Detecting Brain Tumours

Joel Koch and Bastian Schuchardt

Joel Koch & Bastian Schuchardt

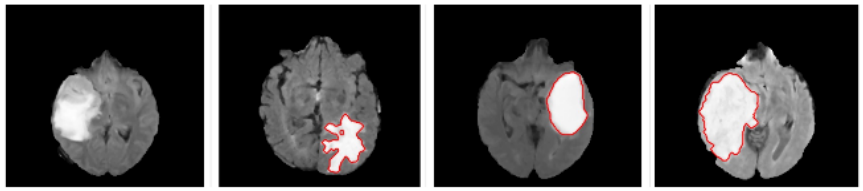
Introduction to the Problem

- Classify malignant & benign brain tumours
  - High accuracy for implementation in medicine
- Two doctors need to check which type of tumour
  - → Replace 2nd doctor with ML model

- Train CNN for malignant/benign tumours

Joel Koch & Bastian Schuchardt
18th of July, 2024

The Dataset



- „Brain Tumor“ kaggle dataset
- 3762 images with labels

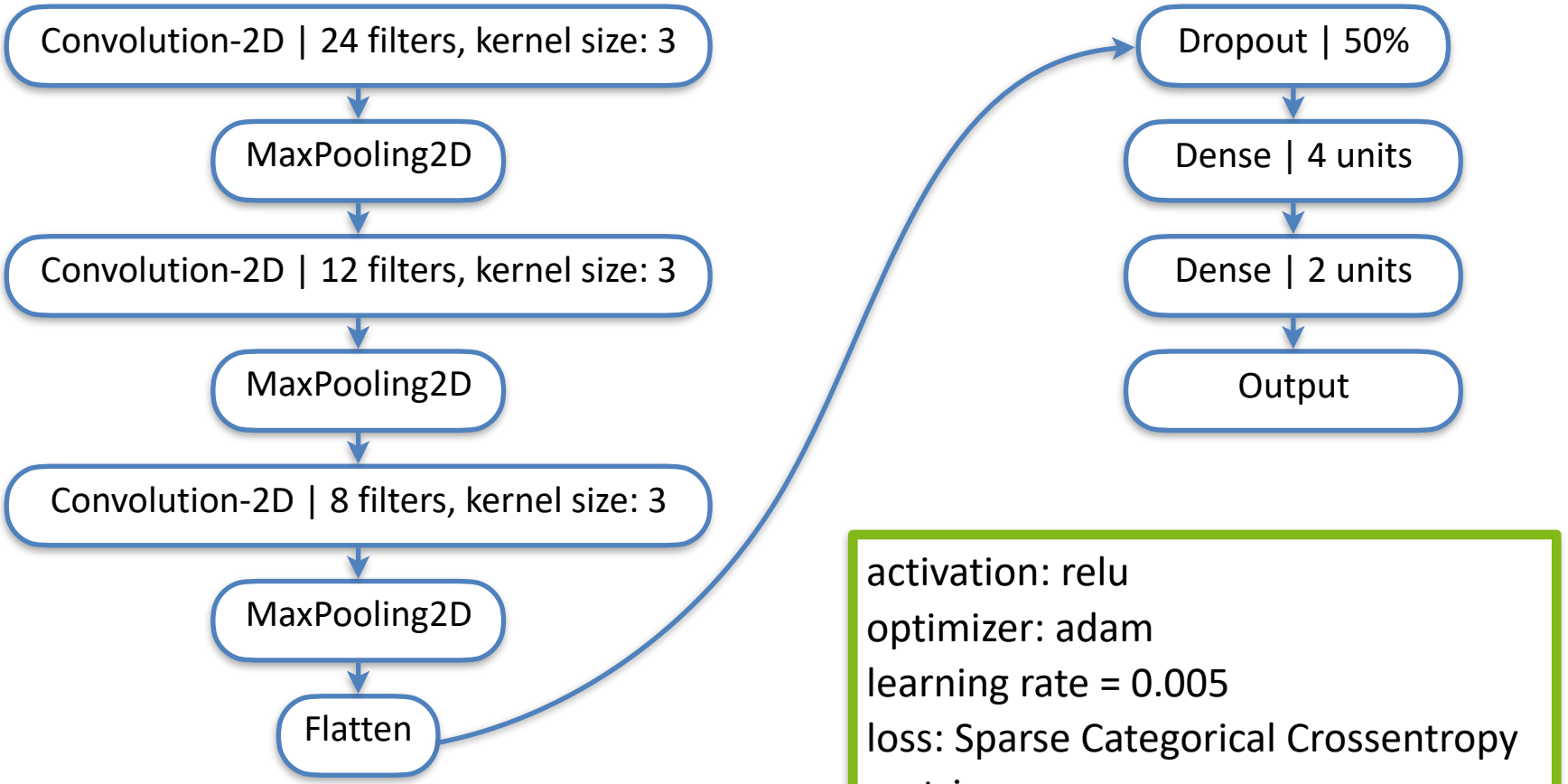
DOI: 10.34740/KAGGLE/DSV/1370629

- First order features:
  - Mean, variance, std. deviation, skewness, kurtosis

- Second order features:
  - contrast, energy, angular second moment (ASM), entropy, homogeneity, dissimilarity, correlation, coarseness

Joel Koch & Bastian Schuchardt
18th of July, 2024

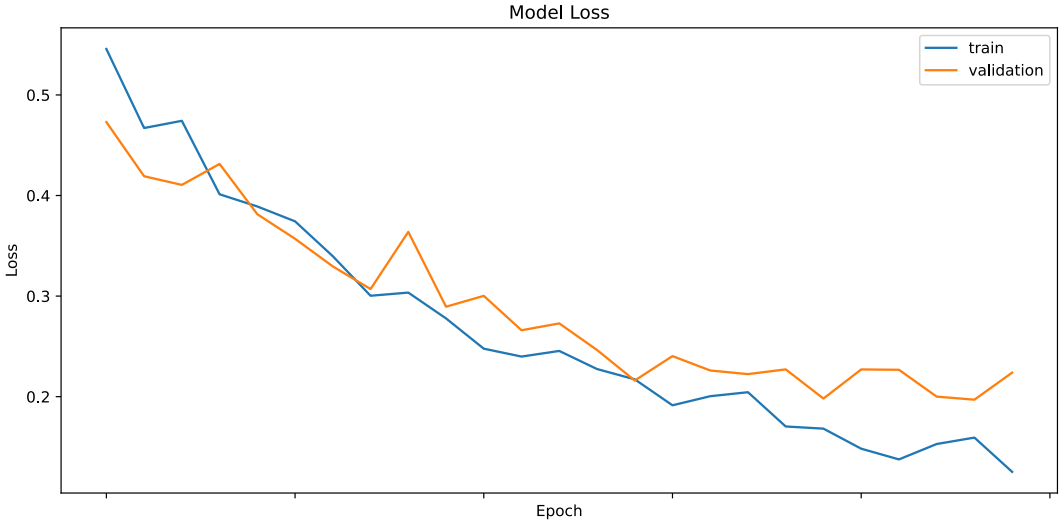
Network architecture

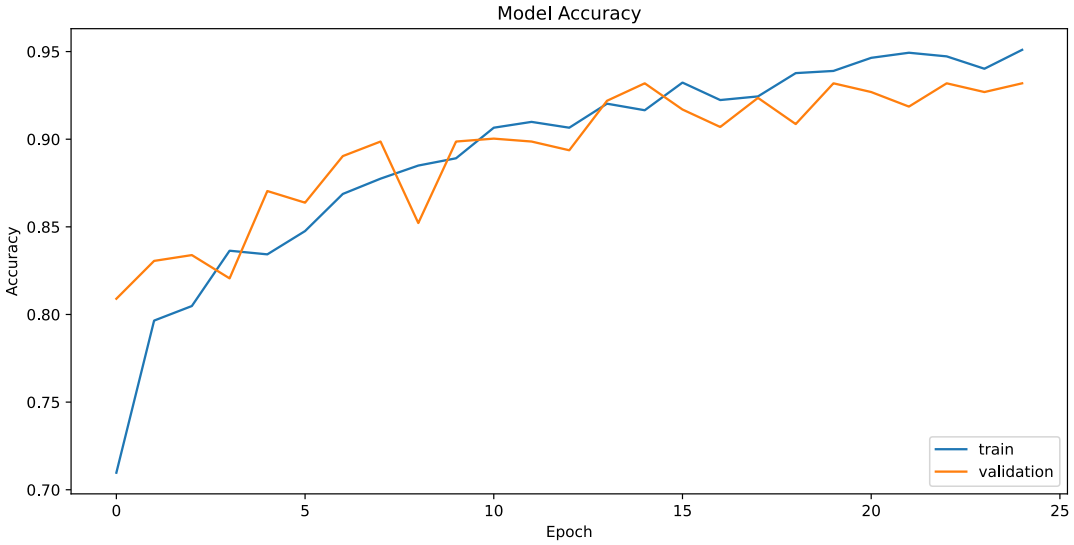


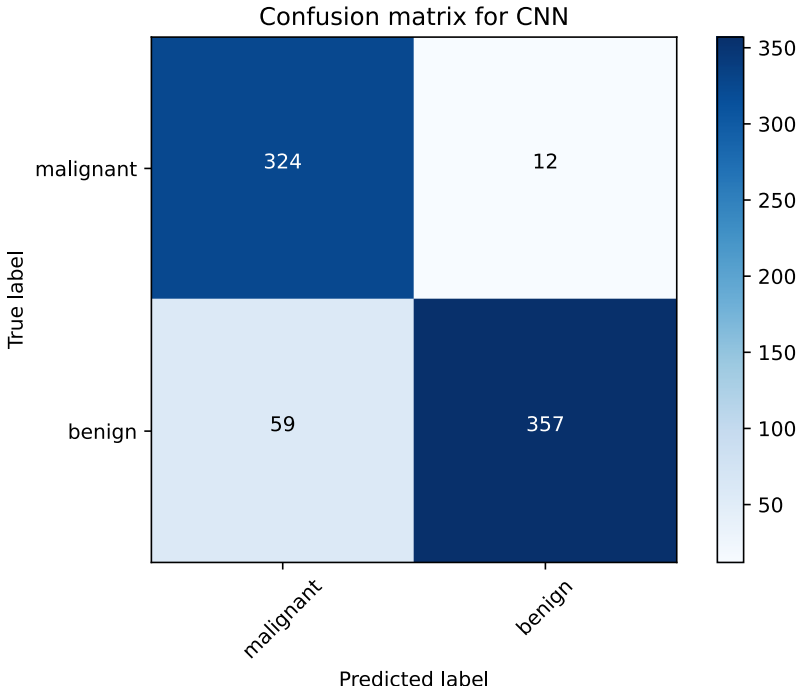
activation: relu
optimizer: adam
learning rate = 0.005
loss: Sparse Categorical Crossentropy
metrics = accuracy

Joel Koch & Bastian Schuchardt
18th of July, 2024

Results







Accuracy: 90.56 %
Mean Squared Error: 0.0944
F1-Score: 0.91

Joel Koch & Bastian Schuchardt

Hyperparameter Optimization / Overtraining checks

- Overtraining was reduced already
  - difference between train & validation was much bigger
- To reduce overtraining:
  - Architecture complexity reduced
  - Introduced Dropout layers
  - Changed learning rate

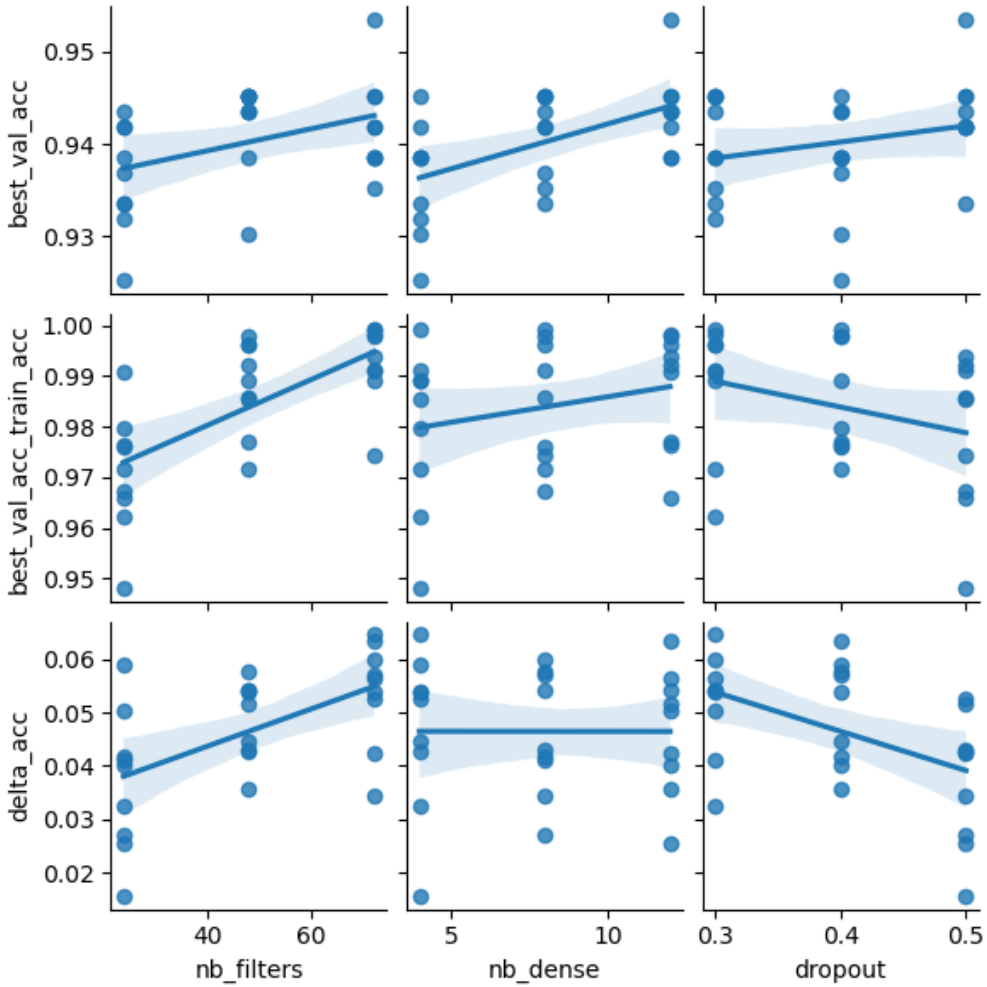
Achieved better accuracy with more overtraining

Joel Koch & Bastian Schuchardt
18th of July, 2024

Hyperparameter Optimization / Overtraining checks

- Filters = [24, 48, 72]
- Dense units = [4, 8, 12]
- Dropout = [.3, .4, .5]

→ bigger Dropout
→ increase dense units
→ check again with filters



Joel Koch & Bastian Schuchardt
18th of July, 2024

Alternative Method

- Comparing image recognition vs. binary classification

	SVM	Logistic Regression	K-NN
Accuracy	98.01%	97.88%	97.61%
Mean Squared Error	0.0199	0.0212	0.0239

Joel Koch & Bastian Schuchardt
18th of July, 2024

Conclusion

- CNN reached good performance with a lot of overtraining
- Reducing overtraining will be the next key task
- Accuracy needs to be increased
  - Run model with more epochs
- Binary classification methods achieve better results than CNN (so far)