STROKE DETECTION USING CONVOLUTIONAL NEURAL NETWORKS

UniversitätsSpital Zürich

Lisa Herzog[1], Elvis Murina[2], Oliver Dürr[2], Susanne Wegener[3], Beate Sick[1,2]

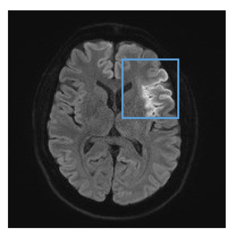
[1]UZH, [2]ZHAW, [3]University Hospital Zurich

Universität Zürich

Abstract

We apply deep learning approaches to magnetic resonance images of stroke and TIA patients. We show how to take the special three-dimensional structure of the data into account in order to improve the model performance. We further utilize MC dropout methods during test time for probabilistic predictions and corresponding confidence measures. For reliable patient-level predictions, we evaluate how to combine the image-based prediction values by considering the uncertainty measurements.

Motivation



Extract predefined features from ROI

Traditional Approach

Prediction model based on extracted features

Deep Learning Approach

Prediction using a CNN considering the whole image

Dropout at test time

Distribution of predictions per image (Uncertainty)

No dropout at test time

One prediction per image

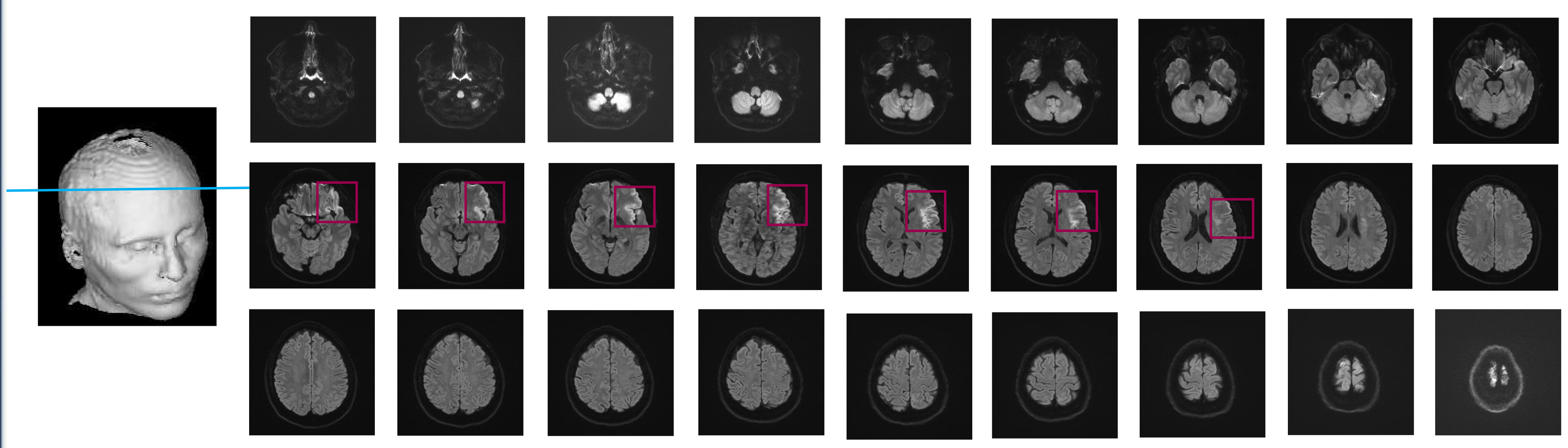
How to combine the predictions ?

- Most of the CNNs used for medical image analysis are still 2D or 2.5D[1] resulting in one prediction per image
- Since there are multiple images per patient, the image-based predictions have to be combined to one patient-level prediction, especially if we want to combine patient data

1. We propose a new way to analyze image volumes in 2D
2. We apply MC-dropout to obtain uncertainties
3. We try to get a reliable patient-level prediction using the image-based predictions and the corresponding uncertainties

Data

- We have n=366 patients (n=120 with stroke and n=140 with TIA)
- We have ~30 DWI slices per patient and each image is expert-labelled as stroke or TIA



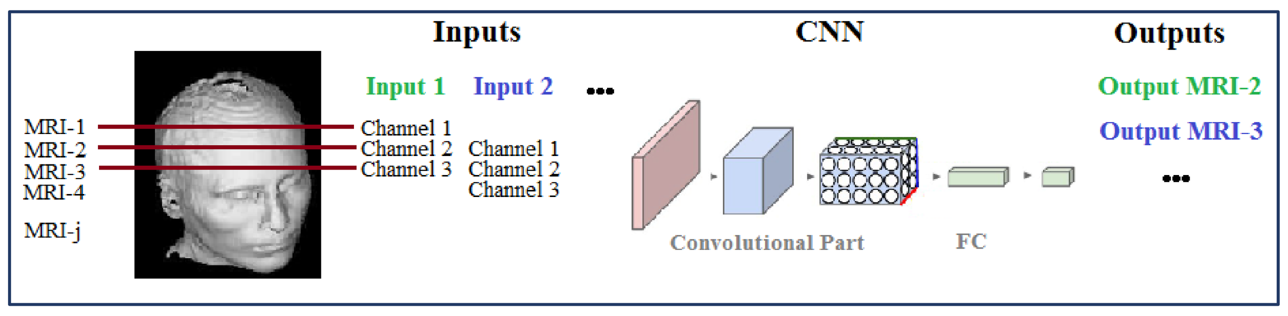
- For image and patient level predictions we used:

	For image level prediction		For patient level predictions		
	train1	valid1	train2	valid2	test
Patients	226	40	40+40	20	40
Images	7559	1128	1128+1120	550	1102

Methods I

Image level predictions

- We developed a VGG-like architecture where we feed in synthetic input images. We include 3 subsequent images within the 3 color channels



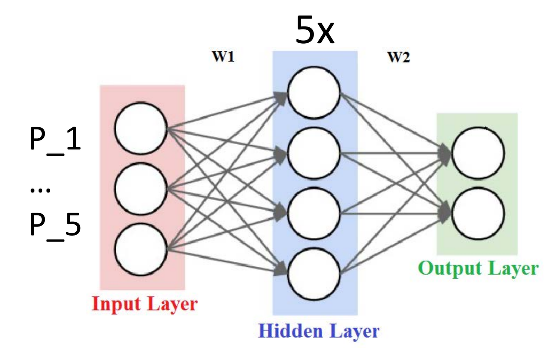
- We trained from scratch using the Adam and applied data augmentation (zooming, rotating, shearing, shifting)
- We used dropout
  1. in the FC part at training time only (Baseline-CNN)
  2. in front of each Convolutional/FC layer at training and test time by what we obtained uncertainty measures: VR*, PE*, MI* (MC-CNN)[2]

Architecture
2 Conv-Stacks*-32
MP*, [Dropout]
2 Conv-Stacks-64
MP, [Dropout]
3 Conv-Stacks-128
MP, [Dropout]
3 Conv-Stacks-256
MP, [Dropout]
3 Conv-Stacks-512
MP, [Dropout]
3 Conv-Stacks-512
Dropout
FC*-4096
BN*, Dropout
FC-2
Softmax

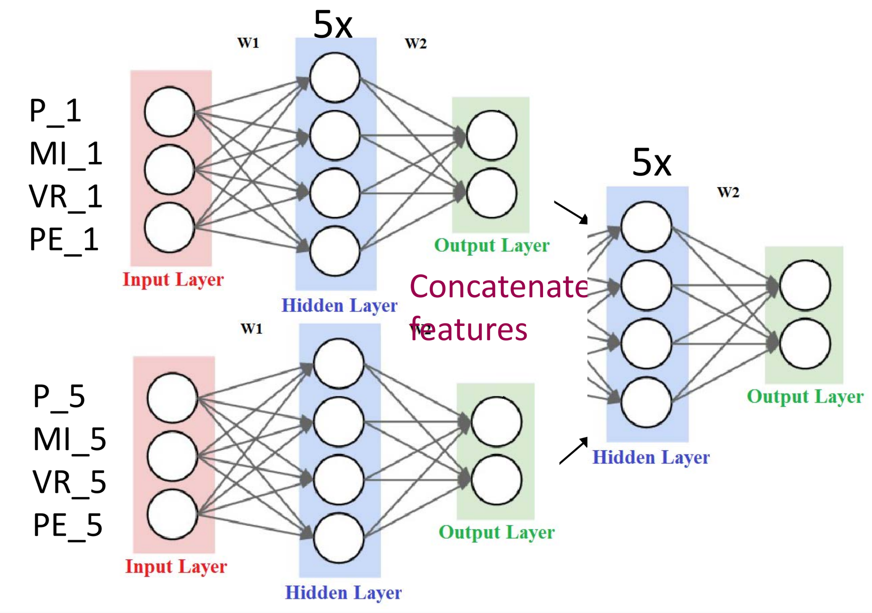
Methods II

Patient level predictions

- **Maximum**: As baseline aggregation method we use the maximum. If one stroke is observed on one image, patient is considered as stroke patient.
- **Neural Network I:** We feed in the predictions (p_1,…,p_5) of the 5 images which are most likely to show a stroke

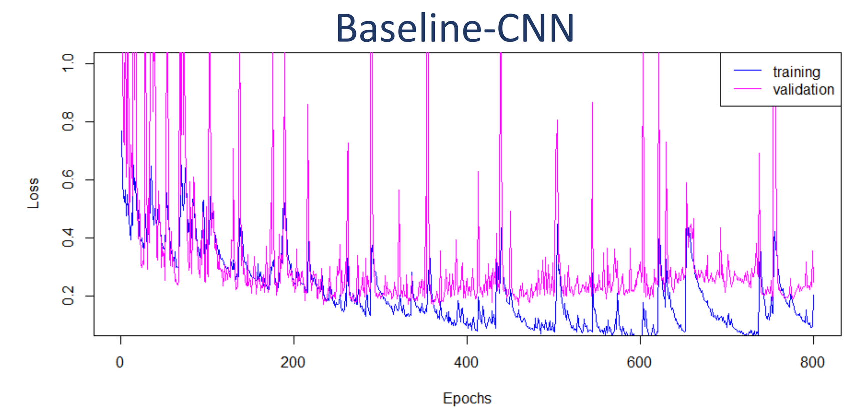


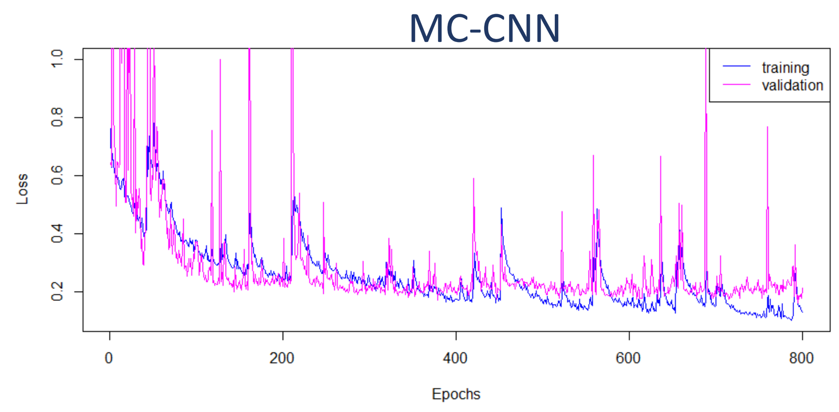
- **Neural Network II:** We feed in the predictions (p_1,…,p_5) and the corresponding uncertainties (PE,MI,VR) of the 5 images which are most likely to show a stroke.



Results

Image level predictions





- **Baseline-CNN:** With the baseline model we reach higher sensitivities

%	Train2	Valid2	Test
Accuracy	94.17	90.18	94.28
Specificity	96.88	92.71	96.73
Sensitivity	79.60	83.55	81.86

- **MC-CNN:** With MC dropout we reach higher accuracies but lower sensitivities

%	Train2	Valid2	Test
Accuracy	94.35	92.36	92.65
Specificity	98.36	96.73	96.63
Sensitivity	72.80	80.92	72.53

Patient level predictions (test)

	Maximum			Neural Net I			Neural Net II		
	Acc	Spec	Sens	Acc	Spec	Sens	Acc	Spec	Sens
BL-CNN	82.50	91.66	78.57	90.00	88.24	91.30	-	-	-
MC-CNN	90.00	82.35	95.65	90.00	93.33	88.00	95.00	93.33	96.00

- Using the maximum seems to be the worst approach. Even if the accuracy is similar to NN I, we detect less strokes (lower sensitivity).
- Taking the uncertainties into account seems to allow the model to better predict the patient outcome. One reason could be that uncertain cases are not considered (high uncertainties).
- On the patient level, one stroke and one TIA patient is misclassified by the best network (NN II).

Conclusion

- We proposed a new way to analyze volumes in a 2D way by generating synthetic input images
- We showed that taking uncertainties into account helps to better predict the patient-level outcome.
- Even if the results are promising, more data has to be collected in order to approve the results.

[1] J. Bernal, Deep Convolutional Neural Networks for brain image analysis on magnetic resonance imaging: a review, 2017

[2] Y.Gal, Uncertainty in Deep Learning, 2016

* Conv-Stack-f = convolutional, batch normalization layer and a ReLU nonlinearity. -f represents the number of filters, MP = Max Pooling, BN = Batch Normalization, FC = Fully connected layer, SD = Standard deviation, VR = Variation Ratio, PE = Predicted Entropy, MI = Mutual Information