Critical Review: *Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks*

Qi Dou, Hao Chen, Lequan Yu, Lei Zhao, Jing Qin, Defeng Wang, Vincent CT Mok, Lin Shi, Pheng-Ann Heng. "Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks" IEEE Transactions on Medical Imaging, Vol. 35, No.5, May 2016

Group 6: Prerna Singh
Mentors: Dr. Jeremias Sulam
April 5, 2020

Project Overview

Clear lipid-exchanged acrylamide-hybridized rigid in-situ-compatible tissue hydrogel (CLARITY) imaging is unique in that it makes brain tissue transparent, allowing for a complete 3-dimensional representation of the brain in a 2-dimensional scan. CLARITY is a critical tool for the connectome project[2], which looks at how neural cells fire together, within a hierarchy of neural activation patterns in the brain. The overall goal of the connectome project is to develop a comprehensive map of neural connections in the brain. Through CLARITY scans, cellular logic and behavior can be understood through cell-type specific excitation throughout the brain.[1] Processing each CLARITY scan usually entails annotating and counting cells by hand. These images can be fairly large, and for instance, in the data set used in this project, there are ~1000 slices per brain. In practice, this can be very time consuming and inefficient, given the amount of information that can be extracted by these scans. Previous methods attempted to solve this problem include filtering, template matching, and blob detection, which had a maximal accuracy of ~59%, which is not very clinically useful.

The project objective is to develop a robust 3-dimensional convolutional neural network (CNN) that can predict, with increased precision and accuracy when compared to other models, how many fluorescent neurons are present within a section of a brain imaged with CLARITY. Development of a 3-D CNN to count the fluorescing cells, will streamline the processing of CLARITY scans, allowing for critical information from these scans to be available almost immediately.

Paper Summary and Background

The authors of the paper "Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks" developed and validated a cascaded framework to automatically detect cerebral microbleeds in MRI images by utilizing 3d convolutional neural networks. Identifying cerebral microbleeds (CMBs) is clinically significant because CMBs have been shown to aid in the diagnosis of several cerebrovascular and cognitive diseases, but also are present in healthy aging individuals. Currently CMBs are laboriously manually labelled by radiologists, resulting in a high error rate.

The authors designed a 2-step framework that first retrieves candidates with high probabilities of being CMBS. Those candidates are further passed through a 3d CNN discrimination model to distinguish true-positive CMBs from hard mimics.

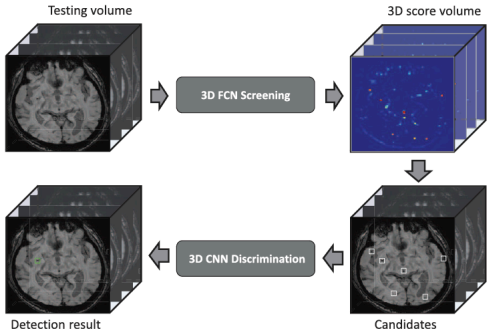

Figure 1: Dou et al. An overview of the cascaded framework for CMB detection.

This framework resulted in a sensitivity rate of 93.16% and precision of 44.31% with an average of 2.74 false positives per CMB, far outperforming previous methods using feature detection or 2D CNNs.

Selection Motivation

In this paper, Dou aimed to solve the same fundamental problem as I aim to: to detect a certain biomarker in volumetric clinical data. CMBs and fluorescent neurons look very similar in these images: rounded hypo/hyper intensities of a small similar size. Further, both CMBs and fluorescent cells are biomarkers found in brains, making these detection tasks very similar.

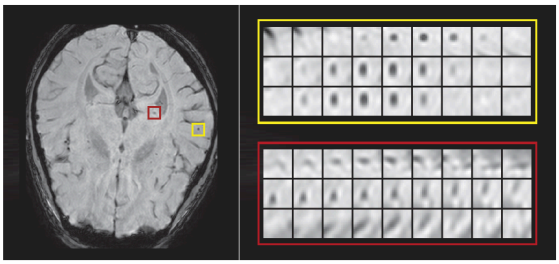

Figure 2: Dou et al. Illustration of a CMB and a CMB mimic denoted with yellow and red rectangles, respectively. In each of the big rectangle from top to down, the rows demonstrate adjacent slices in axial, sagittal and coronal planes.

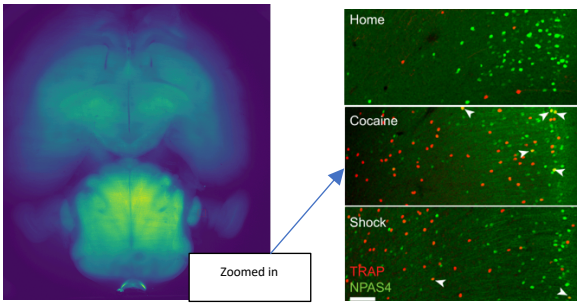

Figure 3: Left: Image from the dataset used in my project. Right: Ye et al. This image is a CLARITY scan zoomed in and the fluorescent cells (round green markers) are what are to be detected.

By reading this paper I aimed to better understand the difficulties and considerations when building a 3D CNN, learn from and critique network architecture choices and justification, and analyze various methods of validation for a biomarker detection task.

Technical Methods

Implementation and Architecture Design

This paper implements a two-stage cascaded framework composed a screening stage with a 3D fully convolutional network and a discrimination stage using a 3D CNN. The input the 3D CNN is the entire volumetric data. The rectifier linear unit (reLU) is used for a non-linear activation function in the C and FC layers. Parameters in the network are tuned using backpropagation with stochastic gradient descent to minimize the cross-entropy loss. Dropout is used to reduce over-fitting and improve the generalization capability of the model. To improve time performance, the typical sliding window strategy is not used because the targets are sparsely distributed. To remedy this, the authors first obtain candidates with a high sensitivity and then perform fine grained discrimination only on those candidates to decrease computational cost.

Screening Stage: Fully Convolutional Network (FCN)

Most traditional fully convolutional layers require that initial inputs have a fixed size, but in this paper, the authors implement a fully convolutional neural network, where the model is transformed to take an

arbitrary-sized input and convolution and max pool kernels sweep over to generate the corresponding sized output.

Layer	Kernel size	Stride	Output size	Feature volumes
Input	-	-	16× 16× 10	1
C1	5× 5× 3	1	12× 12× 8	64
M1	2× 2× 2	2	6× 6× 4	64
C2	3× 3× 3	1	4× 4× 2	64
C3	3× 3× 1	1	2× 2× 2	64
FC1	2× 2× 2	1	1× 1× 1	150
FC2	1× 1× 1	1	1× 1× 1	2

Figure 4: Dou et al. The architecture of the 3D Fully Convolutional Network

The output of the model is a value at each location of score volume that indicates the probability of a CMB. Then the prediction scores can be mapped back to input volume and regions with high probabilities are retrieved as potential candidates.

The model is trained with positive samples from CMB regions. The positive samples are augmented with translation, rotation, and mirroring of positive samples. The model is also trained on negative samples, which are randomly selected non-CMB regions.

Discrimination Stage: 3D Convolutional Neural Network

In the second part of the cascade framework to detect CMBs, a traditional 3D CNN is used to discriminate true positive CMBs from hard mimics. The input to the CNN is 3D blocks centered on screened candidate positions.

Layer	Kernel size	Stride	Output size	Feature volumes
Input	-	-	20× 20× 16	1
C1	7× 7× 5	1	14× 14× 12	32
M1	2× 2× 2	2	7× 7× 6	32
C2	5× 5× 3	1	3× 3× 4	64
FC1	-	-	1× 1× 1	500
FC2	-	-	1× 1× 1	100
FC3	-	-	1× 1× 1	2

Figure 5: Dou et al. The architecture of the 3D CNN Discrimination Model

Because randomly selected non-CMB samples are not strongly representative when the aim is to distinguish true CMBs from close mimics, false positives obtained from the screening stage are taken as negative samples for training the second stage 3D CNN.

Experiments and Results

To validate the cascaded framework, a large data set with 320 SWI images was used. The data set was divided into three sections for training, validation, and testing

Datasets	Stroke		Normal aging		Total	
	Subjects	CMBs	Subjects	CMBs	Subjects	CMBs
Training	91	701	139	223	230	924
Validation	15	81	25	27	40	108
Testing	20	78	30	39	50	117

Figure 6: Dou et al. The details of the datasets.

The ground truth included 1149 radiologist labeled CMBs that were also verified by a neurologist. The metrics calculated induce sensitivity, precision, and average number of false positives per subject. An

optimal threshold of T=.64 was determined for candidates to move to the second CNN, because it yielded best performance on the validation dataset.

A comparison of several methods for CMB screening was also performed. The results are shown below in Figure 7.

Methods	Sensitivity	FP_{avg}	Time per subject (s)
Barnes et al. [15]	85.47%	2548.2	81.46
Chen et al. [18]	90.48%	935.8	**12.00**
3D FCN model	**98.29%**	**282.8**	64.35

Figure 7: Dou et al. Comparison of different CMB Screening Methods

To visualize the outcome of the model, the results of the screening model are projected on the axial and sagittal planes with true CMBs and the score volume generated with the screening model. This method of visualization gives readers a helpful view of how well the model is performing

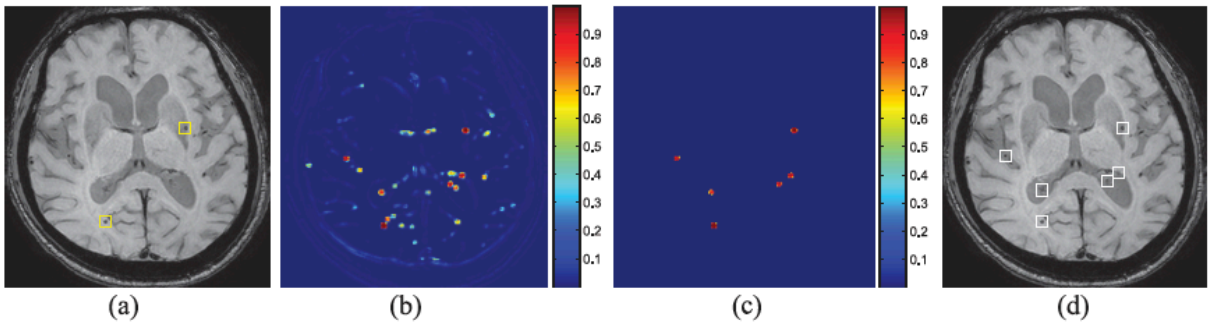

Figure 8: Dou et al. Results of the screening model with score volume projection onto the axial plane. The yellow rectangles are the ground truth and the white rectangles are the retrieved candidates.

In a further comparison of different CMB detection algorithms, the full cascaded framework was compared with other methods. The resulting comparison of FROC curves are displayed in figure 10.

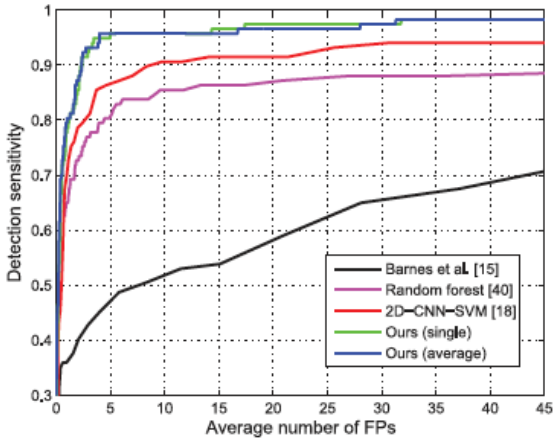

Figure 9: Dou et al. Comparison of FROC curves of different methods. The top two lines are results produced by the framework developed in this paper.

Another interesting validation technique was performed to illustrate the discrimination capability of models using the features extracted by a 2D-CNN-SVM and the 3D CNN discrimination model trained in this paper. The results are shown in figure 11, where we can see that the CMB and non-CMB samples are distinctly separated based on features extracted with the 3DD CNN but embedding of the 2D-CNN-SVM does not show a clear partition boundary.

Figure 10: Dou et al. Feature embedding from 2D-CNN-SVM (LEFT) and the proposed 3D CNN (RIGHT). The red and blue colors represent the CMBs and non-CMBs respectively.

Personal Critique

Positives

Approach

The researchers do a phenomenal job explaining their justification and process for choosing the two-stage cascade framework, mainly to reduce computing time in a sliding window 3D CNN. They further explain the architecture and technical approach of building and training their networks very well at a level that people new to the field can understand, while also explaining higher-level technicalities with detail. Further, the authors do not take for granted the robustness of CNN models, as they test many different input sizes into their CNNs and input parameters. Also, they effectively reduce over-fitting by implementing a dropout layer for overall generalizability.

Training and Testing Methods

This paper demonstrated countless testing and validation methods to analyze the effectiveness of their model. Each method was explained and justified, and all of the figures were helpful and well thought out.

Replicability

The authors have all of their code and data uploaded to a website, to decrease implementation time for others for a similar 3D CNN framework. This also gives replicability to their study.

Areas of Improvement

Fully Convolutional Network

In explaining the Fully Convolutional network, the authors only explain that the FCN successfully eliminates a large number of redundant convolutions, but don't explain why. They only give the mathematical transformation behind it with no connection as to how the network could support an arbitrarily size input with the transformation, which makes it unclear.

Training Data Selection

This paper has two major faults in its training data selection. The paper only trains using scans from a single scanning machine in a single hospital. The training data also only consisted of SWI scans of patients with strokes and aging patients, limiting the generalizability of the network to those sects of patients on a single MRI machine.

Takeaways

- Learned standard methods of validation of a 3D CNN framework and how to present results of validation in a paper
- Understood how to build 3D CNNs, and why certain architectures are chosen
- Learned how to solve common 3D CNN pitfalls, and ways to reduce computational time

Future Steps

- Integrate multi-scale/size information during feature representation phase into the 3D models
- Apply 3D CNN framework to other medical detection tasks

Conclusion

Overall, I learned immensely from this paper on how to build and improve my current 3D CNN architecture to produce potentially improved and quicker performance in biomarker detection in brain images. The paper provided a concise yet informative lens into the approach of building such a model and verification of their results.

References

1. L. Ye, W. E. Allen, K. R. Thompson, Q. Tian, B. Hsueh, C. Ramakrishnan, A.-C. Wang, J. H. Jennings, A. Adhikari, C. H. Halpern, I. B. Witten, A. L. Barth, L. Luo, J. A. Mcnab, and K. Deisseroth, “Wiring and Molecular Features of Prefrontal Ensembles Representing Distinct Experiences,” Cell, vol. 165, no. 7, pp. 1776–1788, 2016.
2. “NIH Launches the Human Connectome Project to Unravel the Brains Connections,” PsycEXTRA Dataset, 2009.
3. Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V. C. Mok, L. Shi, and P.-A. Heng, “Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks,” IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1182–1195, 2016.