Implementation Transfer Learning on Convolutional Neural Network for Tuberculosis Classification

Adya Zizwan Putra1*, Reynaldi Prayugo2, Rizki Mudrika Alfanda Siregar3, Rizky Syabani4, Allwin M Simarmata5

1,2,3,4,5 Universitas Prima Indonesia

1) adyazizwanputra@unprimdn.ac.id, 2) reynaldi.prayugo1711@gmail.com, 3) alfandarizky2@gmail.com, 4) syahbanirisky@gmail.com, 5) allwinmsimarmata@unprimdn.ac.id

Submitted : Jun 4, 2024 | Accepted : Jun 23, 2024 | Published : Jul 1, 2024

Abstract: Tuberculosis (TB) is an infectious disease that can have serious effects on the lungs and is among the top 10 causes of death worldwide. This disease is caused by the transmission of Mycobacterium tuberculosis bacteria through the air when coughing or sneezing. Without treatment, pulmonary tuberculosis can result in permanent lung damage and can be life-threatening. Accurate and early diagnosis is crucial for effective treatment and control of the disease. The challenge lies in the accurate classification of tuberculosis from lung images, which is essential for timely diagnosis and treatment. Traditional diagnostic methods can be time-consuming and sometimes lack precision. To address this issue, this research aims to achieve high accuracy in classifying tuberculosis using the Convolutional Neural Network (CNN) algorithm through transfer learning methods. By utilizing visual images of tuberculosis-affected and normal lungs, we propose a solution that leverages advanced deep learning techniques to enhance diagnostic accuracy. This approach not only expedites the diagnostic process but also improves the reliability of tuberculosis detection, ultimately contributing to better patient outcomes and more effective disease management.

The dataset applied consists of two labels: tuberculosis and normal. This dataset contains 4200 lung images of individuals with tuberculosis and normal lungs. By applying the transfer learning method, Transfer learning is a machine learning method where a pre-trained model is used as the starting point for a new, related task. it was found that the ResNet50 model achieved the highest accuracy at 99%, followed by InceptionV3 at 97%, and lastly, DenseNet121 at 91%.

Keywords: Tuberculosis; Classification; CNN; Transfer Learning; Deep Learning;

INTRODUCTION

An infectious disease that primarily affects the lungs and is among the top 10 causes of death is tuberculosis. This disease is caused by the airborne transmission of bacteria when coughing or sneezing (Tuberkulosis (TBC), Kenali Gejala, Penyebab Dan Cara Penularan, 2024). If left untreated, pulmonary tuberculosis can cause permanent lung damage, potentially life-threatening (Kenali Penyebab Dan GejalaTBC Paru Yang Menular, 2021). In Indonesia, the disease ranks third after India and China, with the number of cases reaching 824,000 and 93,000 deaths, equivalent to 11 deaths per hour (Sehat Negeriku Sehatlah Bangasaku, 2022). A pulmonary tuberculosis patient with a positive smear test has
the potential to transmit the disease to 10-15 people around them (Kristini & Hamidah, 2020). In 2025, the national target for tuberculosis control includes a morbidity rate of around 50% and a mortality rate of around 70% (Zaidi et al., 2022). In connection with the global strategy to combat TB (Tuberculosis) and the Indonesian government's commitment, including the National Medium-Term Development Plan (RJPMN) 2020-2024, the National Tuberculosis Prevention Strategy of Indonesia 2020-2024 has been developed. This document outlines strategies, interventions, activities, and ambitious targets to reduce TB cases as quickly as possible (Germas, 2021). TB infections and deaths are increasing due to transmission, lack of early diagnosis, and a shortage of radiologists in developing areas where TB (Tuberculosis) is common (Oloko-Oba & Viriri, 2020). X-ray imaging is required for an accurate diagnosis, although it is not always specific to TB (Sathithatanacheewin et al., 2020). TB (Tuberculosis) manifestations often resemble those of other lung diseases, making X-ray image interpretation difficult (Melni, 2024). Certainly, diagnosing patients becomes increasingly challenging for medical professionals. Recent studies highlight deep learning in transfer learning as an effective method for classifying lung X-ray images, aiming to diagnose TB (Tuberculosis) more quickly and accurately.

In previous research conducted by (Achmad et al., 2023) classification of tuberculosis disease was performed using a convolutional neural network model, which was trained for 75 epochs. The training and validation data achieved accuracies of 96.40% and 80.58%, respectively. The model was then tested on the test data, resulting in an accuracy of 88%. The model correctly classified 38 positive TB images, with 5 errors. For negative TB images, it correctly classified 37 cases, with 6 errors. In the research by (Anggiratih et al., 2021) using transfer learning models EfficientNet B3 and MobileNet V3 to classify rice plant diseases, they achieved an accuracy of 99% with a training loss of 0.012 for EfficientNet B3, whereas MobileNet V3 reached an accuracy of 57% with a loss of 0.007. In the research conducted by (Aras et al., 2022) to classify batik patterns, we adapted transfer learning using EfficientNet-B2 on a CNN algorithm, utilizing training data, validation data, and test data. The training data consists of 169 images, the validation data has 22 images, and the test data has 22 images. Testing on the data resulted in an accuracy of 72%. With additional fine-tuning and the implementation of ColorJitter and Contrast techniques, the accuracy increased to 90%. The same approach was also undertaken in the research conducted by (Harahap, Em Manuel Laia, et al., 2022) to detect COVID and non-COVID diseases using feature extraction from transfer learning models. Several models were utilized in the study, including VGG19, InceptionNetV2, MobileNetV2, ResNet50V2, ResNet101V2, and ResNet152V2. These models were tested using a COVID dataset consisting of 352 training data, 100 test data, and 88 validation data. Based on the experimental results conducted on Google Colab, ResNet50V2 exhibited high accuracy compared to other models, achieving around 95% (95%) with precision of 0.96, recall of 0.973, F1-score of 0.966, and support of 74. Furthermore, in the research conducted by (Fitroh & 'Uyun, 2022) skin cancer classification in dermoscopic images was performed using transfer learning on a CNN algorithm. The study utilized a skin cancer dataset consisting of malignant and benign cases, divided into 3 subsets: 160 images for validation, 200 images for testing, and a total of 640 training images for malignant and benign cases, summing up to 2000 images. This research employed transfer learning models VGG-16 and ResNet-50. The results showed that VGG-16 achieved an accuracy of 91% with a loss of 0.2712, while ResNet-50 achieved an accuracy of 94% with a loss of 0.2198. ResNet-50 was optimized using RMSprop optimizer, with a batch size of 64, learning rate of 0.0001, and trained for 100 epochs. This approach demonstrated promising performance and could aid dermatology specialists in early skin cancer detection. Next research by (Naufal & Kusuma, 2021) used transfer learning models Xception and MobileNetV2 to classify face mask images. In this study, Xception achieved an accuracy of 0.988 with a computation time of 18,274 seconds, while MobileNetV2 reached an accuracy of 0.981 with a computation time of 40,181 seconds.

Transfer Learning is a recent technique to expedite training in Convolutional Neural Networks (CNN) and enhance performance in classification (S et al., 2020). Transfer Learning is also a machine learning strategy where knowledge from one task is transferred to another task. The process begins by selecting a model trained on general tasks such as image recognition or natural language processing. This model is then adapted to a specific task by modifying some of its final layers or adding new layers (Yan et al., 2024). During adaptation, the model is adjusted to the new patterns from the task data. Fine-tuning is performed with a lower learning rate than the initial training. Afterward, the model is evaluated...
on validation or test data, and additional adjustments can be made to improve performance (Iman et al., 2023). With transfer learning, models can be enhanced by leveraging existing knowledge, saving time and resources needed to train from scratch. This makes it a useful tool for developing and adapting models for various tasks in machine learning.

The solution offered by Transfer Learning is the enhancement of models by leveraging existing knowledge, thus saving significant time and resources. This process not only accelerates the development of models but also improves their performance by building on previously acquired insights. This makes it a useful tool for developing and adapting models for various tasks in machine learning, providing a more efficient pathway to achieving high accuracy and reliability in new applications.

**LITERATURE REVIEW**

Previous studies have demonstrated that Transfer Learning CNNs have proven to provide optimal accuracy in image classification. Research conducted by (Ali et al., 2021) utilized a Convolutional Neural Network (CNN) to identify sounds, employing 16 convolution layers, 3 fully connected layers, and 5 max-pooling layers connected to the visual geometry group of the VGG19 architecture, achieving 95% accuracy on training data and 98% on test data, capable of recognizing 20 letters. Research conducted by (Rochmawanti et al., 2021) employed five pre-trained models provided by Keras: ResNet50, DenseNet121, MobileNet, Xception, InceptionV3, and InceptionResNetV2. This study achieved the highest accuracy of 91.57% in classifying TB disease using DenseNet121. Research conducted by (Saxena et al., 2021) conducted research on tumor classification in MRI images using the VGG16, InceptionV3, and ResNet-50 architectures, achieving accuracies of 90% for VGG-16, 55% for InceptionV3, and 95% for ResNet-50. Research conducted by (Esdras Chaves, Caroline B. Gonçalves, Marcelo K. Albertini, Soojeong Lee, Kwanggil Jeon, 2020) utilized a Convolutional Neural Network (CNN) with five architectures: AlexNet, GoogLeNet, ResNet-18, VGG-16, and VGG-19, to classify breast cancer using 440 infrared images from 88 patients, achieving approximately 90% accuracy with ResNet-18, which outperformed other architectures. Research conducted by (Viera Valencia & Garcia Giraldo, 2019) research comparing Transfer Learning (TL) classification on CNN for brain tumor using MRI. Transfer learning models utilized were VGG16 and DenseNet169 with a total of 3,264 brain Magnetic Resonance Imaging (MRI) image data. DenseNet169 achieved an accuracy of 97%, while VGG16 achieved an accuracy of 75%. Research conducted by (Harahap, Anjelli, et al., 2022) research utilizing transfer learning models in classifying Diabetic Ulcers in Diabetes Mellitus patients. The transfer learning models used were VGG19, MobileNetV2, InceptionResNetV2, ResNet50V2, ResNet101V2, and ResNet152V2. Among the tested transfer learning models, the highest result was obtained from the ResNet152V2 model, which achieved the highest accuracy of 99%.

**METHOD**

The study category falls under experimental research. In the conducted experiment, the research objective is to achieve high accuracy in classifying tuberculosis using the Convolutional Neural Network (CNN) algorithm through transfer learning methods in classifying images of tuberculosis-infected lungs and normal lungs in individuals.

In the context of the research, image processing procedures using Convolutional Neural Network (CNN) are executed through a series of steps. Firstly, chest X-ray images from the Kaggle dataset are categorized into tuberculosis and normal folders, merged, and then split into 3,360 training and 840 testing images. Classification assigns a label of 1 for tuberculosis and 0 for normal.

The test image data undergoes preprocessing steps to ensure optimal input quality for the classification model:

1. The images are rotated to correct any misalignment, ensuring that all lung images are oriented in a standard way which aids in consistent feature extraction.
2. Contrast enhancement techniques, such as histogram equalization or adaptive contrast enhancement, are applied to improve the visibility of features in the X-ray images. This step makes the details in the images more pronounced and easier for the model to learn from.
3. Noise removal is performed using filtering techniques like Gaussian blur or median filtering to eliminate unwanted random variations in the images. This helps in smoothing the images and

*advayizwanputra@unprimdn.ac.id*

This is a Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
reducing the impact of noise on the model's performance, leading to more accurate classification results.

Following this, the images are inputted into the process to classify individuals with tuberculosis using transfer learning models: DenseNet121, InceptionV3, and ResNet-50. The use of these three different models is motivated by their unique architectures and strengths in image classification tasks. DenseNet121 is known for its dense connections between layers, which help in improving the flow of information and gradients throughout the network, leading to better performance on complex image data. InceptionV3, with its inception modules, efficiently captures multi-scale features by using convolutional filters of different sizes, making it adept at recognizing diverse patterns in the images. ResNet-50, on the other hand, leverages residual connections that allow the model to train deeper networks without suffering from the vanishing gradient problem. The connected layers act as classifiers, extracting features and evaluating probabilities on objects in the test image data to determine the final performance of the model. The evaluation of the model's performance is based on accuracy, precision, recall, and F1 score in classifying tuberculosis and normal lungs.

The study utilized Google Colab for programming in Python, while the image data came from an open-source dataset on Kaggle accessible through the provided link. The dataset consists of 700 tuberculosis images and 3500 normal images. An example visualization of the images in the dataset used in the study can be seen in Fig 1. Link for dataset https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset.

![Fig. 1 Examples of tuberculosis and normal data in the dataset.](https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset)

The research applies transfer learning on Convolutional Neural Networks, namely DenseNet121, InceptionV3, and ResNet-50, to classify tuberculosis and normal lung images.

![Fig. 2 The framework of the DenseNet121 model.](https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset)
The Fig 2 illustrates a workflow for detecting tuberculosis in chest X-ray images using a architecture DenseNet121. The process begins with input images of chest X-rays, including both normal and processed versions. These raw images undergo preprocessing techniques to enhance their quality and prepare them for the subsequent stages. Preprocessing may include normalization, resizing, noise reduction, and contrast enhancement. Once preprocessed, the images are fed into a deep learning model, specifically a CNN architecture. The CNN architecture consists of several layers, including convolutional layers, pooling layers, and dense (fully connected) layers. The process starts with an initial convolutional layer (Conv1) that performs the initial convolution operations on the image to create feature maps. This is followed by multiple dense blocks (DB) and transition layers (TL), characteristic of DenseNet architectures. These dense blocks consist of several convolutional layers that receive input from all previous layers, promoting feature reuse and efficiency. The architecture includes further feature extraction layers, such as Dense Blocks 4, 5, and 6, which likely feed into a global pooling layer or a fully connected layer before classification. Finally, the model outputs a classification result, determining whether the X-ray indicates "Normal" or "Tuberculosis." This workflow highlights the use of deep learning, specifically CNNs with dense connectivity (DenseNet architecture), to automate the detection of tuberculosis in chest X-ray images by extracting relevant features and making a binary classification.

![Fig. 2 Workflow for detecting tuberculosis](image)

Fig. 3 The framework of the InceptionV3 model.

The Fig 3 illustrates the workflow for detecting tuberculosis in chest X-ray images using a convolutional neural network (CNN) with the InceptionV3 model. This process is divided into several stages:

**Input Image:** The process begins with a chest X-ray image, showing both the raw and preprocessed versions.

**Applying Preprocessing Techniques:** The raw chest X-ray undergoes various preprocessing techniques to enhance its quality and prepare it for analysis. These techniques may include normalization, resizing, noise reduction, and contrast enhancement.

**Feature Recognition with InceptionV3:** The preprocessed image is then fed into a deep learning model, specifically a CNN with the InceptionV3 architecture. The diagram shows several convolutional layers designed to extract features from the images. The structure of InceptionV3 is detailed, highlighting layers such as convolutional layers, activation functions, pooling layers, and possibly batch normalization layers. These layers work together to learn and extract relevant features from the input image.

**Results:** After passing through the InceptionV3 network, the model produces a classification result, determining...
whether the X-ray shows "Normal" or "Tuberculosis." The overall process emphasizes the use of deep learning, particularly CNN with the InceptionV3 architecture, to automate tuberculosis detection in chest X-ray images. By applying preprocessing techniques and feeding the enhanced images into an advanced neural network, the system can effectively extract important features and make accurate classifications.

The Fig 4 illustrates the workflow of detecting tuberculosis in chest X-ray images using deep learning techniques with the ResNet-50 model. The process begins with the input of black-and-white chest X-ray images. Next, these X-ray images undergo preprocessing techniques, which include contrast enhancement, noise removal, segmentation, and color map transformation to highlight certain features in the images. After preprocessing, the images are fed into the ResNet-50 model, a convolutional neural network (CNN) architecture consisting of 50 layers. ResNet-50 is known for its robust ability to identify important features in images and its effectiveness in addressing the vanishing gradient problem that often occurs in very deep networks. This model learns complex features such as edges, textures, and patterns relevant to tuberculosis. After passing through several layers in ResNet-50, the final result is a classification of whether the X-ray image shows normal lungs or signs of tuberculosis. This classification is performed at the final stage of the neural network using fully connected layers that output two categories: "Normal" or "Tuberculosis." Thus, the use of deep learning, particularly through the ResNet-50 model, enables the automatic and accurate analysis of medical images to assist in the diagnosis of tuberculosis.

RESULT

The platform used for testing in the conducted research is Google Colab, serving as the tool for model creation using the Python programming language. Data is directly retrieved from Kaggle using the Google Colab API, establishing a connection between Google Colab and Kaggle. Various variations in transfer learning model architectures, including DenseNet121, InceptionV3, and ResNet50, are required for performance analysis in classifying tuberculosis and normal lung diseases in individuals. Before starting the testing, preprocessing of the available images is performed by resizing the images to 224 x 224 pixels (scaling), rotating the images, and reducing noise to enhance focus on the test image data. All models used in the research are trained using the Adam optimization algorithm. Here are the evaluation results of the transfer learning models used in the study.
In Fig 6, the graph illustrates the accuracy and loss metrics of the DenseNet121 model, which underwent testing across 15 epochs. During training, the DenseNet121 achieved an accuracy of 0.9613 and a corresponding loss of 0.1189, while validation accuracy was maintained at 0.961. Transitioning to Fig 7, the graph showcases the performance of the InceptionV3 model over 10 epochs. It achieved a
commendable training accuracy of 0.9827 with a loss of 0.0756, and during validation, it maintained an accuracy of 0.9714 with a validation loss of 0.1561. Following this in Fig 8, the ResNet50 model's accuracy and loss trends over 10 epochs are displayed. Notably, the ResNet50 achieved a remarkable training accuracy of 0.9925 with a minimal training loss of 0.0328. Its validation accuracy was equally high at 0.9935, with a validation loss of 0.0950. These metrics are crucial in evaluating the models' effectiveness in distinguishing between tuberculosis-infected and normal lungs based on their respective labels (0 for normal lungs and 1 for tuberculosis-infected lungs). The graphical representations provide insights into how each model performs throughout the training and validation phases, highlighting their robustness and reliability in classification tasks. For a more comprehensive assessment of the models' classification capabilities, refer to the confusion matrix presented below, which further details their predictive accuracy and performance across different diagnostic scenarios.

The Fig 9 shows three different confusion matrices, each for a different deep learning model: ResNet50, DenseNet121, and InceptionV3. These confusion matrices are used to evaluate the performance of classification models by comparing the model's predictions to the actual labels. For the ResNet50 model, for label 0, the number of True Negatives (TN) is 523, and False Positives (FP) is 2. For label 1, False Negatives (FN) is 2, and True Positives (TP) is 103. This model has an accuracy of 99.37%, with both precision and recall at 98.10%, resulting in an F1-Score of 98.10%. In the DenseNet121 model, for label 0, there are 644 True Negatives and 56 False Positives. For label 1, there are 18 False Negatives and 122 True Positives. The accuracy of this model is 91.19%. The model's precision is 68.54%, recall is 87.14%, and F1-Score is 76.95%. Meanwhile, the InceptionV3 model shows 700 True Negatives and 0 False Positives for label 0. For label 1, there are 24 False Negatives and 116 True Positives, with an accuracy of 97.14%. The precision of this model is 100%, but its recall is lower at 82.86%, resulting in an F1-Score of 90.63%. From these three models, it can be observed that each model has its own strengths and weaknesses. ResNet50 demonstrates a very good balance between precision and recall. DenseNet121 shows more balanced performance but not as high in precision as the other two models. On the other hand, InceptionV3 has very high precision but lower recall compared to ResNet50.

*advazizwanputra@unprimdn.ac.id

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
The Table 1 compares three convolutional neural network models, namely ResNet50, InceptionV3, and DenseNet121, based on evaluation metrics for two labels (0 and 1): Accuracy, Precision, Recall, F1-Score, and Support. ResNet50 demonstrates the best performance with an accuracy of 0.99, and Precision, Recall, and F1-Score all at 1.00 for label 0, as well as Precision, Recall, and F1-Score of 0.98 for label 1. InceptionV3 has an accuracy of 0.97, with Precision at 0.97, Recall at 1.00, and F1-Score at 0.98 for label 0, and Precision at 1.00, Recall at 0.83, and F1-Score at 0.91 for label 1. DenseNet121 shows the lowest performance with an accuracy of 0.91, Precision at 0.97, Recall at 0.87, and F1-Score at 0.77 for label 1. Based on these results, ResNet50 is the best model, followed by InceptionV3 and DenseNet121.

CONCLUSION

The research utilized several transfer learning models for testing on the dataset, such as ResNet-50, InceptionV3, and DenseNet121, to classify tuberculosis and normal conditions in individual lungs. The process involved testing on a dataset divided into training data and validation data. Subsequently, preprocessing was conducted for testing purposes. The experimental results showed that ResNet-50 achieved the highest accuracy of 99%. The model with the lowest accuracy was DenseNet121 at 91%, followed by InceptionV3 with an accuracy of 97%. Therefore, it can be concluded that ResNet-50 excels in classifying tuberculosis and normal conditions in individual lungs compared to other models.

For further research, it is recommended to conduct testing using different models such as YoloV5, VGG16, and others. Additionally, the current models should be tested using larger-scale datasets. Exploring the impact of using larger and more diverse datasets can improve the generalizability of the models. Investigating the effectiveness of other state-of-the-art models and combining multiple models (ensemble methods) could potentially enhance classification performance. Finally, implementing techniques to reduce computational complexity and increase processing speed would be beneficial for real-time applications in medical diagnostics.

REFERENCES


Melnì, A. (2024). *HUBUNGAN FASE PENGOBATAN ANTI TUBERKULOSIS PARU DENGAN RASIO NEUROFIL LIMFOSIT (RNL) DAN RASIO MONOSIT LIMFOSIT (RML) PADA PASIEN EFUSI PLEURA TB DI RSUD H. ABDUL MOELOEK PROVINSI LAMPUNG*.


*advizwanputra@unprimdn.ac.id*


*advazizwanputra@unprimdn.ac.id*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.