Analysis of the Impact of Convolutional Neural Network Parameters on Classification Accuracy in a Small Medical Image Dataset

RESEARCH SEMINAR PRESENTATION

ŽAN PETERNELJ, 89212061

MENTOR: PETER ROGELJ, PhD

Agenda

- Introduction
- Data structure
- Neural network configuration and learning
- Data classification
- Results
- Conclusion

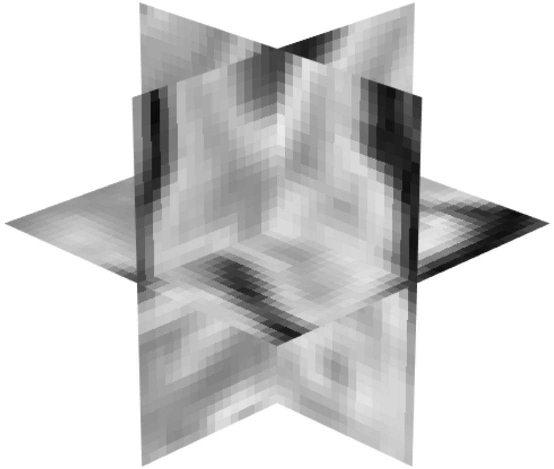
Introduction

- Analysis of impact of individual CNN parameters on classification accuracy between classes.
- Previous research on effect of adjusting the CNN parameters show that significant growth in performance and accuracy can be achieved.
  - K. G. Pasi and S. R. Naik, "Effect of parameter variations on accuracy of Convolutional Neural Network," 2016 International Conference on Computing, Analytics and Security Trends (CAST), 2016.
  - S. Maitra, R. K. Ojha and K. Ghosh, "Impact of Convolutional Neural Network Input Parameters on Classification Performance," 2018 4th International Conference for Convergence in Technology, 2018.
- Extension of previous collaborative research with University of Basel.
  - CNN for classification of brain lesions types, based on the effect of lesion presence on development of multiple sclerosis.
- 8 different CNN architecture configurations, 4 different classifications, 32 models.

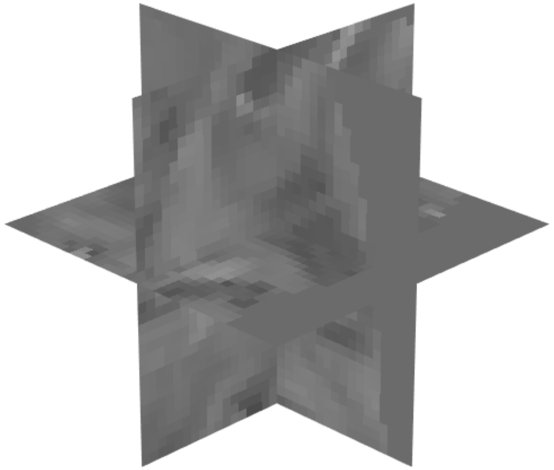
Data structure

- 3D brain lesion patches from MRI images.
- Size of the patches: 35 x 35 x 35 voxels.
- Each patch case is represented using two methods:
  - Fluid attenuated inversion recovery (FLAIR) – Part of almost all protocols for brain imaging.
  - Quantitative susceptibility mapping (QSM) – Used for detecting blood products, calcifications, etc.
- Images of each method are stored as a channel alongside the mask which represents the area of the lesion on the patch.
- Data object consists of four dimensions, where first three dimensions are spatial, and the fourth dimension represents the image channel.

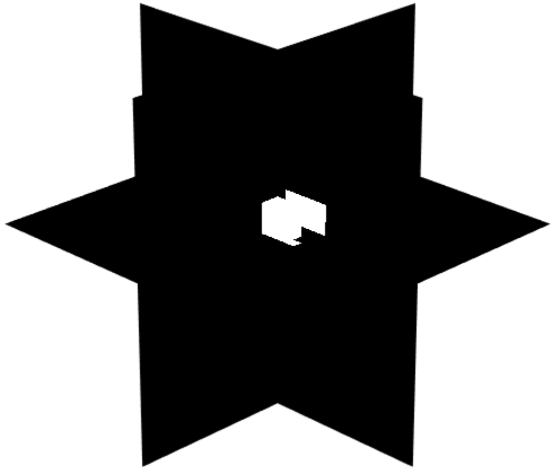
Data structure



Channel 1 - FLAIR



Channel 2 - QSM



Channel 3 - Mask

Data structure

- 6 types of lesions based lesion properties.
- Due to limited number of cases for certain lesion types and lack of training data overall, images were augmented using rotation and mirroring.

#	Lesion type	Definition
1	/	Lesion can not be classified.
2	Isointense	A lesion, but inside and around the lesion the values are the same.
3	Hyperintense rim	The border of the lesion has higher values compared to inside and outside the lesion
4	Hypointense rim	The border of the lesion has lower values compared to inside and outside the lesion.
5	Hyperintense lesion	The value inside the lesion has higher values compared to the surrounding, no specific delineation of the border is present
6	Hypointense lesion	The value inside the lesion has lower values compared to the surrounding, no specific delineation of the border is present

Lesion type	Number of cases
1	3664
2	460
3	214
4	19
5	841
6	71
Sum	5269

Convolutional neural network

- Network architecture for deep learning.
- Learns directly for data.
- Used for finding patterns in images to recognize objects, faces, etc.
- Also used for classifying non-image data such as audio, signal data, etc.
- Medical image classification.

Neural network configuration

- A basic form of a CNN with one input layer, a convolutional layer and an output layer with two classes was used as a baseline architecture.

#	Layer	Parameter
1	3-D Image Input	[35x35x35] x 3
2	Convolution	[5x5x5] x 10
3	Batch Normalization	
4	ReLU	
5	3-D Max Pooling	[2x2x2]
6	Fully Connected	
7	Softmax	
8	Classification Output	

Neural network configuration

- Baseline architecture was then configured using the following parameters:
  - Number of convolutional layers – increased to two and three layers.,
  - size of convolution filters – reduced to [3,3,3] and increased to [7,7,7],
  - number of filters – reduced to 5 and increased to 20 and 30.

#	Architecture	Description
1	baseline	Baseline architecture.
2	conv[3,3,3]	Convolution kernel size decreased to [5,5,5].
3	conv[7,7,7]	Convolution kernel size increased to [7,7,7].
4	filter 5	Number of convolution kernels decreased to 5.
5	filter 20	Number of convolution kernels increased to 20.
6	filter 30	Number of convolution kernels increased to 30.
7	2 conv	Two convolution layers with same parameters as baseline.
8	3 conv	Three convolution layers with same parameters as baseline.

Classification

- Each type of lesion was classified against all the types, using all different architecture configurations.
- Due to uncertain effect on development of disease type 1 and due to low case count type 4 were omitted.
- Classifications:
  - 2 vs Others
  - 3 vs Others
  - 5 vs Others
  - 6 vs Others

Neural network training

- Matlab, version R2022a
- PC, Windows 11, Intel I9-12900K, Nvidia RTX3080 and 32GB RAM.
- Adam optimizer and same training settings were used to train all CNN configurations.

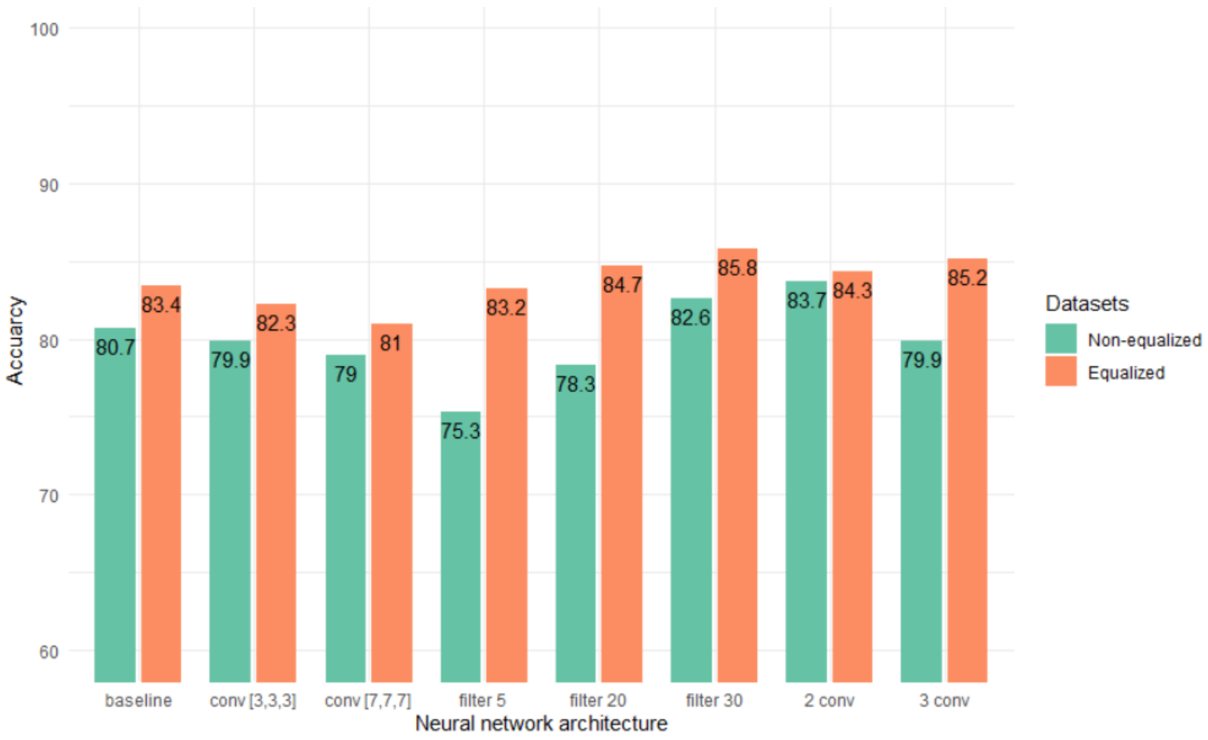
Option	Value
Initial learn rate	1e-3
Max Epochs	10
Shuffle	every-epoch
Learn Rate Schedule	piecewise
Execution Environment	GPU
Mini Batch Size	50
L2 Regularization	0.0005

Neural network training

- Initial data was split into train dataset and test dataset in the ratio of 70% to 30%.
- Only train dataset was augmented.
- In total 32 different models were trained.
- Each model was tested using two different test datasets:
  - Equalized – By decreasing number of cases to match the class with the smaller number of cases.
  - Non-equalized – Imbalance between classes was preserved.

Results

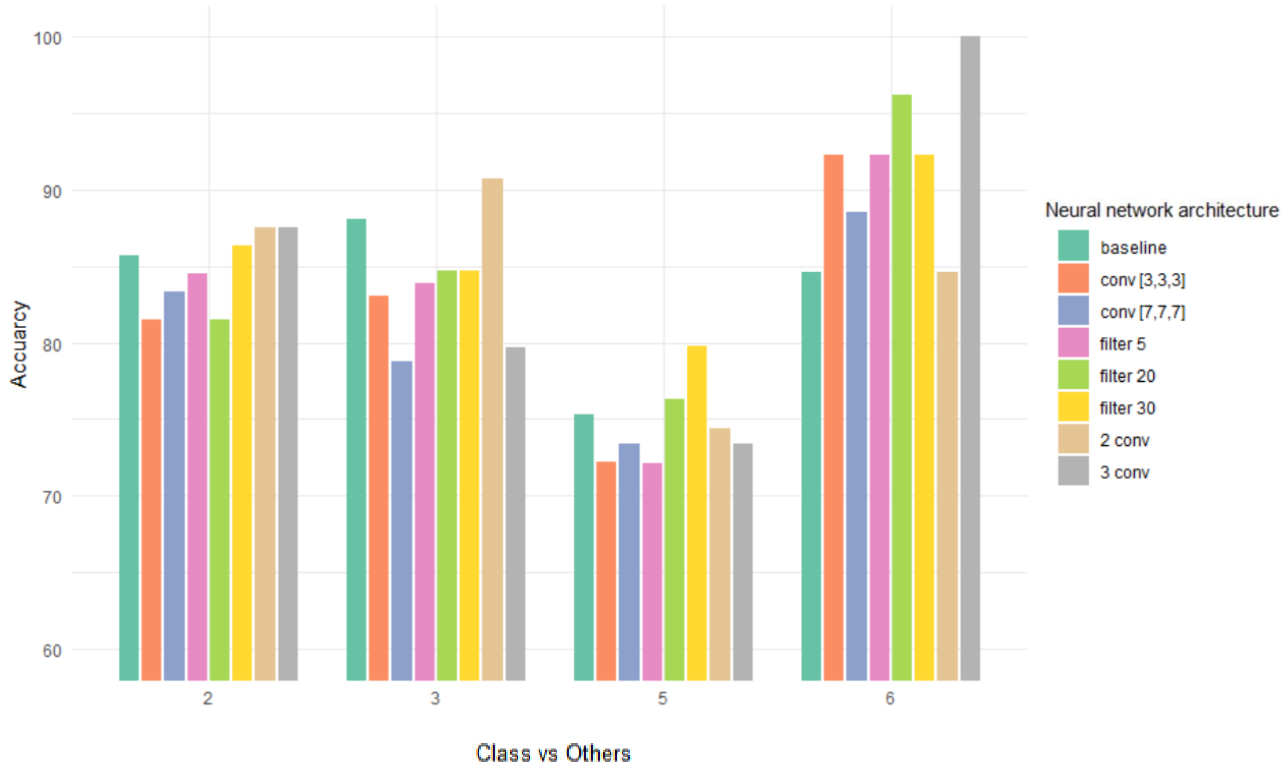
- The non-equalized dataset achieves on average 3.9% lower overall classification accuracy.



Results

- Equalized test dataset:
  - Highest average accuracy: *filter 30*
  - Lowest average accuracy: *conv[7,7,7]*

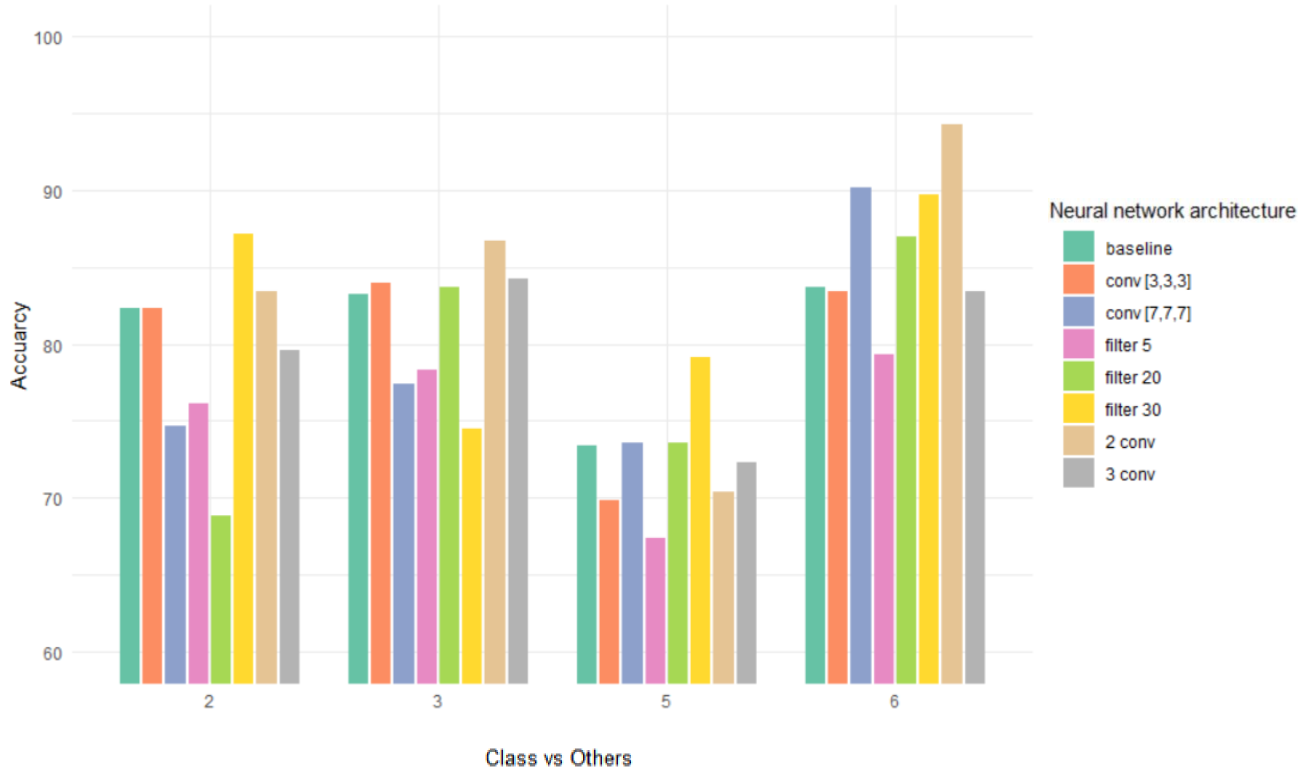
Class	Best	Worst
2 vs Others	2 conv 3 conv	conv[3,3,3] filter 20
3 vs Others	2 conv	conv[7,7,7]
5 vs Others	filter 30	conv[3,3,3] filter 5
6 vs Others	3 conv	baseline 2 conv



Results

- Non-equalized test dataset:
  - Highest average accuracy: *2 conv*
  - Lowest average accuracy: *filter 5*

Class	Best	Worst
2 vs Others	filter 30	filter 20
3 vs Others	2 conv	filter 30
5 vs Others	filter 30	filter 5
6 vs Others	2 conv	filter 5



Results

- The biggest improvement on average in comparison to baseline architecture was achieved by the filter 30 on the equalized dataset, and by 2 conv on the non-equalized datatset.
- Decreasing or increasing filter sizes on average performed worse that baseline, and never achieved the highest accuracy for any classification.
- Decreasing number of filters achieved on average worse results than baseline architecture, except for class 6 vs Others in the equalized dataset.

Conclusion

- This research analyzed effect of CNN layer parameters on classification accuracy using an unbalanced small size medical dataset.
- Changing the parameters effects the classification accuracy both when compared to baseline architecture, both on the same classification or between different classifications.
- In this specific case using a single architecture for all classifications would not achieve the optimal accuracy, as some types are classified more accurately by certain architectures than the others.
- Future work is required to analyze optimal way of combining different types of architectures used for each type of classification into one combined.