Transfer Learning for Image Classification

Aaron Modiyil Joseph

22018497

Code: https://colab.research.google.com/drive/1vOLhSgV7Ds34iRZsS6KB4YfAM3Pvap2M?usp=sharing

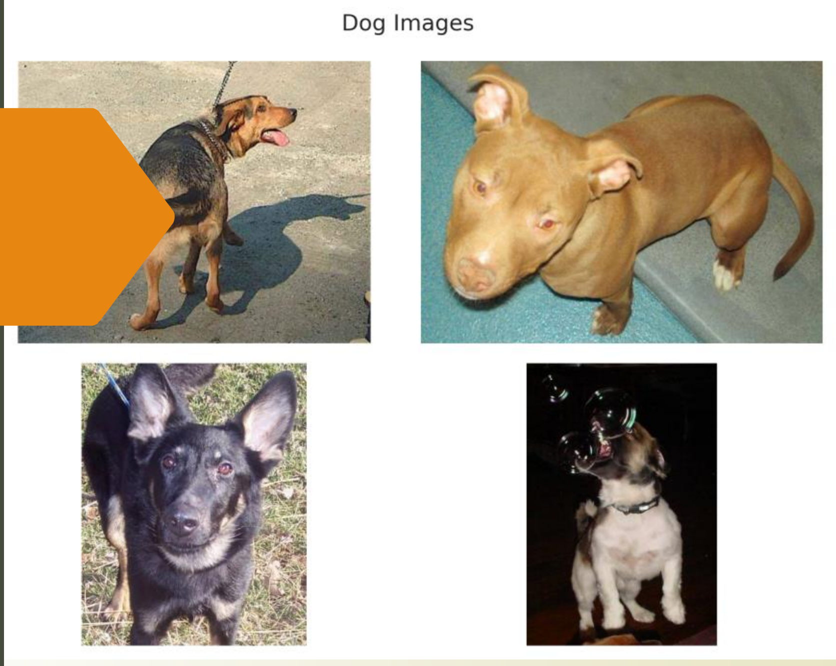
University of Hertfordshire UH

Dataset Overview

The Dataset used is the Dogs vs. Cats Redux: Kernels Edition. The dataset is available on Kaggle and is used for a binary classification task to classify images of dogs and cats.

The folder names ‘train’ includes 25,000 images of each class. Each image's filename includes the label. The classes are balanced with 12,500 images each.

(Cukierski, 2016)



Transfer Learning

Transfer learning, a subset of Machine Learning, is the re usage of previous models to address present challenges. The basic concepts of pre-existing training data are used to improve the performance of the current challenge, removing the need to create it from scratch.

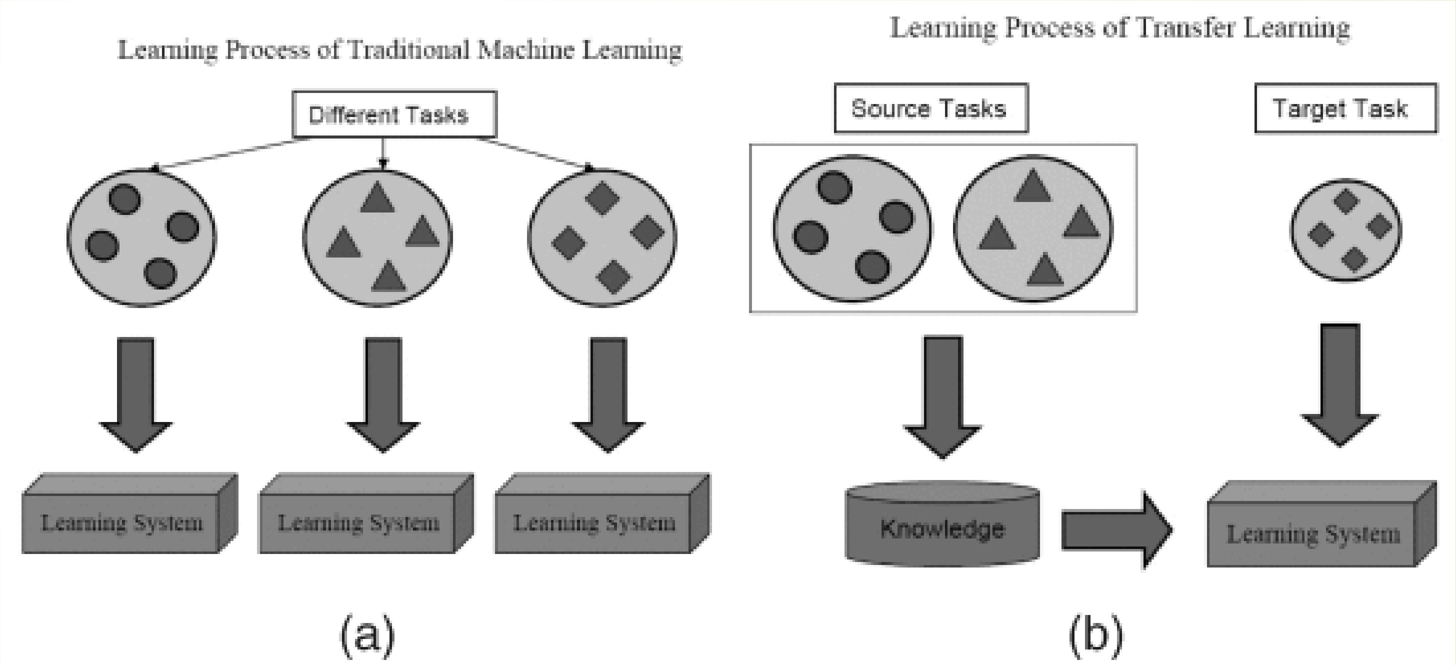


Figure 1: Learning process difference between (a) traditional machine learning and (b) transfer learning (Hosna et al., 2022) .

(Hosna et al., 2022)

Benefits and Importance

Contradiction	Traditional ML	Transfer learning
Large data versus little annotation	Expensive human labeling	Transfer knowledge from existing fields
Big data versus poor computation	Exclusive computation device	Model transfer
Limited data versus generalization ability	Poor performance for OOD data	Meta-learning, domain generalization, etc.
Pervasive versus personal need	Cannot meet personal demands	Adapt to each individual
Particular uses	No solution for cold starts	Transfer of data from other fields

Table 1: Benefits and importance of transfer learning

(Wang J. and Chen Y., 2023)

Pre-Trained Model: VGG16

K. Simonyan and A. Zisserman of the University of Oxford put forward VGG16, a convolutional neural network model, in their paper "Very Deep Convolutional Networks for Large-Scale Image Recognition".

VGG16 was trained on the ImageNet dataset over many weeks using NVIDIA Titan Black GPUs. The ImageNet comprises about 15 million images with high resolutions labelled into around 22,000 categories.

VGG16 is a convolutional neural network that accepts a 224 × 224 RGB image as input and sends it through a succession of convolutional layers with small receptive fields (3x3 or 1x1) while maintaining spatial resolution. It employs max-pooling across 2x2 with stride 2. The network further consists of three dense layes: the first two contain 4096 channels each, while the last one conducts 1000-way ILSVRC classification. All hidden layers employ the rectification (ReLU) non-linearity. Except for one configuration, the networks do not use Local Response Normalisation (LRN). The final layer uses softmax.

On ImageNet, VGG16 has 92.7% test accuracy, ranking in the top five.

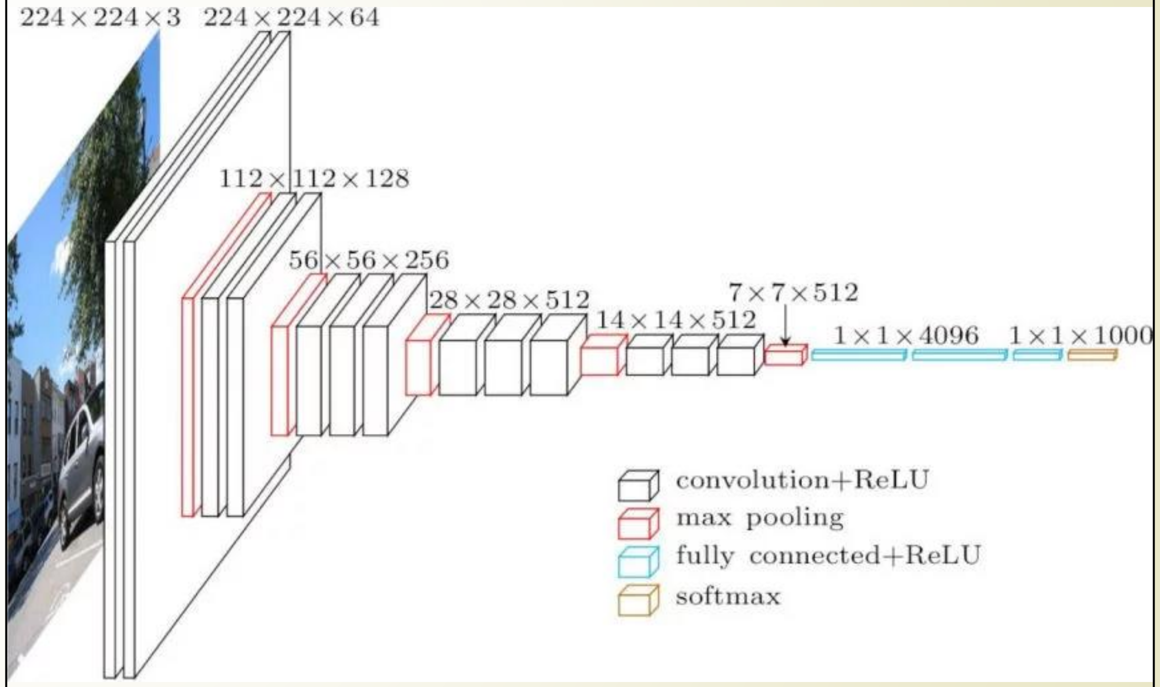


Figure 2: Architecture of VGG16 (Hassan, 2018).

((Hassan, 2018), (Simonyan K. and Zisserman A., 2015))

University of Hertfordshire UH

Data Pre-processing

Extracting the Dataset

- Extract the zipped dataset from Google Drive into a specified directory.

Directory Setup

- Create directories for training and validation datasets, further divided for cat and dog images.

Dataset Splitting

- The first 10,000 cat and dog images form the training set, and the next 2,500 form the validation set.

Data Augmentation

- Techniques including rescaling, random rotation, width and height shifts, shear transformation, zooming, and horizontal flipping are applied to diversify the training set.

Loading Images

- The images are loaded using flow_from_directory. The images are resized to 150x150 pixels and are set up in batches of 50. The class_mode is set to 'binary' for binary classification.

Methodology

Transfer Learning: The VGG16 model is used as a feature extractor. The top layers of the model, which are responsible for classification, are not included. The output of this model is then passed through a Flatten layer and two Dense layers. The final Dense layer uses a sigmoid activation function, making this setup suitable for binary classification. The VGG16 model is frozen and trained for 20 epochs. The RMSprop optimizer is used with loss function - binary cross-entropy and metric - accuracy.

Fine-Tuning: After the initial training, last two blocks of the VGG16 model are unfrozen. This is done because the model's initial layers learn more general traits, whilst the later levels acquire more specialized features. The learning rate is reduced to avoid large updates that could destroy the pre-learned features. The model is then recompiled and trained for another 20 epochs.

Training from Scratch: Finally, the VGG16 model's layers are all unfrozen, allowing the entire model to be trained on the new dataset from scratch. A new model is built and compiled in the same way as before, and then trained for 20 epochs.

	Transfer Learning Model	Fine-tuned Model
Loss	0.2617	0.2393
Accuracy	0.8879	0.8958
Val Loss	0.2033	0.1912
Val Accuracy	0.9166	0.9230

Table 2: The transfer learning model and the fine-tuned model, at the epoch of best validation accuracy of each model.

- Each parameter has become better by the fine-tuning, which is evident from figure 3 and 4, and table 2.
- The model has adapted and learned more specific features related to our dataset through fine-tuning.

Results

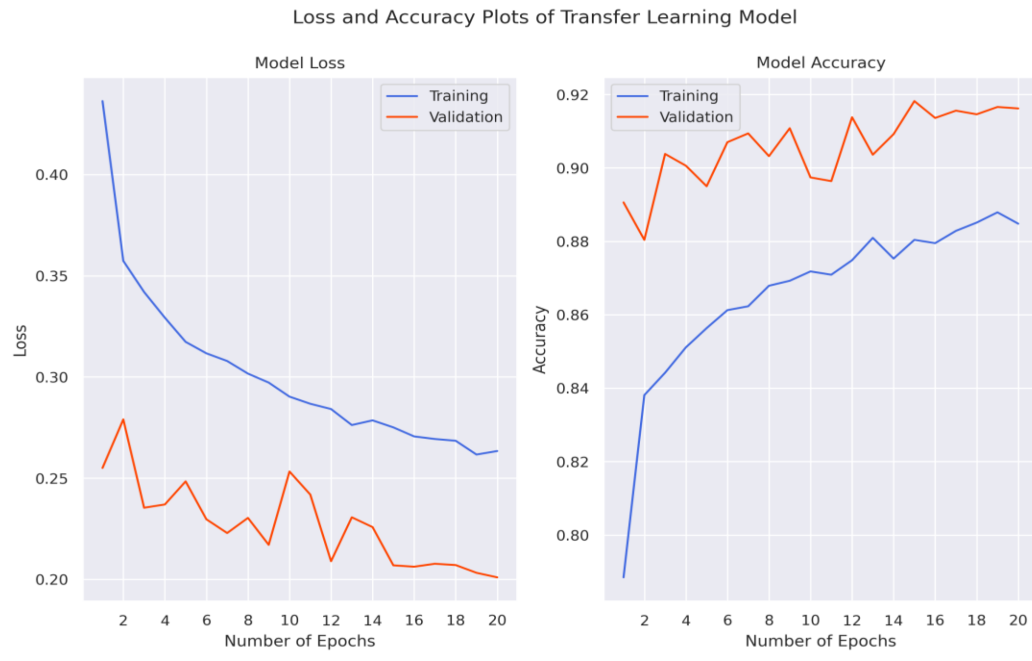


Figure 3: Loss and accuracy plots of transfer learning model

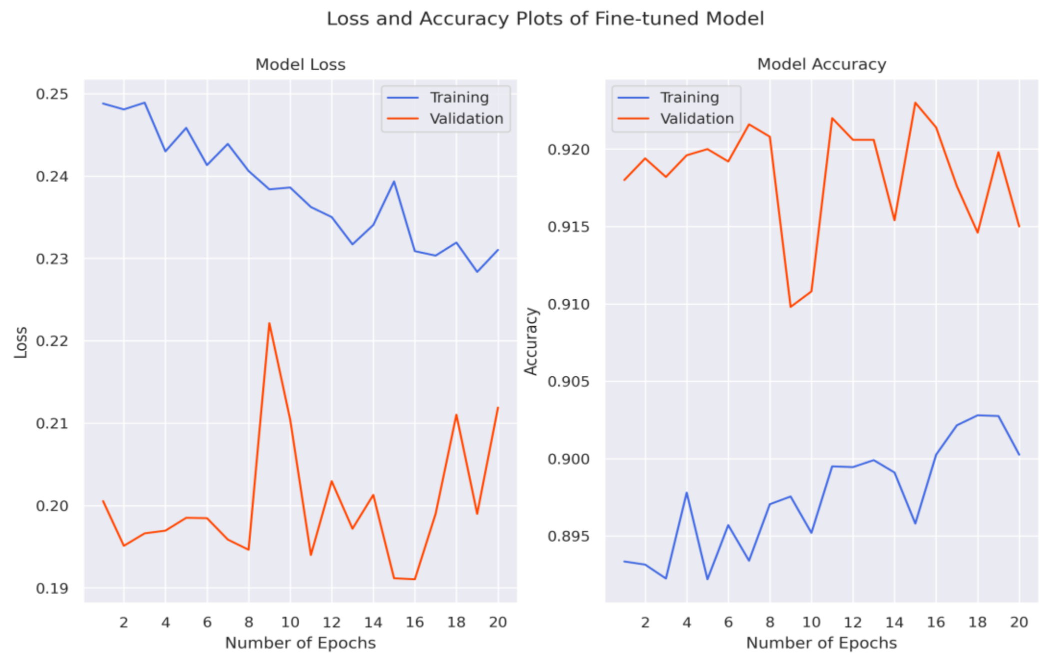


Figure 4: Loss and accuracy plots of fine-tuned model

- It is evident from figure 4 and 5, and table 3 that each parameter is better in the fine-tuned Model.
- Fined-tuned model performs better than the model trained from scratch.
- The fine-tuned model generalises better and is less prone to overfitting, as evidenced by the smaller difference between accuracy and validation accuracy, as well as between loss and validation loss compared to the model trained from scratch.

	Fine-tuned Model	Model Trained from Scratch
Loss	0.2393	0.2651
Accuracy	0.8958	0.8838
Val Loss	0.1912	0.2004
Val Accuracy	0.9230	0.9178

Table 3: The fine-tuned model and the model trained from scratch, at the epoch of best validation accuracy of each model.

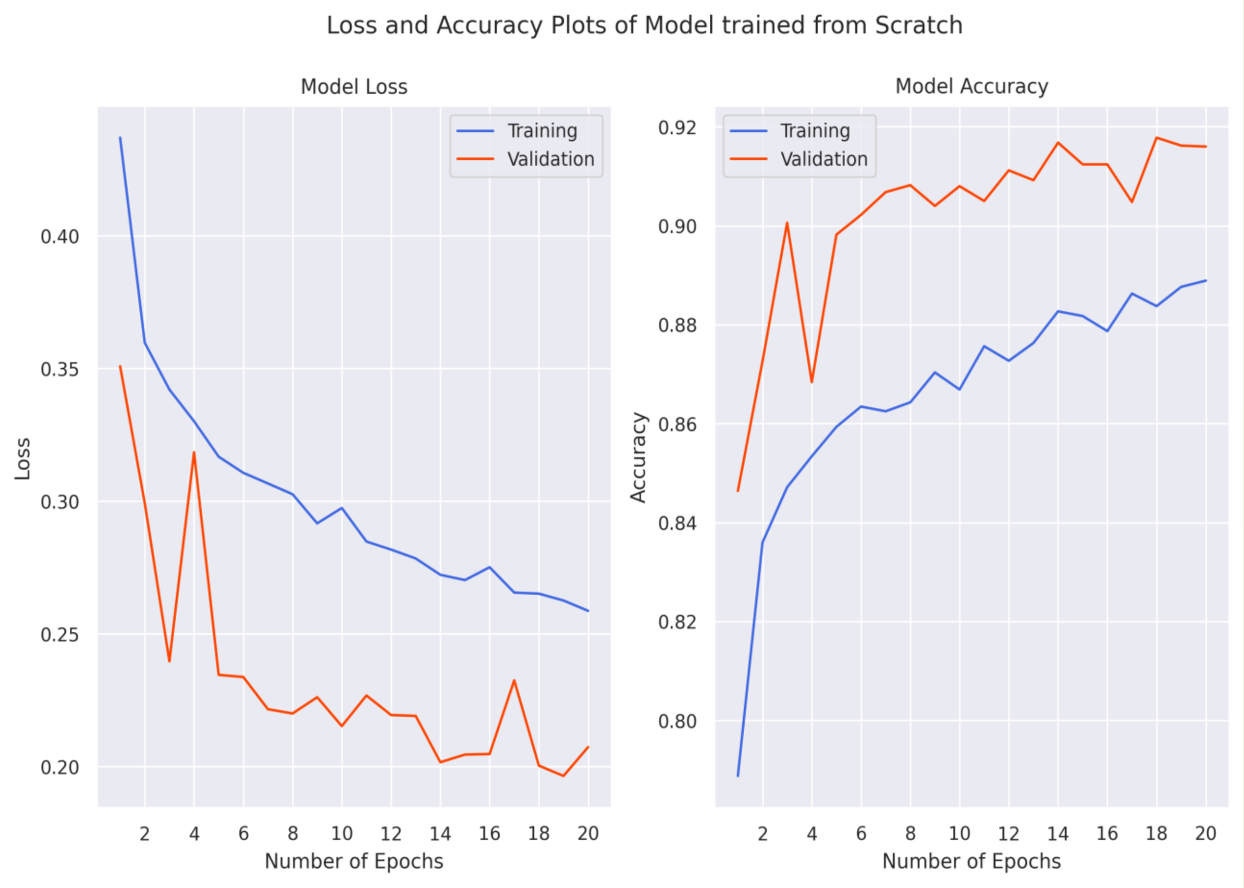


Figure 5: Loss and accuracy plots of model trained from scratch.

- Fine-tuned transfer learning model requires less computional power than the model trained from scratch, as the former has already learned useful features from the pre-training task.
- In conclusion, fine-tuning enhances the performance of a transfer learning model and a fine-tuned transfer learning model generalizes better, is less prone to overfitting, and requires less computational power compared to a model trained from scratch. Therefore, fine-tuning and transfer learning proves to be a highly effective strategy for image classification tasks.

Limitations, Potential Pitfalls and Ethical Implications

Arbitrary Configuration: Transfer learning is arbitrarily configured. This might result in inferior performance if the pre-trained model or transfer learning technique is not appropriate for the job.

Usage of Multiple Layers: The use of several layers and combinations in transfer learning models can impede the ability to detect and compare individual layers and their impact on performance.

Pre-trained Model Suitability: The pre-trained model may not be appropriate for the job at hand. In some circumstances, the model may need to be retrained to produce optimal results. Furthermore, the pre-trained model may be too huge to do the new task.

Limited Interpretability: Models for transfer learning have limited interpretability. This may pose ethical concerns when understanding the decision-making process of a model is crucial.

Data Privacy: Concerns about confidentiality and privacy arises when training datasets are directly accessible.

Data Retention: Transfer learning models may keep input data for reasons such as model training, potentially exposing personal information to third parties. Be aware of data retention policies and ensure that sensitive data is not stored or is properly safeguarded.

((Kardas, 2023), (Kim, Cosa-Linan et al., 2022), (Noema, 2022), (Zhang, Zhu et al., 2022))

Potential Areas of Improvement

Optimal Configuration: Transfer learning's efficacy may be improved by empirically evaluating several ways to determine the ideal configuration.

Using other Models: Using deeper and better models, such as ResNet, Xception or Inception, as feature extractors can reduce time and computing costs without paying the price of predictive power.

(Kim et al., 2022)

References

- Cukierski W. (2016). Dogs vs. Cats Redux: Kernels Edition. Kaggle. Available at: https://kaggle.com/competitions/dogs-vs-cats-redux-kernels-edition (Accessed: 16 January 2024).
- Hassan M. U. (2018). VGG16 – Convolutional Network for Classification and Detection. Nuerohive. Available at: https://neurohive.io/en/popular-networks/vgg16/ (Accessed: 16 January 2024).
- Hosna A., Merry E., Gyalmo J., Alom Z., Aung Z. and Azim M. A.. Transfer learning: a friendly introduction. J Big Data 9, 102 (2022). Available at: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00652-w (Accessed: 16 January 2024).
- Kardas R. (2023). AI, Data Security, and Confidentiality in Accountancy. PICPA. Available at: https://www.picpa.org/articles/cpa-now-blog/cpa-now/2023/06/28/ai-data-security-and-confidentiality-in-accountancy (Accessed: 16 January 2024).
- Kim H. E., Cosa-Linan A., Santhanam N., Jannesari M., Maros M. E. and Ganslandt T. (2022). Transfer learning for medical image classification: a literature review. BMC Med Imaging 22, 69. Available at: https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-022-00793-7 (Accessed: 16 January 2024).
- Noema Y. (2022). Image Classification With Transfer Learning (TL). Medium. Available at: https://medium.com/imagescv/image-classification-with-transfer-learning-tl-3031c8be1ad7 (Accessed: 16 January 2024).
- Simonyan K. and Zisserman A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. Available at: https://arxiv.org/abs/1409.1556 (Accessed: 16 January 2024).
- Wang J. and Chen Y. (2023). Introduction to Transfer Learning: Algorithms and Practice. Springer.
- Zhang J., Zhu H., Wang F., Zhao J., Xu Q., Li H. (2022). Security and Privacy Threats to Federated Learning: Issues, Methods, and Challenges, Security and Communication Networks, vol. 2022, Article ID 2886795. Available at: https://www.hindawi.com/journals/scn/2022/2886795/ (Accessed: 16 January 2024).