Machine Learning (COMP09012)

Assignment 2

Maksim Gusev

Atlantic Technological University

Introduction:

Kaggle dataset - GTSRB - German Traffic Sign Recognition Benchmark.

Transfer learning - take a pre-trained VGG16 network and retrain it to recognize traffic signs.

Methods:

Data preprocessing:

- Keras & TF import;
- Dataset loading;
- Dataset analysis;
- Image rescaling;
- Image cropping;
- Image centering;
- Train and verification sets;
- Test set;

CNN – VGG16:

- Instantiate VGG16 base model;
- Load pre-trained weights;
- Adding a new layer as the output (new model);
- Hyperparameter tuning;
- Train the new model on GTSRB dataset;
- Model validation;
- Metrics and assessment;

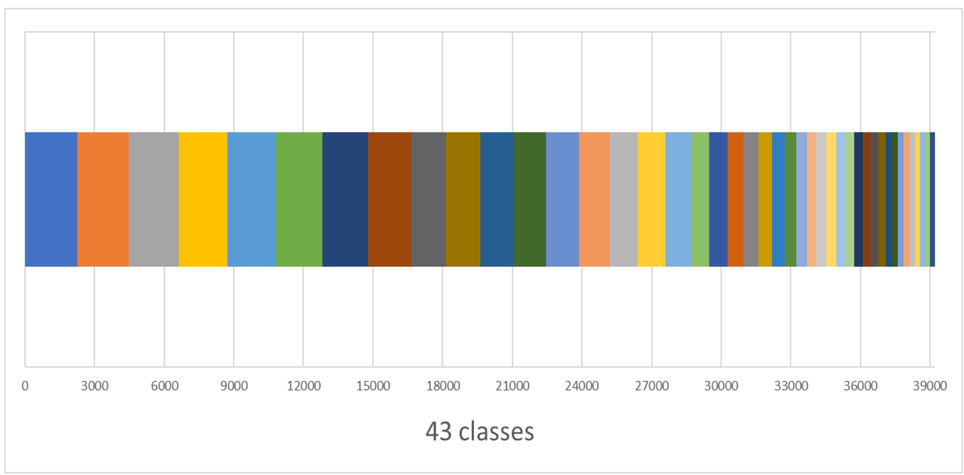
Results:

Test dataset analysis:

Total: 43 classes = 39209 png = 100%

Majority: 17 classes = 28710 png = 73.2%

Minority: 26 classes = 10499 png = 26.7%

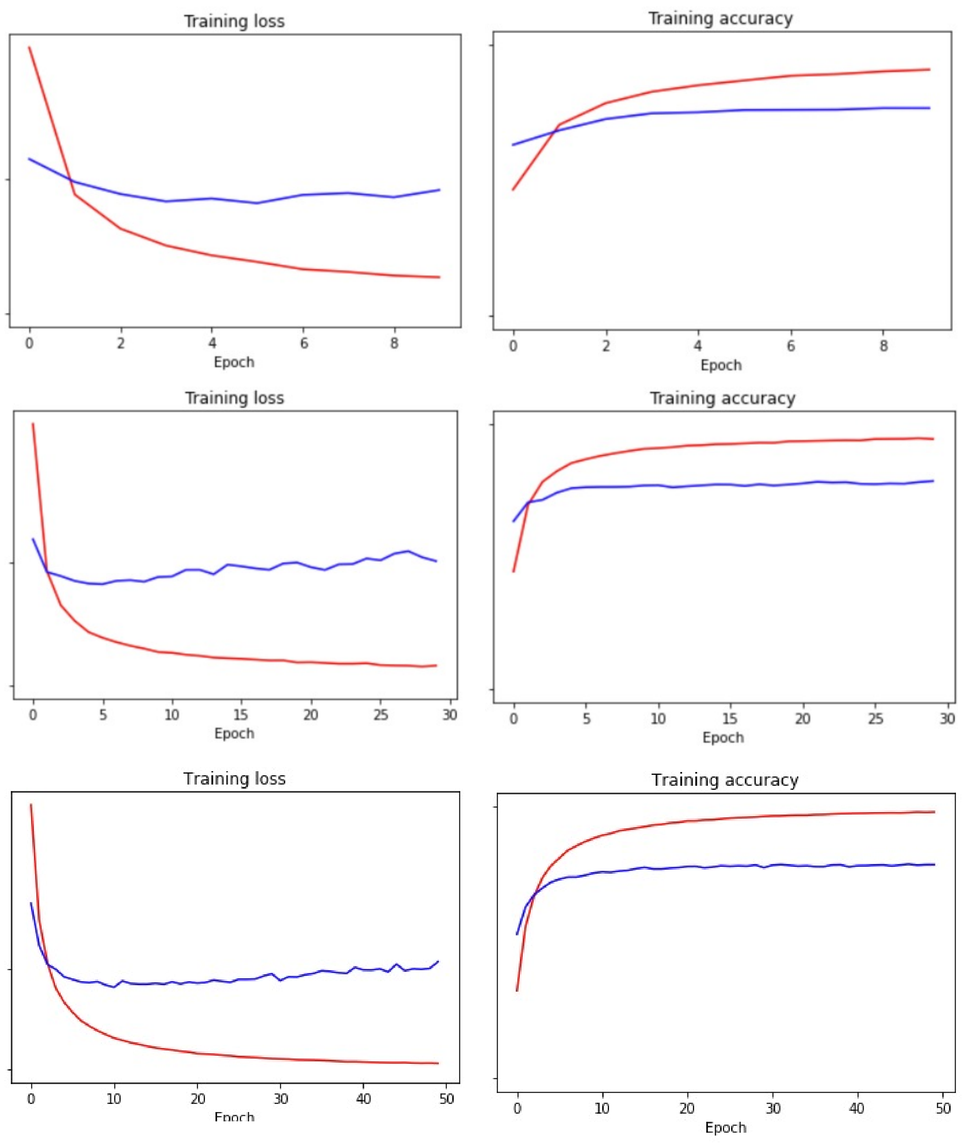


Model results:

Epoch 10: Train Acc. = 90.8% / Test Acc. = 76.6%

Epoch 30: Train Acc. = 94.4% / Test Acc. = 78.5%

Epoch 50: Train Acc. = 98.0% / Test Acc. = 78.6%



Conclusions:

Transfer learning saves time.

- 10 epochs - 23 min.
- 30 epochs - 65 min.
- 50 epochs - 116 min.

Dataset is imbalanced.

- 39209 pictures for 43 classes.
- 1 class takes from 5.7% to 0.5% (10+ times difference).

More epochs & accuracy required more training data.

- 10 epochs – underfit. (90,8% train acc. and 76,6% test acc.);
- 30 epochs vs 10 – fit. (+3,6% train acc., but +1,9% test acc.);
- 50 epochs vs 30 – overfit. (+3,6% train acc., but +0,1% test acc.);

Literature cited

Jake Vander Plas, Published O'Reilly in 2022. Python Data Science Handbook. Essential Tools for Working with Data.

Joel Grus, Published O'Reilly in 2022. Data Science from Scratch: First Principles with Python.

Acknowledgments

Ollscoil Teicneolaíochta an Atlantaigh

Atlantic Technological University

Further information

https://github.com/GusevPortfolio/Machine-Learning-1-assignment-.git

https://www.kaggle.com/datasets/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign?resource=download