Transfer Learning-based Convolutional Neural Networks in Pneumonia Recognition

Kaixi Xu*

Dept. of Computer Science, University of Ottawa, 550 R Cumberland St, Ottawa, ON, CANADA K1N 6N8

*Corresponding author: kxu092@uottawa.ca

Abstract. It is very hard for humans to distinguish the lungs of different kinds of people by looking at lung x-ray pictures with their eyes. However, artificial intelligence such as deep learning methods can help people distinguish these diverse images. The dataset of lung images has four categories, namely normal, COVID, viral pneumonia, and lung occupancy. Many models such as VGG and AlexNet can achieve high accuracy, but these models cannot guarantee that they can be used in other fields. Therefore, transfer learning can be considered to make sure one model can be used in multi-disciplines. In this work, the imagenet as the first dataset was used. After training in the imagenet dataset, the model got some edge characteristics from imagenet so that this model can distinguish other stuff. The model would distinguish the types of the lung images and give a correct justification. The transfer learning model only has three layers. By using the Adam optimizer and 0.0001 learning rate, the precision had arrived at 85% and the loss had arrived at 0.37. After importing the transfer learning, the machine can be used in several disciplines such as human recognition, flower recognition.

Keywords: Transfer learning, Convolutional neural network, Lung images detection.

1. Introduction

Since the year 2020, everything in the world such as economics and entertainment have become much worse due to the COVID-19 pandemic. The COVID-19 virus can be transferred through human contact. Some patients were attacked by this kind of COVID-19 virus and eventually died. Up till now, approximately 6.36 million people have died as a result of this illness, make it clear that this is a very big disaster. People worldwide are afraid of this kind of COVID-19 virus and would like to get rid of it as soon as possible. The only way to solve this problem is to cure all the patients who get COVID-19. However, firstly healthcare workers must confirm whether one patient is a COVID-19 patient or not and continue their treatments. The COVID-19 symptom starts in the lungs, so the lungs of COVID-19 patients are quite different from the lungs of normal people or those patients who get normal pneumonia. However, it is very difficult for people even doctors to 100 percent correctly differentiate whether a person is a normal person, a normal pneumonia patient, or a COVID-19 patient through the x-ray images of lungs because those x-ray images look the same. Many people died because the doctors could not detect them as normal pneumonia or COVID-19. In this case, deep learning-based methods can be considered an ideal alternative to solve this issue to a certain extent due to their excellent performance.

From the prior research, researchers employed deep learning and lung images to develop a single model based on the machine [1-5], in order to predict the results. Some experimenters use the Convolutional Neural Network (CNN) model that is built by themselves, the others use some existing models to predict the results, such as LeNet, AlexNet, Visual Geometry Group (VGG), and ResNet [6]. Some models can achieve high accuracy. For instance, both AlexNet and Visual Geometry Group can achieve an average of 90% accuracy in differentiating these lung pictures. However, these models do not have any sustainability and reusability. If one CNN model can acquire 95% accuracy in distinguishing the animals, then maybe it cannot have very high accuracy in differentiating the vehicles such as cars or trucks. In order to ensure one model can differentiate almost everything, a new technology called transfer learning can be introduced to maintain the sustainability and reusability of the model.

Transfer learning is a method of deep learning [7-10]. It transfers the model that has been trained to differentiate many images of many different species such as animals, food and etc., to a new mission that differentiates different kinds in a new or old species. The dataset from the previous task has a greater number of pictures than the number of images from the current(second) dataset. If the second mission dataset image number is larger than the first task's, the accuracy will not be very high since this model may not have seen some of these images from the second task. In this project, a large dataset called imagenet is introduced to be the dataset of the first task, and a model has been trained to recognize all the images in imagenet. The imagenet dataset has about 14 million images over 20000 categories, so the model which can recognize all the images in imagenet can also differentiate the dataset in this project. Therefore, this model will also recognize the x-ray images of lungs in this project.

2. Methodology

2.1. Datasets

In this project, two datasets are used in transfer learning, since normally there are two or more datasets in transfer learning. The machine studies some low-level features such as edges, and colors from the first dataset, and is used in the second or other datasets. For the first dataset, this paper used the imagenet. Imagenet has a large number of data, and it does not have any medical images. After the machine had been well-trained in the first dataset: imagenet, the machine studied the low-level feature such as edges information from the imagenet. The machine used these low-level features to recognize the categories on the second dataset, which was about the lung pictures of four lung situation categories (COVID-19, normal, lung opacity, and viral pneumonia). These data were retrieved from Kaggle (https://www.kaggle.com/datasets/preetviradiya/covid19-radiography-dataset). The number of these four categories is 3616: 10192: 6012: 1345. Therefore, the total number would be 20,000. However, this number is greatly less than the number of pictures in imagenet, and that is normal. The train, validation, and test data rate are 6:2:2. That is, this paper would randomly choose 60% data in the training set, and 20% in both validating and testing sets. These images are really large amounts of data, and it is not needed to do undersampling or oversampling to deal with the imbalance between categories. Figure 1 shows the sample data in the dataset.

Figure 1. The lung x-ray image of a COVID-19 patient.

2.2. Model

The model of this transfer learning was imported from the TensorFlow.keras. The structure of this transfer learning model has three layers. The first layer is called the base model, which is used for "transfer". This layer transfers some edge characters from imagenet to the lung image data. It also shapes the model input, which is (128,128,3). At the same time, this paper made the trainable of this layer to be true. The second layer is the GlobalAveragePooling layer, which seizes the input shape one more step, from 3d-shape to 1d-shape. The last layer is a dense layer, this layer divides the data into four categories. The activation function is softmax due to multi-class.

2.3. Implementation details

The current precision arrives at 85%. This processed the hyperparameter tuning. The learning rate is 0.0001, the optimizer now is Adam, the batch size is 32, the loss function is CategoricalCrossentropy, the metric is accuracy, and the epoch is 20.

3. Results and discussion

3.1. Results

By adjusting the hyperparameters, the accuracy has arrived at 85.05% accuracy and 0.3907 loss. This result is not very high, but it is also not very low. Initially the machine could only achieve only 24.36% accuracy and 2.0653 loss, which was a pretty bad situation. To improve this low accuracy and high loss problem, it is necessary to do hyperparameter tuning. This study did not change the epoch and batch size parameter. These two hyperparameters are maintained at 32(batch size) and 20(epoch). However, other hyperparameters such as optimizer and learning rates remanded the revision. In this project, several parameters were tried to test the dataset.

Table 1. The accuracy under different optimizers and learning rates

	0.001	0.0001	0.00001
Adam	84.22%	85.05%	83.14%
RMSprop	82.34%	80.43%	76.54%
SGD	83.45%	79.56%	68.50%

From Table 1, the experiments show that the optimizer=Adam, and the learning rate =0.0001 has the best effect. This study also changed the trainable from False to True. Therefore, the model can be trained more efficiently.

3.2. Discussion

From the result, it can be still known that 15% of the pictures are not correct. This may be because the transfer learning only transfers the edge characters of the imagenet. The actual imagenet does not contain the medical images. Therefore, the machine which has studied the distinguishing of the images in imagenet has no experience in distinguishing the medical images. Also, the pictures of the lungs look similar, such as some lung opacity pictures and viral pneumonia pictures, even though humans' eyes, they are pretty difficult to distinguish. Although better than humans, the machine is sometimes also difficult to distinguish the similar images. So, it is normal for machines to make some inaccuracies.

Besides this minor inaccuracy, the machine behaves well. It can still correctly about 16,000 out of the 20,000 pictures. During the 20 epochs of the training and validating sections, the accuracy got the highest during the 20th epoch of both parts. Worth to say, the accuracy in the 20th epoch of the validation part is even higher than the accuracy of the final test. There still exists high potential to improve the accuracy in the final test. Compared to some designed convolutional neural network such as AlexNet and VGG, the transfer learning can still get the same or even higher accuracy. At the same time, after training on this dataset in transfer learning, the model can also be trained in other datasets or other fields to achieve higher accuracy. Transfer learning can make sure the model is multi-usable.

During the construction of the model, the accuracy could not be over 52% for a long time. This study tried every hyperparameters such as optimizer and learning rates, but the result could not be improved. However, it can be found that the trainable attribute of the first layer of the model was set to be False, which means that all the layers' weights become non-trainable during being trained in the imagenet. The state of the first layer would be freezed, this state will not be updated during the training step in the imagenet. After this study changed this attribute to True, the accuracy became higher, and the loss became lower.

4. Conclusion

In this work, transfer learning was applied in lung image recognition to distinguish different lung conditions such as normal or COVID or pneumonia through the imagenet dataset. The model achieved 85% accuracy, which is a good result. The experiment shows that the result of the transfer learning is as good as other models such as AlexNet and VGG. At the same time, transfer learning can also be used in other domains such as distinguishing transportations or food recognition. In the future, further study plans to improve the accuracy of the transfer learning model and to utilize this kind of model in other fields.

References

[1] Xu, R., et al. "Histopathological Tissue Segmentation of Lung Cancer with Bilinear CNN and Soft Attention." BioMed Research International 2022 (2022).

[2] Xu, Z., et al. "ISANET: Non-small cell lung cancer classification and detection based on CNN and attention mechanism." Biomedical Signal Processing and Control 77 (2022): 103773.

[3] Tejaswini, C., et al. "CNN Architecture for Lung Cancer Detection." 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT). IEEE, 2022.

[4] Praveena, M., et al. "Lung Cancer Detection using Deep Learning Approach CNN." 2022 7th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2022.

[5] Yuchen, G., et al. "Lung cancer diagnosis system based on 3D CNN." Second International Conference on Medical Imaging and Additive Manufacturing (ICMIAM 2022). Vol. 12179. SPIE, 2022.

[6] Qiu, Y., et al. "Semantic segmentation of intracranial hemorrhages in head CT scans." 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS). IEEE, 2019.

[7] Torrey, L., and Jude, S., "Transfer learning." Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. IGI global, 2010. 242-264.

[8] Weiss, K., et al. "A survey of transfer learning." Journal of Big data 3.1 (2016): 1-40.

[9] Zhuang, F., et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.

[10] Pan, S. et al. "A survey on transfer learning." IEEE Transactions on knowledge and data engineering 22.10 (2009): 1345-1359.