# DEEP LEARNING ARCHITECTURES FOR IMAGE CLASSIFICATION

Michael Polonio

# INTRO

- Image classification, a fundamental task of computer vision, has applications in many fields such as the automotive industry, medical field, manufacturing, and many more.
- We depend on advanced techniques such as machine learning/deep learning to analyze the images efficiently and perform the classification
- Deep learning methods can vary in complexity and performance
- Transfer learning allows us to leverage robust models that were pre-trained on a similar problem

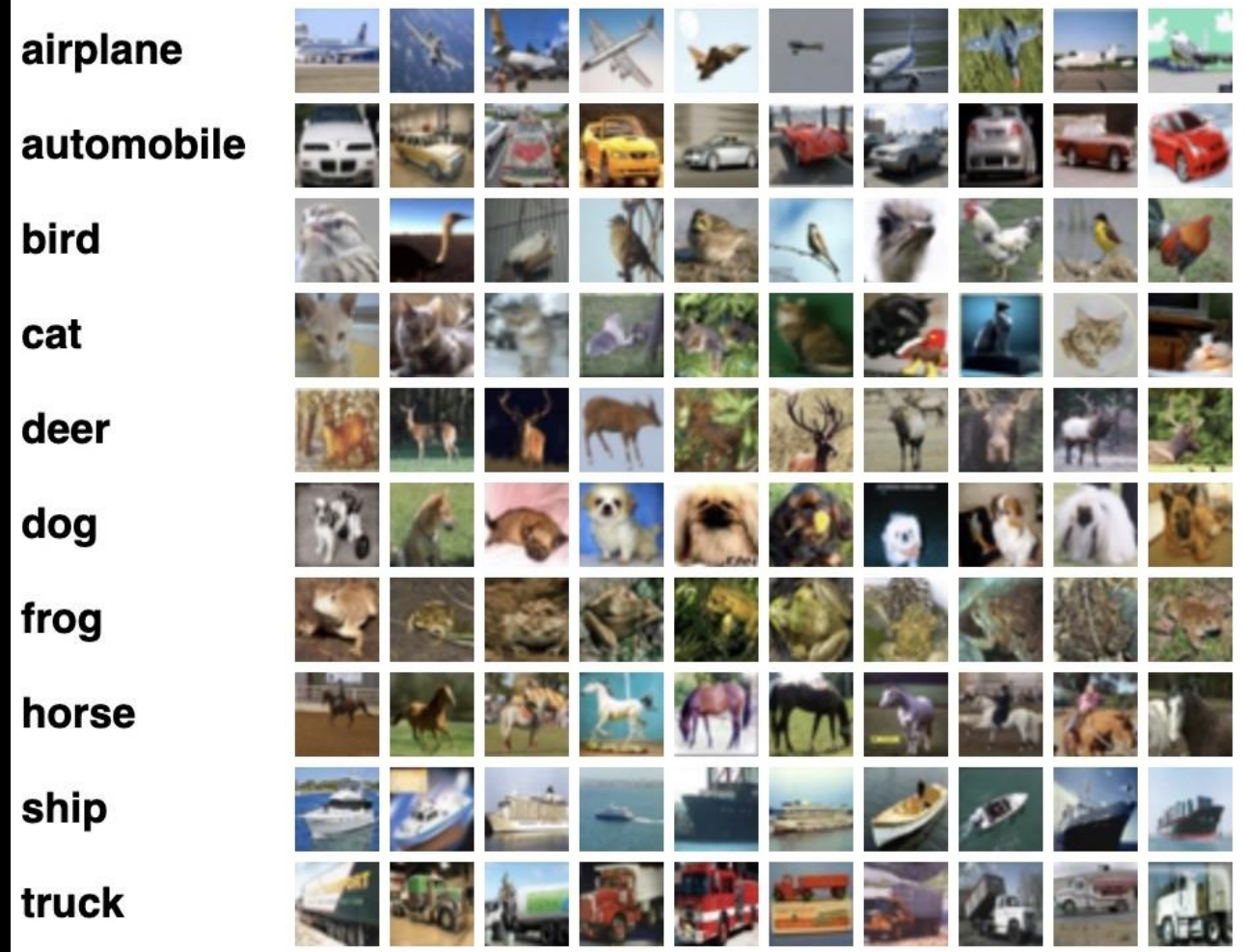
# PROBLEM STATEMENT

Can using *transfer learning* greatly increase the accuracy of an image classification model?

What are the performance differences between deep learning models of varying architecture and complexity?

# DATA : CIFAR-10

- 60,000 32x32 pixel RGB images
- 10 classes
- 50,000 training / 10,000 test



# METHODS

- Python (Keras, Matplotlib)
- Normalize data values from 0 to 1
- Construct Four Deep Learning Models with different architectures including two transfer learning models
- Hyperparameter tuning
- Evaluate learning curves and accuracy, confusion matrix

# MODELS

*Multilayer Perceptron (MLP):*

-1 hidden layer with 128 nodes and a sigmoid activation function

-20% dropout to avoid overfitting

-Output layer has 10 nodes, one for each label (softmax)

-SGD optimizer & categorical crossentropy loss function

*-25 epochs*

*Convolutional Neural Network (CNN):*

-First two convolutional layers have 32 filters, a 3x3 kernel, and a ReLU activation function. One layer has a 30% dropout and the other has max pooling

-The next two layers have the same hyperparameters but with 64 filters

-The next two layers have the same hyperparameters but with 128 filters

-Then data is flattened and passed to the last hidden layer, with 128 nodes.

-Output layer has 10 nodes (softmax)

-SGD optimizer & categorical crossentropy loss function

# MODELS - TRANSFER LEARNING

VGG-19:

-VGG-19 is a CNN with 16 layers and trained on 14 million images belonging to nearly 1000 classes (Boesch).

- data was trained on VGG then flattened and passed through a succession of hidden layers with 256, 128, 64 nodes with ReLU activation functions

- SGD optimizer

- each with a dropout to avoid overfitting

-5 epochs

ResNet-50:

-Resnet is a CNN that is 50 layers deep & trained on more than one million images

- data was trained on ResNet-50 then flattened and passed through two hidden layers with 64 and 128 nodes. Both ReLU activation function and a dropout of 30%

-Adam optimizer

-20 epochs

# RESULTS

The VGG-19 transfer learning model outperformed all other models with an accuracy of 85%

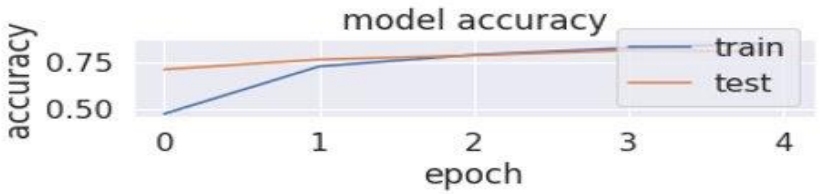
The ResNet-50 model performed worse than the less complex CNN, possibly indicating ResNet is overkill for this problem

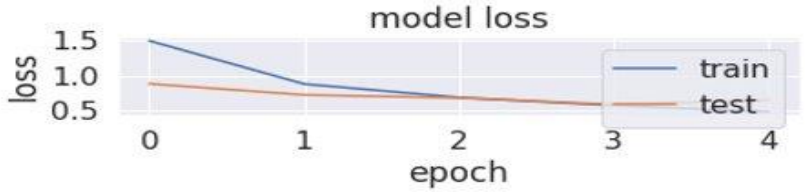
The models showed a wide range of performance, varying by up to a difference of 35% (MLP vs VGG-19)

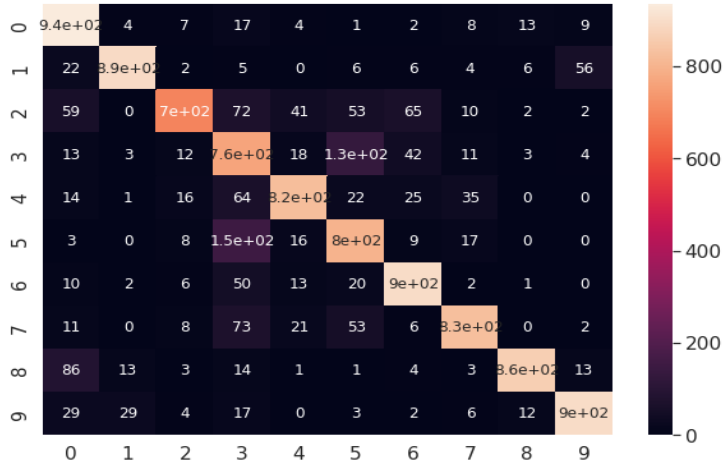
| Model | Accuracy |
| --- | --- |
| Multilayer Perceptron | 50% |
| Convolutional Neural Network | 74% |
| VGG-19 Transfer Learning | 85% |
| ResNet-50 Transfer Learning | 71% |

# RESULTS CONT'D

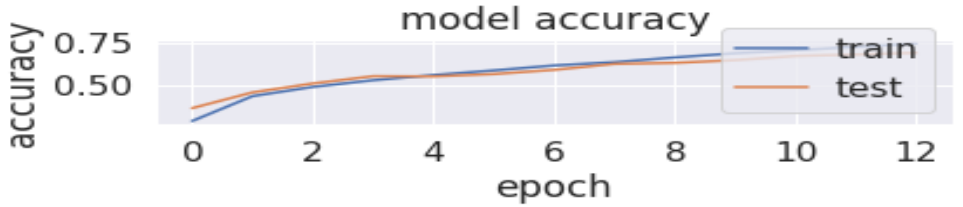
## VGG-19 model 85% accuracy

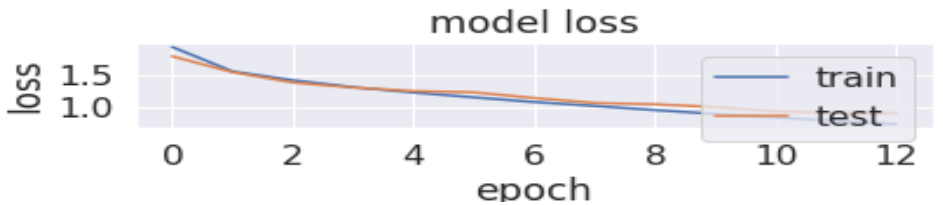


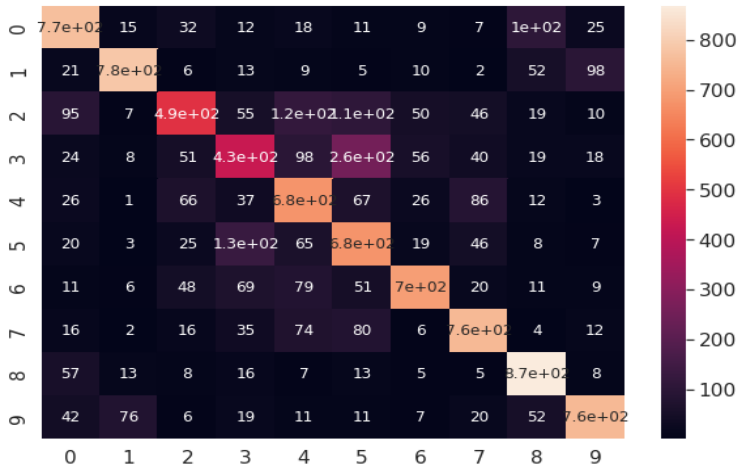




## CNN model 74% accuracy



# CONCLUSION

- An additional 11% accuracy was achieved by using *transfer learning, VGG-19 i*n particular.
- The various architectures yielded a wide range of results
- In transfer learning, multiple different models should be trained. Just because a model is more robust does not mean it will perform better. VGG-19 performed better than the more robust ResNet-50.

# REFERENCES

CIFAR-10 Dataset: CIFAR-10 and CIFAR-100 datasets (toronto.edu)

Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images.

Boesch, Gaudenz (2020). A Complete Guide to Image Classification in 2022. A Complete Guide to Image Classification in 2022 - viso.ai

# THE END