal images, and Viral Pneumonia images. Table I provides further details regarding the dataset [19].

TABLE I
Distribution of images according to classes

Image Class	Number
COVID 19	3616
Normal	10192
Viral Pneumonia	1345

The distribution of visual data among different classes exhibits an imbalance, as illustrated in Table I. This imbalance can lead to incorrect learning by the model and subsequently generate inaccurate results. In an effort to mitigate this issue, only 3500 normal images from the dataset were utilized. The distribution of the dataset is depicted in Figure 1, encompassing

a total of 8461 images.

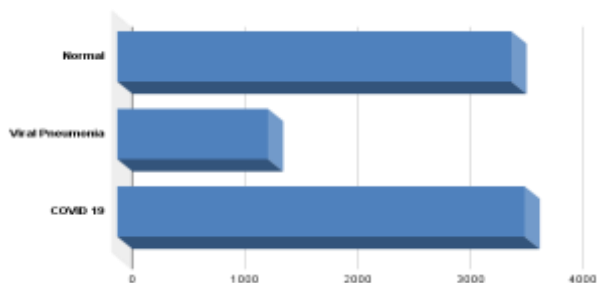


Fig.1. Dataset distribution

Figure 2 displays image examples from three distinct classification clusters within the dataset. These image samples were obtained following the application of data augmentation techniques on the images.

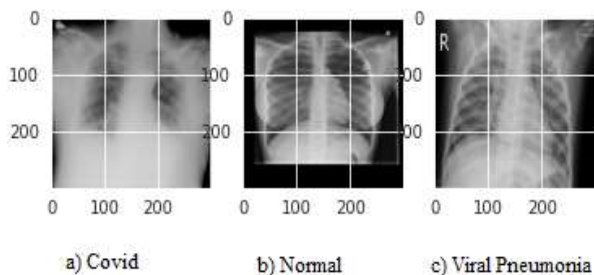


Fig 2. Examples of 3 different image classes

Several studies have been conducted in the literature using this dataset:

Aslan et al. [20] conducted a study where they utilized features extracted from various well-known Convolutional Neural Network (CNN) models, such as AlexNet, Inceptionresnetv2, ResNet18, Inceptionv3, ResNet50, MobileNetv2, Densenet201, and GoogleNet. These extracted features were then used for classification using different machine learning algorithms, including SVM, k-NN, Naive Bayes, and Decision Tree. The DenseNet + SVM method achieved a classification accuracy of 96.29%.

Sohan et al. [21] employed ResNet-18 and VGG16 transfer learning models on the same dataset, achieving an accuracy of 97% in their study.

Sahlol et al. [22] proposed a hybrid classification approach that utilized the Marine Predators algorithm to select a swarm-based feature from the InceptionV3 transfer learning model. In their study, they achieved an accuracy of 98.7% and an F-score of 98.2%.

Abdollahi et al. [23] achieved an accuracy of 97.99% in their study using the VGG16 transfer learning model.

Abdrakhmanov et al. [24] employed few-shot learning techniques to classify images with a small amount of training data and achieved a classification accuracy of 97.7%.

These studies demonstrate the effectiveness of various transfer learning models and approaches in achieving high accuracies on the dataset.

B. Pre Processing

Before being inputted into the model, the images underwent pre-processing steps. During this stage, the image dimensions were resized to 224x224 pixels. To prevent overfitting and enhance the diversity of the dataset for improved model performance, data augmentation techniques were employed. The data augmentation process included applying a 15% rotation, 10% shifting to the right and left, horizontal flipping, and a 20% zoom in and out of the images.

Following the pre-processing and data augmentation steps, the images were split into three distinct sets: training, validation, and testing. Specifically, 15% of the images were allocated for testing, while the remaining 85% were assigned to the training set. Within the training set, an additional 15% of the images were set aside for validation purposes.

III. METHODOLOGY

In order to address the challenges and time constraints associated with building a model from scratch for our specific dataset, we turned to the powerful technique of transfer learning. Transfer learning involves utilizing the knowledge and features learned by a pre-trained model on a different task or dataset and applying it to a new task or dataset, aiming to achieve improved performance.

Transfer learning brings several advantages to the table, as it allows us to leverage the wealth of information captured by pre-trained models and adapt them to our specific task. By employing transfer learning, we aimed to overcome the difficulties inherent in starting from scratch and enhance the performance of our model on our dataset.

In our experimental studies, the primary objective was to identify the models that exhibited the best performance on our dataset. To this end, we employed transfer learning models, and the performance metrics of these models when applied to our dataset can be found in Table II. To optimize the training process, we utilized the Adam optimization function with a fixed learning rate of 0.0003. Furthermore, we implemented the ReduceLRonPlateau and EarlyStopping methods, which dynamically reduce the learning rate and automatically halt the training of the model when the accuracy fails to improve over a certain number of epochs.

The utilization of transfer learning, along with the aforementioned optimization techniques, enabled us to effectively leverage pre-existing knowledge and adapt it to our specific task. This approach not only accelerated the training process but also allowed us to achieve superior results on our dataset. The detailed performance metrics of the transfer learning models applied in our experimental studies are presented in Table II, providing insights into the effectiveness of our approach.

TABLE II
PERFORMANCE OF TRANSFER LEARNING MODELS

	Accuracy	Precision	Recal	F1-Score

Xception	0.9779	0.9784	0.9779	0.9779
InceptionV3	0.9832	0.9832	0.9832	0.9832
ResNet50V2	0.9799	0.9799	0.9799	0.9799
DenseNet201	0.9819	0.9819	0.9819	0.9819

According to Table II, the InceptionV3 and DenseNet201 transfer learning models achieved the best results for the dataset. After identifying these models through initial experimental studies, model pruning was conducted on them. Model pruning involves removing unnecessary layers from transfer learning models that are not essential for the dataset. Various methods can be employed for model pruning. In this study, the inter-layer visualization or data visualization method was utilized. The inter-layer visualization method generates visual representations of the output produced by each layer involved in processing the input image of an image

classification model. By analyzing the output of each layer, we gain insights into the features learned by the model at each stage and how these features are propagated to subsequent layers. For example, it helps us understand that the initial layers generally learn object contours or colors, while deeper layers focus on more specific object features. This information assists in identifying the layers where learning occurs and those where it does not.

In the inter-layer visualization method, the appearance of an image in all layers of the DenseNet201 and InceptionV3 models trained on the dataset was examined. Through this examination, the layers in which learning occurred were determined in these two models, and the layers where learning did not occur were removed from the models. Figure 3 illustrates the image appearance at different layers of the InceptionV3 model for the dataset. As depicted in Figure 3, learning did not occur after the mixed3 layer.

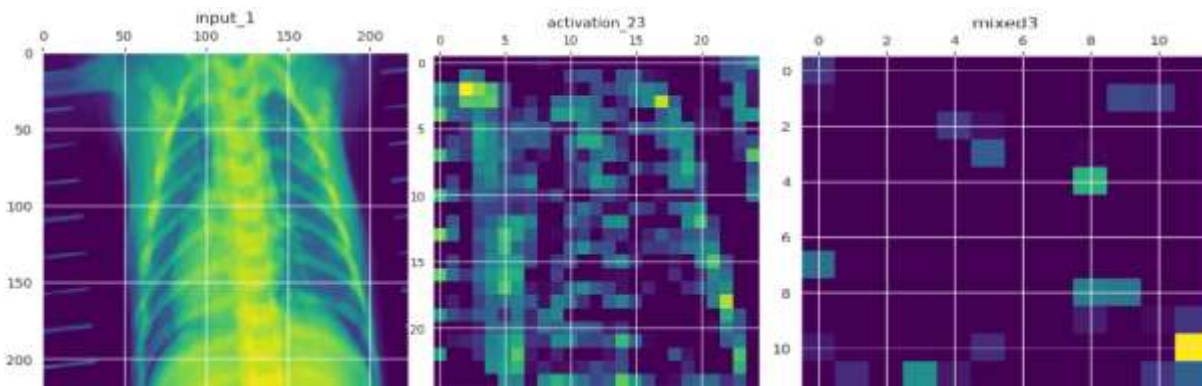


Fig 3. Examples of 3 different image classes

As a result of the process, it was observed that learning occurred up to the mixed3 layer in the InceptionV3 model and up to the pool4 layer in the DenseNet201 model. Based on this observation, new and modified models were created from the transfer learning models.

In the Mixed3 model, only the layers up to the mixed3 layer of the InceptionV3 model were utilized, and the remaining layers were discarded. Similarly, in the Pool4 model, only the layers up to the pool4 layer from the DenseNet201 model were retained, and the rest were removed to obtain a new model.

Figure 4 illustrates the structure of the new Convolutional Neural Network (CNN) model obtained through inter-layer image visualization from the InceptionV3 transfer learning model. The mixed3 model was used for feature extraction in the new model, which were subsequently fed into an artificial neural network. The neural network consists of two layers with 256 and 128 neurons, respectively. The output layer of the artificial neural network consists of three classes, and the Softmax activation function was employed.

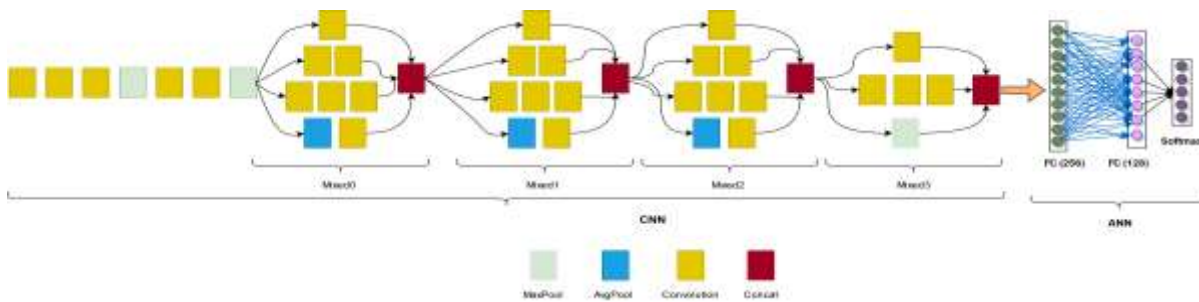


Fig 4. Mixed3 CNN model

After the new models were obtained, experimental studies were performed. The results of the experimental studies with Mixed3 and Pool4 models are shown in Table III. The train and validation, loss and accuracy graphs of the Mixed3 model are

shown in Figure 5. The confusion matrix of the Mixed3 model is shown in Figure 6 and the classification report is shown in Table IV.

TABLE III
ACCURACY RESULTS OF NEW MODELS

	Train Acc	Val Acc	Test Acc
Mixed3	0.998	0.979	0.986
Pool4	0.999	0.988	0.988

TABLE IV
Mixed3 classification report

	Precision	Recall	F1-score
Normal	0.98	0.986	0.983
Viral Pneumonia	0.99	0.985	0.988
Covid	0.989	0.986	0.989

accuracy			0.986
macro avg	0.987	0.986	0.986
weighted avg	0.986	0.986	0.986

TABLE V
Pool4 classification report

	Precision	Recall	F1-score
Normal	0.975	0.997	0.986
Viral Pneumonia	1	0.963	0.981
Covid	0.997	0.989	0.993

accuracy			0.988
macro avg	0.991	0.983	0.987
weighted avg	0.988	0.988	0.988

The evaluation of classification performance was conducted using various metrics, including accuracy, precision, recall, and F1 score for each class. Additionally, the classification report provides average precision, recall, and F1 score, offering an overall assessment of the model's performance across all classes. Table IV demonstrates that the Mixed3 model derived from the InceptionV3 model achieved superior results. While the InceptionV3 model attained an accuracy rate of 98.3% on the same dataset, the Mixed3 model achieved a success rate of 98.6%. Notably, images belonging to the Covid-19 class were detected with a sensitivity of 98.9%. The classification report for the Pool4 model derived from the DenseNet201 model is displayed in Table V. The DenseNet201 model achieved an accuracy rate of 98.19%, while the Pool4 model achieved 98.8% accuracy. Remarkably, images belonging to the Covid-19 class were detected with 99.7% accuracy.

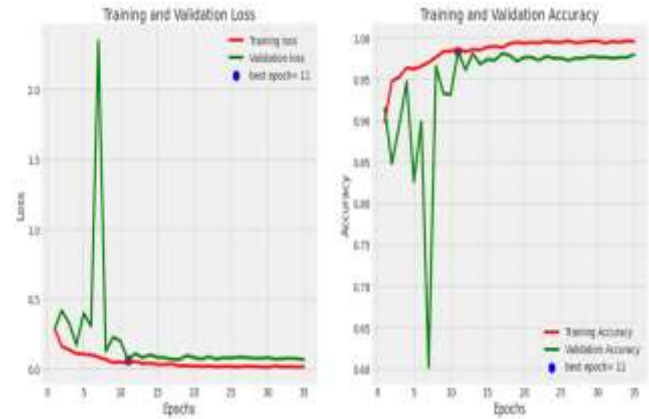


Fig 5. Loss and accuracy graph for Mixed3 model

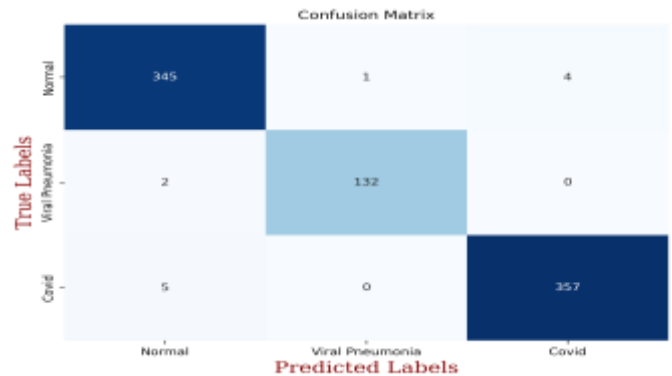


Fig 6. Confusion matrix of the Mixed3 model

Figure 7 illustrates the train and validation loss and accuracy graph of the Pool4 model, providing insights into the model's performance during the training process. Furthermore, Figure 8 presents the confusion matrix for the Pool4 model, allowing for a visualization of the model's performance in terms of correct and incorrect predictions for each class.

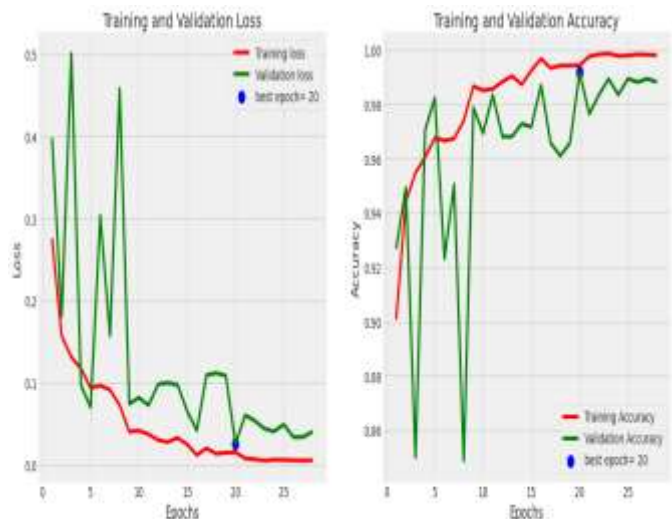


Fig 7. Loss and accuracy graph for Pool4 model

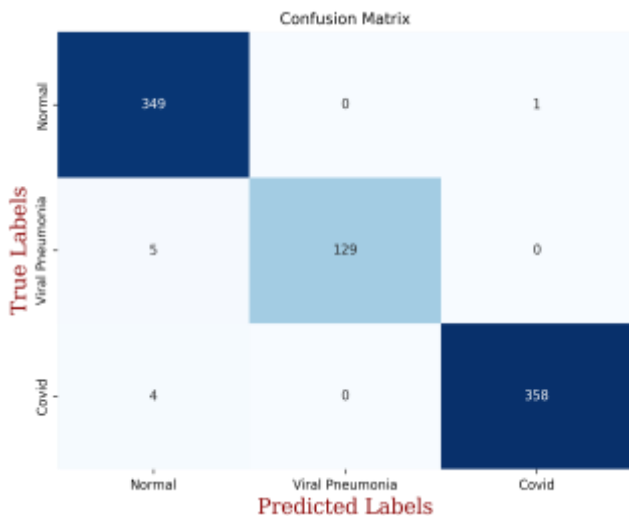


Fig 8. Confusion matrix of the Pool4 model

The experimental studies demonstrated that the new models derived from transfer learning models yielded more successful results. The InceptionV3 model achieved an accuracy of 98.32%, while the Mixed3 model achieved an even higher accuracy of 98.6%. Similarly, the DenseNet201 model achieved an accuracy of 98.19%, while its derived Pool4 model achieved a higher success rate of 98.8%. These results indicate that lightweight and fast models obtained through the interlayer image visualization method can yield more successful outcomes.

IV. CONCLUSIONS

This study proposes an approach that utilizes the interlayer visualization method from model pruning techniques to derive new models from transfer learning models, instead of designing a new model specifically for the COVID-19 dataset. The aim is to address the challenges associated with creating a new model, which is often time-consuming and laborious. By leveraging transfer learning models, the focus is on adapting them to the dataset and creating a new model. It is crucial to identify the layers in transfer learning models that may not extract relevant features from the dataset, as this significantly impacts the model's success and complexity.

To evaluate the performance, transfer learning models were employed on the COVID-19 dataset, rather than building a model from scratch. In these experimental studies, the Densenet201 and InceptionV3 transfer learning models achieved the highest scores. The interlayer visualization method was utilized to identify unnecessary layers in the transfer learning models. The experimental studies conducted with the derived new models demonstrated more successful results.

The new models created using the interlayer visualization method are lightweight, fast, and have fewer parameters. They facilitate faster training processes and yield improved performance in terms of model accuracy. The InceptionV3 model achieved an accuracy rate of 98.3%, while the Mixed3 model derived from it achieved a higher accuracy of 98.6%. Moreover, images belonging to the Covid-19 class were

detected with a sensitivity of 98.9%. Similarly, the DenseNet201 model achieved an accuracy rate of 98.19%, and the Pool4 model derived using the interlayer visualization method achieved an even higher accuracy of 98.8%. Images belonging to the Covid-19 class were detected with a sensitivity of 99.7%.

This study demonstrates that the interlayer visualization method can effectively design the most suitable CNN model for a given dataset by creating new models from transfer learning models. Leveraging the interlayer visualization method provides an effective approach to overcome the challenges associated with creating a new model.

REFERENCES

- [1] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, 'Transfer learning for medical image classification: a literature review', *BMC Med. Imaging*, vol. 22, no. 1, p. 69, Apr. 2022.
- [2] S. Atasever, N. Azginoglu, D. S. Terzi, and R. Terzi, 'A comprehensive survey of deep learning research on medical image analysis with focus on transfer learning', *Clin. Imaging*, vol. 94, pp. 18–41, Feb. 2023.
- [3] I. Atas, C. Ozdemir, M. Atas, and Y. Dogan, 'Forensic dental age estimation using modified deep learning neural network', *arXiv [eess.IV]*, 21-Aug-2022.
- [4] Y. Dogan and H. Y. Keles, 'Iterative facial image inpainting based on an encoder-generator architecture', *Neural Comput. Appl.*, vol. 34, no. 12, pp. 10001–10021, Jun. 2022.
- [5] M. Ataş, C. Özdemir, İ. Ataş, B. Ak, and E. Özeroğlu, 'Biometric identification using panoramic dental radiographic images with few-shot learning', *TURK. J. OF ELECTR. ENG. COMPUT. SCI.*, vol. 30, no. 3, pp. 1115–1126, Jan. 2022.
- [6] Y. Dogan and H. Yalim Keles, 'Stability and diversity in generative adversarial networks', in *2019 27th Signal Processing and Communications Applications Conference (SIU)*, Sivas, Turkey, 2019.
- [7] C. Ozdemir, M. A. Gedik, and Y. Kaya, 'Age estimation from left-hand radiographs with deep learning methods', *Trait. Du Signal*, vol. 38, no. 6, pp. 1565–1574, Dec. 2021.
- [8] M. Iman, K. Rasheed, and H. R. Arabnia, 'A review of Deep Transfer Learning and recent advancements', *arXiv [cs.LG]*, 18-Jan-2022.
- [9] M. D. Zeiler and R. Fergus, 'Visualizing and understanding convolutional networks', in *Computer Vision – ECCV 2014*, Cham: Springer International Publishing, 2014, pp. 818–833.
- [10] G. Urban et al., 'Do deep convolutional nets really need to be deep and convolutional?', *arXiv [stat.ML]*, 17-Mar-2016.
- [11] D. Bau, B. Zhou, A. Khosla, A. Oliva, and A. Torralba, 'Network dissection: Quantifying interpretability of deep visual representations', in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017.
- [12] H. Li, Z. Xu, G. Taylor, C. Studer, and T. Goldstein, 'Visualizing the loss landscape of neural nets', *arXiv [cs.LG]*, 28-Dec-2017.
- [13] H. Harapan et al., 'Coronavirus disease 2019 (COVID-19): A literature review', *J. Infect. Public Health*, vol. 13, no. 5, pp. 667–673, May 2020.
- [14] J. Elliott, M. Whitaker, B. Bodinier, O. Eales, S. Riley, H. Ward, ... & P. Elliott, 'Predictive symptoms for COVID-19 in the community: REACT-1 study of over 1 million people', *PLoS medicine*, 18(9), e1003777., 2021.
- [15] U. Jain, 'Effect of COVID-19 on the Organs', *Cureus*, 12(8), 2020
- [16] D. L. Weiner, V. Balasubramaniam, S. I. Shah, J. R. Javier, and Pediatric Policy Council, 'COVID-19 impact on research, lessons learned from COVID-19 research, implications for pediatric research', *Pediatr. Res.*, vol. 88, no. 2, pp. 148–150, Aug. 2020.
- [17] Ertuğrul, Ö. F., Emrullah, A. C. A. R., Öztekin, A., & Aldemir, E. (2021). Detection of Covid-19 from X-ray images via ensemble of features extraction methods employing randomized neural networks. *European Journal of Technique (EJT)*, 11(2), 248-254.
- [18] Kaya, Y., Yiner, Z., Kaya, M., & Kuncan, F. (2022). A new approach to COVID-19 detection from X-ray images using angle transformation with GoogleNet and LSTM. *Measurement Science and Technology*, 33(12), 124011
- [19] Kaggle. COVID-19 Radiography Database. <https://www.kaggle.com/datasets/tawsiurrahman/covid19-radiography-database>. Accessed 18 January 2023

- [20] M. F. Aslan, K. Sabanci, A. Durdu, and M. F. Unlarsen, 'COVID-19 diagnosis using state-of-the-art CNN architecture features and Bayesian Optimization', *Comput. Biol. Med.*, vol. 142, no. 105244, p. 105244, Mar. 2022.
- [21] M. F. Sohan, A. Basalamah, and M. Solaiman, 'COVID-19 detection using machine learning: a large scale assessment of x-ray and CT image datasets', *J. Electron. Imaging*, vol. 31, no. 04, Mar. 2022.
- [22] A. T. Sahlol, D. Yousri, A. A. Ewees, M. A. A. Al-Qaness, R. Damasevicius, and M. A. Elaziz, 'COVID-19 image classification using deep features and fractional-order marine predators algorithm', *Sci. Rep.*, vol. 10, no. 1, p. 15364, Sep. 2020.
- [23] J. Abdollahi and L. Mahmoudi, 'An artificial intelligence system for detecting the types of the epidemic from X-rays : Artificial intelligence system for detecting the types of the epidemic from X-rays', in *2022 27th International Computer Conference, Computer Society of Iran (CSICC)*, Tehran, Iran, Islamic Republic of, 2022.
- [24] R. Abdrakhmanov, M. Altynbekov, A. Abu, A. Shomanov, D. Viderman, and M.-H. Lee, 'Few-shot learning approach for COVID-19 detection from X-ray images', in *2021 16th International Conference on Electronics Computer and Computation (ICECCO)*, Kaskelen, Kazakhstan, 2021.