Detection of Hyperperfusion on Arterial Spin Labeling using Deep Learning

Nicholas Vincent, **Noah Stier**, **Songlin Yu**, **David S. Liebeskind**, **Danny JJ Wang**, and **Fabien Scalzo**

Neurovascular Imaging Research Core, Department of Neurology, University of California, Los Angeles (UCLA)

Abstract

Hyperperfusion detected on arterial spin labeling (ASL) images acquired after acute stroke onset has been shown to correlate with development of subsequent intracerebral hemorrhage. We present in this study a quantitative hyperperfusion detection model that can provide an objective decision support for the interpretation of ASL cerebral blood flow (CBF) maps and rapidly delineate hyperperfusion regions. The detection problem is solved using Deep Learning such that the model relates ASL image patches to the corresponding label (normal or hyperperfused). Our method takes into account the regional intensity values of contralateral hemisphere during the labeling of a pixel. Each input vector is associated to a label corresponding to the presence of hyperperfusion that was manually established by a clinical researcher in Neurology. When compared to the manually established hyperperfusion, the predicted maps reached an accuracy of $97.45 \pm 2.49\%$ after crossvalidation. Pattern recognition based on deep learning can provide an accurate and objective measure of hyperperfusion on ASL CBF images and could therefore improve the detection of hemorrhagic transformation in acute stroke patients.

I. Introduction

Acute ischemic stroke occurs when a blood vessel to the brain becomes obstructed. It is recognized as a complex heterogeneous process that remains only partially understood. Researchers have attempted to shed light on its complex pathophysiology with various techniques such as molecular biology, genomics, and imaging. Currently, magnetic resonance imaging (MRI) plays a major role in the diagnosis of acute ischemic stroke [21] as it can provide location of the clot, extent of lesions, and maps of tissue at risk that is still viable. Yet, despite the wealth of information extracted from MRI, the information is mostly used qualitatively by visual review. In this paper, we present a tool to automatically delineate likelihood maps of hyperperfusion which is associated to subsequent hemorrhagic transformation (HT); a major risk and a potentially life-threatening complication in patients who receive reperfusion therapy.

Although they have been proven efficient, current treatments of stroke using mechanical clot-retrieval are associated with complications [5]: hemorrhagic transformation (HT), arterial dissection, in situ thrombosis, emboli, vasospasm, perforation, and reperfusion injury. HT is the most common complication and can be life-threatening. It is associated with changes in bloodbrain barrier (BBB) permeability and loss of cerebral blood flow

(CBF) autoregulation, both of which can be observed through post-treatment hyperperfusion. Imaging studies based on PET and MRI suggest that hyperperfused tissue may experience metabolic failures and tend to develop infarction [28]. However, the automatic detection of hyperperfusion is beyond current methods as threshold based methods on CBF maps are greatly affected by noise.

Beyond hyperperfusion, other markers have provided insights regarding potential risks for HT. For example, it has been shown that the DWI/PWI volume is linked to an increased risk of HT after endovascular therapy [6]. Other studies have found a variety of imaging features correlated with an increased risk of HT, such as: leukoaraiosis [14], prior cerebral microbleeds visualized with T2*-weighted MRI sequences [8], early parenchymal enhancement [9], and early colony stimulating factor hyperintensity [10]. In addition, specific permeability parameters [1], [27], [25], [2], [18] derived from perfusion-weighted images (PWI) have been shown to provide distinctive markers to identify patients with an increased risk of HT.

In recent decades, improvements in computer hardware and interest in big data have led to advancements in machine learning. One sub-field of machine learning that holds immense promise for biomedical imaging applications is Deep learning. It is now well established as an effective method of pattern recognition and has been applied to a wide variety of problems, including handwritten character recognition [12], face detection [15], anatomical classification [17] and speech recognition [11]. These systems could provide valuable inputs to physicians in terms of computer-aided diagnosis, image segmentation, image annotation, image registration, and multimodal image analysis.

This paper introduces a machine learning model for the automatic detection of hyperperfusion on ASL CBF maps. A key property of the model is to use the matched contralateral ASL CBF to predict the likelihood of hyperperfusion at a target location. This is similar to the methodology used by stroke neurologists for visual examination of ASL CBF maps. In addition, the model attempts to capture the regional distribution that characterize truly hyperperfused ASL CBF using a convolutional neural network.

II. Methods

A. Dataset

1) Patient Selection—MRI data was collected from patients identified with symptoms of ischemic stroke and admitted at the University of California-Los Angeles Medical Center from May 2010 to September 2013. The use of these data was approved by the local Institutional Review Board (IRB) and was introduced in a previous study [28]. Inclusion criteria were as follows: (1) acute ischemic lesions occurred within the middle cerebral artery (MCA) distribution on diffusion-weighted imaging (DWI); (2) baseline MRI was performed within 24 hours of symptom onset; (3) ASL imaging was acquired along with routine clinical MRI, and (4) the absence of previous intracranial hemorrhage, brain surgery, or large territorial lesion.

2) Imaging Protocols—Details about imaging protocols can be found in the original paper [28]. In summary, all patients underwent MRI on Siemens 1.5 T Avanto or 3.0 T TIM Trio systems (Erlangen, Germany), using 12 channel head coils. Arterial spin labeling was performed at various time points after stroke onset as part of a routine clinical MRI protocol including DWI, GRE, FLAIR, and perfusion imaging. A 3D GREASE pseudo-continuous ASL pulse sequence was applied with the following parameters: repetition time (TR)/echo time (TE)/label time/ postlabel delay (PLD), 4,000/22/1,500/2,000 ms; field of view, 22 cm; matrix size, 64×64, 26×5 mm slices; GRAPPA factor of 2, 4/8 partial k-space along slice direction with zero-filling for image reconstruction, 30 pairs of label and control images with a scan time of 4 minutes.

B. Pre-processing

1) Cerebral Blood Flow—Data feature extraction from ASL was performed with Interactive Data Language (IDL, Boulder, CO, USA) software programs developed at UCLA. Motion correction was performed on ASL images of each PLD. Pairwise subtraction between label and control images was performed followed by averaging to generate the mean difference image. Quantitative CBF maps were calculated based on a previously published model [26].

2) Atlasing and Manual Annotation—Cerebral blood flow maps, GRE, and FLAIR images were coregistered with DWI for each time point in each subject using SPM8. All coregistered imaging modalities were projected into the Montreal Neurological Institute template using SPM8 and were displayed on the same axial slices for ratings.

A free-form region of interest (ROI) was hand drawn on multislice CBF maps to delineate hyperperfusion areas by two experienced readers independently. Hyperperfusion was defined as patchy areas with visually perceivable increased CBF on ASL maps either within or around the corresponding lesion observed on DWI images when compared with the homologous contralateral hemisphere. After the annotation process, each voxel of the groundtruth is labeled to 1 if it is hyperperfused and 0 if it is normal.

3) Patch Sampling—For training, we exploit a set of ASL CBF images F at onset, and their corresponding label images L. The dataset $\{X, Y\}$ used to train and to evaluate the predictive model is created by extracting local patches [19], [22], [20] of fixed size $w \times l$ among input images with their corresponding label. Each patch $p \in \mathbb{R}^s$ is described by its raw voxel values, yielding an input vector of $s = w \times l$ numerical attributes. Our method extracts a large number of patches at random positions from training images. In practice, given a sampled location $\{i, j, k\}$, we extract a patch p_F in the ASL CBF image at $F(i, j, k)$, a corresponding patch in the mirrored, contralateral hemisphere p'_F in the ASL CBF, and a corresponding label in the label map at $L(i, j, k)$.

For efficient retrieval of similar cuboids across patients [4], [13], it is desirable to obtain patches that are invariant to rotations. Rotation invariance is also useful when considering patches that present partially high CBF patterns but at different directions. If no rotational normalization is performed, the patches would have a different appearance and the model would require additional training examples. For these reasons, the patches are normalized

with respect to the direction θ of the image gradient using a rotation performed with a bilinear interpolation,

$$c_{F'} = \text{imrotate}(c_F, -\theta_F) \qquad \theta_F = \tan^{-1} \frac{L^\sigma_{y,F}(x,y,z)}{L^\sigma_{x,F}(x,y,z)} \qquad (1)$$

where $L^\sigma_{x,F}, L^\sigma_{x,F}$ are the Gaussian derivatives in X and Y directions, computed from the ASL CBF image F,

$$L^\sigma_{x,F} = \frac{\partial}{\partial x}\mathscr{G}_\sigma \otimes F \qquad L^\sigma_{y,F} = \frac{\partial}{\partial y}\mathscr{G}_\sigma \otimes F \qquad (2)$$

where $\mathscr{G}_\sigma$ is a 2D isotropic Gaussian filter with standard deviation $\sigma = 3$ in our experiments. The two patches are merged into a single, multi-modal patch as follows $x = \{c_{F'} + (c_{F'} - c_M)\}$. Each multi-modal patch x is then labeled with the intensity y of the central voxel in the corresponding label image $y = L(i, j, k)$. The data set consists of the set of patches $x \in X$ and their corresponding outputs $y \in Y$ that represent the corresponding label of hyperperfusion.

C. Predictive Model

Deep learning is very promising field that uses machine learning algorithms to model abstractions in large data sets. It holds great promise for the automation of complex data processing [4]. Deep learning is especially applicable to predictive tasks that use 2-dimensional images [3]. Since our goal is to extract data from CBF images that we have generated, we believe deep learning can provide an automated solution to detecting hyperperfusion. Researchers agree that deep learning is a top contender for pattern recognition [23].

To apply deep learning to stroke diagnosis, we use a set of CBF images and associated voxel label (hyperperfusion or no hyperperfusion) to train a network. After running a variable number of training "epochs", the network's parameters are trained so that the network can produce predictive values when it is given newly acquired data. Each epoch consists of a "feed forward" step that involves feeding inputs, followed by a "back-propagation" step that involves adjusting the network's filter parameters. No prior knowledge of the features has been incorporated in the model. The training occurs when the initial filter parameters are adjusted through back-propagation as the network processes labeled inputs. To test the accuracy of the trained network, we provide independent CBF data, and compare the outcomes to the values manually assigned by experts.

Our overall goal with the neural network is to use the predicted values from each patch to construct a full prediction image. We can use our knowledge of the patch sampling process to associate the CNN's outputs with coordinates, and construct a full image. Then the entire image can be interpreted by a clinician, or used for further data-driven analysis. To quantify the error in our predictions, we use an area under the ROC curve (AUROC) approach.

The model is a convolutional neural network (CNN) with six total layers. Our CNN (Figure 2) was implemented using the DeepLearnToolbox [16] in Matlab. A Fast Fourier Transform-based convolution function was used for faster computation on our particular data-set. Instead of treating the CBF values for the two hemispheres are two input layers to the same networks, we took the difference between the CBF from each image and the CBF in the contralateral hemisphere.

The network has an input layer, two convolutional layers, two sampling layers, and an output layer. The input layer can accept varying patch sizes, and the output layer will always produce one binary output per input patch. The convolutional layers outputted 5 maps and 12 maps respectively, but the kernel size was varied to account for different patch sizes. We worked with a variety of input sizes while developing the model, but experimented primarily on 13×13 and 23×23 sized input patches.

Each kernel of the convolutional layer is a filter with varying weights. These filter parameters are initially very small, and change as part of the training process. The units of a convolutional layer's filter are often called "neurons". Through the convolution operation, these layers allow the network to learn features. Ideally, after the network has learned a feature, the kernel weights will be set such that when a new, independent data set is received that has the same features as the training set, the network's parameters will produce the appropriate output. In our case, this was the connection between CBF values and hyperperfusion of a voxel.

The sampling layers perform a "max-pooling" operation after each convolutional layer; the purpose of this to reduce computational size and prevent overfitting. No learning takes place in the sampling layers, but they are important to make the CNN usable for time-sensitive applications. The final output layer actually consists of a 2D map of output probabilities. We can convert these output probabilities into a binary mask that represents presence of hyperperfusion.

The main hyper-parameters we had to vary were the number of epochs (iterations of training without resetting the CNN), the size of the kernels, and the amount of output maps of each layer. These hyper-parameters must be adjusted for different data sets, and currently there aren't standardized or codified recommendations for selecting these hyper-parameters in the Deep learning community. Hyper-parameter selection was performed using a manual approach: we compared accuracy, quality of predictive images, and computational time. Selecting parameters that create long computational times can make it difficult to reproduce results and lower the practical value of the CNN.

D. Experiments

All of our Deep learning experiments described used patches extracted from 2D CBF images. Our experiment was meant to examine how the predictions of our Deep learning network compared with the manual annotations. We wanted to quantify how accurate a Deep learning approach is, and how viable the predictive abilities are for future research and clinical applications.

For our experiment, we used a inputs with a patch size of 13×13. Samples from 65 patients were used for training and testing. These were split into two groups, so that each group could be tested with a CNN trained by the other group (2-fold validation). Each of these test patients had an existing "ground truth" image that were produced by Neurology experts. We directly compared the predictive output to the ground truth images, in addition to using the area under ROC curve as a quantitative measure of success.

We also compared the predictive results we obtained from the CNN with predictive results obtained by simply thresholding the CBF values. We created a binary image by thresholding at 75 percent of the maximum CBF, and used this binary image to try to predict hyperperfusion. There was no Deep learning involved in this secondary comparison prediction. The purpose of this was to verify that it was Deep learning that produced accurate predictive results, and not simply the trends in CBF.

III. Results

Among a total of 361 ASL scans were collected from 221 AIS patients (age = 72 ± 17 years; 45% males) from May 2010 to September 2013, excluding 5 ASL scans rated as nondiagnostic, Hyperperfusion was detected in 76 patients.

We produced predictive images from those scans that display tissue hyperperfusion predictions based on the probabilities determined by the CNN (a subset of 18 of them is shown in Figure 3). By comparing these images with manually annotated images, we obtained a comparison of accuracy using area under ROC curve. Through visual inspection, we can quantitatively analyze the accuracy of the images; this is relevant to the way imaging is used quantitatively in clinical practice. In addition, we can qualitatively assess the results; this approach is relevant to data-driven medical solutions. The AUROC for our CNN gave an accuracy of $97.45\pm2.49\%$. The AUROC for a simple model that converted CBF to a binary image by thresholding at 75 percent of the max CBF was $59.62 \pm 10.70\%$.

IV. Discussion

There are many advantages to Deep learning, and specifically convolutional neural networks. They are relatively fast to develop and use computationally efficient techniques. In particular, having kernel sizes that are smaller than the input layer gives CNNs a unique advantage compared to similar Deep learning approaches [3] that don't have smaller kernel sizes. The networks are modular, easy to implement in a variety of coding languages, and can be customized to do learning on a variety of input types (for example, different parameters obtained from stroke imaging). The main downside is the lack of standardized hyper-parameter selection [24]. This can make it difficult to apply neural networks to new problems without manual manipulation. In the future, a more systematic approach to hyper-parameter selection could improve results. In addition, we could improve the amount of training data, as this is a consistent way to improve neural network performance [24]. Another promising improvement would be the use of 3D patches. There is already evidence of the effectiveness of 3D patches for CNNs that use spatial and temporal information [7]. In future experiments, it could be advantageous to use a larger training set; in this experiment,

we were limited by the availability of quality data sets. Additionally, to properly test the validity of the CNN, we could not include any patient we were testing in the training set. One major improvement for future studies would be to have additional experts produce manual annotations that were used as ground truth. All ground truth was based on manual annotations from two experts; it is possible that the manual annotations could be closer to an actual ground truth if we averaged results from many experts.

V. Conclusion

Overall, while convolutional neural networks can be challenging to design optimally, they have powerful predictive abilities and can be implemented quickly using existing toolboxes and frameworks. We conclude that the use of a 6 layer convolutional neural network was computationally adequate to make detect brain tissue hyperperfusion based on cerebral blood flow. Our results imply that a computationally simple CNN has great potential as a very accurate method to improve stroke diagnosis. This has potential to aid decision making in a clinical setting and provide clinicians with a powerful tool.

References

1. Bang OY, Buck BH, Saver JL, Alger JR, Yoon SR, Starkman S, Ovbiagele B, Kim D, Ali LK, Sanossian N, Jahan R, Duckwiler GR, Vinuela F, Salamon N, Villablanca JP, Liebeskind DS. Prediction of hemorrhagic transformation after recanalization therapy using T2*-permeability magnetic resonance imaging. Ann. Neurol. 2007 Aug; 62(2):170–176. [PubMed: 17683090]

2. Bang OY, et al. Patterns and predictors of blood-brain barrier permeability derangements in acute ischemic stroke. Stroke. 2009 Feb; 40(2):454–461. [PubMed: 19038915]

3. Bengio, Y., Goodfellow, IJ., Courville, A. Deep learning. MIT Press; 2015. book in preparation for

4. Brown M, Szeliski R, Winder S. Multi-image matching using multi-scale oriented patches. CVPR. 2005:510–517.

5. Darkhabani Z, Nguyen T, Lazzaro MA, Zaidat OO, Lynch JR, Fitzsimmons BF, Linfante I. Complications of endovascular therapy for acute ischemic stroke and proposed management approach. Neurology. 2012 Sep; 79(13 Suppl 1):S192–S198. [PubMed: 23008397]

6. Fiehler J, Remmele C, Kucinski T, Rosenkranz M, Thomalla G, Weiller C, Zeumer H, Rother J. Reperfusion after severe local perfusion deficit precedes hemorrhagic transformation: an MRI study in acute stroke patients. Cerebrovasc. Dis. 2005; 19(2):117–124. [PubMed: 15640606]

7. Ji S, Xu W, Yang M, Yu K. 3d convolutional neural networks for human action recognition. IEEE TPAMI. 2013; 35(1):221–231.

8. Kidwell C, Alger J, Saver J, Mattiello J, Woods R, Starkman S, Liebeskind D, Vespa P, Jahan R, Gobin Y, Sayre J, Duckwiler G, Vinuela F. Mr signatures of infarction vs salvageable penumbra in acute human stroke: a preliminary model. Stroke. 2000; 31(1):285.

9. Kim EY, Na DG, Kim SS, Lee KH, Ryoo JW, Kim HK. Prediction of hemorrhagic transformation in acute ischemic stroke: role of diffusion-weighted imaging and early parenchymal enhancement. AJNR. 2005 May; 26(5):1050–1055. [PubMed: 15891158]

10. Latour LL, Kang DW, Ezzeddine MA, Chalela JA, Warach S. Early blood-brain barrier disruption in human focal brain ischemia. Ann. Neurol. 2004 Oct; 56(4):468–477. [PubMed: 15389899]

11. LeCun, Y., Bengio, Y. Convolutional networks for images, speech, and time-series. In: Arbib, MA., editor. The Handbook of Brain Theory and Neural Networks. MIT Press; 1995.

12. LeCun, Y., Jackel, LD., Bottou, L., Brunot, A., Cortes, C., Denker, JS., Drucker, H., Guyon, I., Muller, UA., Sackinger, E., Simard, P., Vapnik, V. Comparison of learning algorithms for handwritten digit recognition. In: Fogelman, F., Gallinari, P., editors. ICANN. Paris: EC2 & Cie; 1995. p. 53-60.

13. Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. IEEE TPAMI. 2005 Oct. 27:1615–1630.

14. Neumann-Haefelin T, Hoelig S, Berkefeld J, Fiehler J, Gass A, Humpich M, Kastrup A, Kucinski T, Lecei O, Liebeskind DS, Rother J, Rosso C, Samson Y, Saver JL, Yan B. Leukoaraiosis is a risk factor for symptomatic intracerebral hemorrhage after thrombolysis for acute stroke. Stroke. 2006 Oct; 37(10):2463–2466. [PubMed: 16931786]

15. Osadchy M, LeCun Y, Miller M. Synergistic face detection and pose estimation with energy-based models. JMLR. 2007 May.8:1197–1215.

16. Palm, RB. Master's thesis. Technical University of Denmark; Prediction as a candidate for learning deep hierarchical models of data. DTU Informatics, reception@imm.dtu.dk, Asmussens Alle, Building 305, DK-2800 Kgs. Lyngby, Denmark, 2012, supervised by Associate Professor Ole Winther, owi@imm.dtu.dk, DTU Informatics, and Morten Mørup, mm@imm.dtu.dk, DTU Informatics

17. Roth HR, Lee CT, Shin H, Seff A, Kim L, Yao J, Lu L, Summers RM. Anatomy-specific classification of medical images using deep convolutional nets. CoRR. 2015 vol. abs/1504.04003.

18. Scalzo F, Alger JR, Hu X, Saver JL, Dani KA, Muir KW, Demchuk AM, Coutts SB, Luby M, Warach S, Liebeskind DS. Multi-center prediction of hemorrhagic transformation in acute ischemic stroke using permeability imaging features. Magn Reson Imaging. 2013 Jul; 31(6):961–969. [PubMed: 23587928]

19. Scalzo F, Hao Q, Alger JR, Hu X, Liebeskind DS. Regional prediction of tissue fate in acute ischemic stroke. Ann Biomed Eng. 2012 Oct; 40(10):2177–2187. [PubMed: 22592291]

20. Scalzo F, Hu X, Liebeskind DS. Neuroimaging. 2012 InTech ch. Tissue Fate Prediction from Regional Imaging Features in Acute Ischemic Stroke.

21. Scalzo F, Nour M, Liebeskind DS. Data science of stroke imaging and enlightenment of the penumbra. Front Neurol. 2015; 6:8. [PubMed: 25798125]

22. Scalzo, F., Hao, Q., Alger, J., Hu, X., Liebeskind, D. Advances in Visual Computing, ser. LNCS. Vol. 6454. Springer Berlin Heidelberg: 2010. Tissue fate prediction in acute ischemic stroke using cuboid models; p. 292-301.

23. Schmidhuber J. Deep learning in neural networks: An overview. Neural Networks. 2015; 61:85–117. published online 2014; based on TR arXiv:1404.7828 [cs.NE]. [PubMed: 25462637]

24. Simard PY, Steinkraus D, Platt JC. J.c.: Best practices for convolutional neural networks applied to visual document analysis. ICDAR. 2003:958–963.

25. Thornhill RE, Chen S, Rammo W, Mikulis DJ, Kassner A. Contrast-enhanced MR imaging in acute ischemic stroke: T2* measures of blood-brain barrier permeability and their relationship to T1 estimates and hemorrhagic transformation. AJNR Am J Neuroradiol. 2010 Jun; 31(6):1015–1022. [PubMed: 20190209]

26. Wang DJ, Alger JR, Qiao JX, Hao Q, Hou S, Fiaz R, Gunther M, Pope WB, Saver JL, Salamon N, Liebeskind DS. The value of arterial spin-labeled perfusion imaging in acute ischemic stroke: comparison with dynamic susceptibility contrast-enhanced MRI. Stroke. 2012 Apr; 43(4):1018–1024. [PubMed: 22328551]

27. Wu S, Thornhill RE, Chen S, Rammo W, Mikulis DJ, Kassner A. Relative recirculation: a fast, model-free surrogate for the measurement of blood-brain barrier permeability and the prediction of hemorrhagic transformation in acute ischemic stroke. Invest Radiol. 2009 Oct; 44(10):662–668. [PubMed: 19724234]

28. Yu S, Liebeskind DS, Dua S, Wilhalme H, Elashoff D, Qiao XJ, Alger JR, Sanossian N, Starkman S, Ali LK, Scalzo F, Lou X, Yoo B, Saver JL, Salamon N, Wang DJ. Postischemic hyperperfusion on arterial spin labeled perfusion MRI is linked to hemorrhagic transformation in stroke. J. Cereb. Blood Flow Metab. 2015 Apr; 35(4):630–637. [PubMed: 25564233]
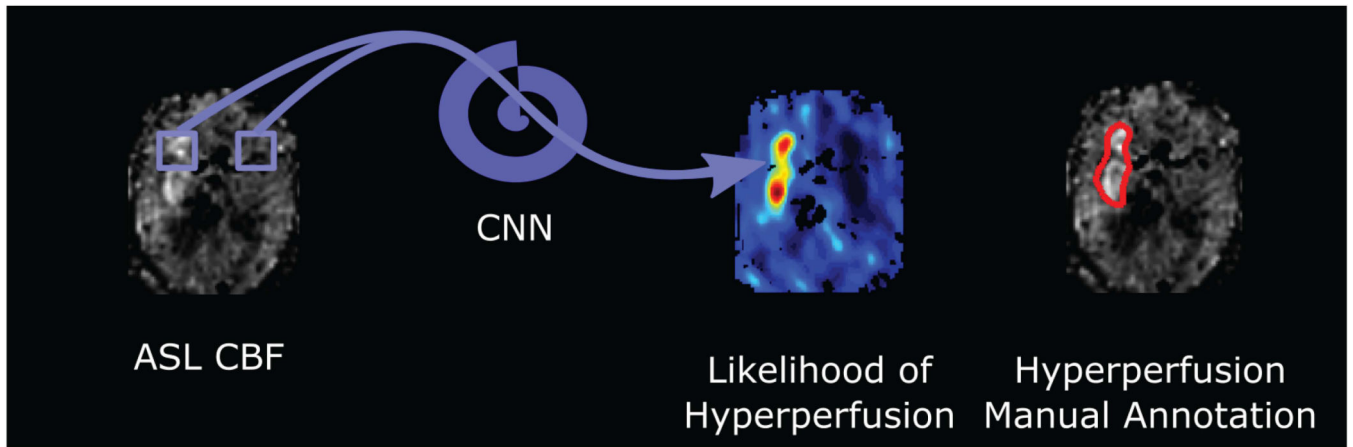
Fig. 1.
Illustration of a pair of patches extracted at mirrored locations on an ASL CBF images and used by the CNN model to predict likelihood of hyperperfusion.
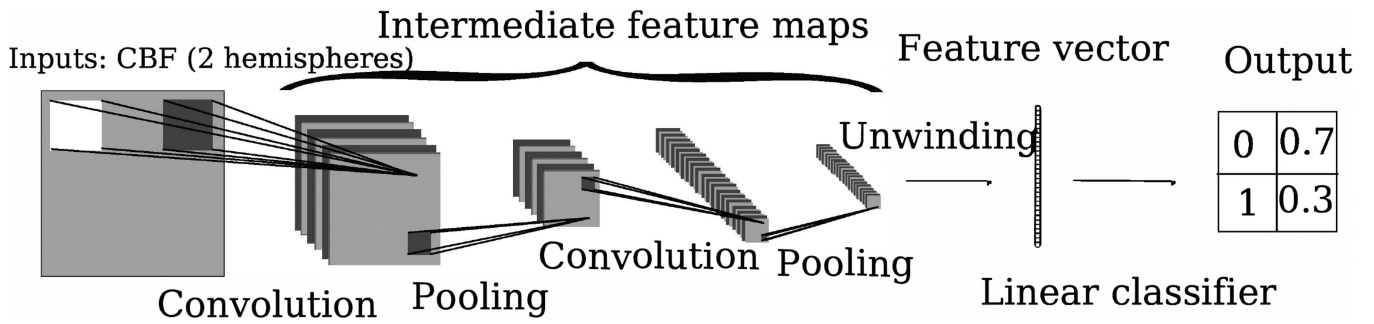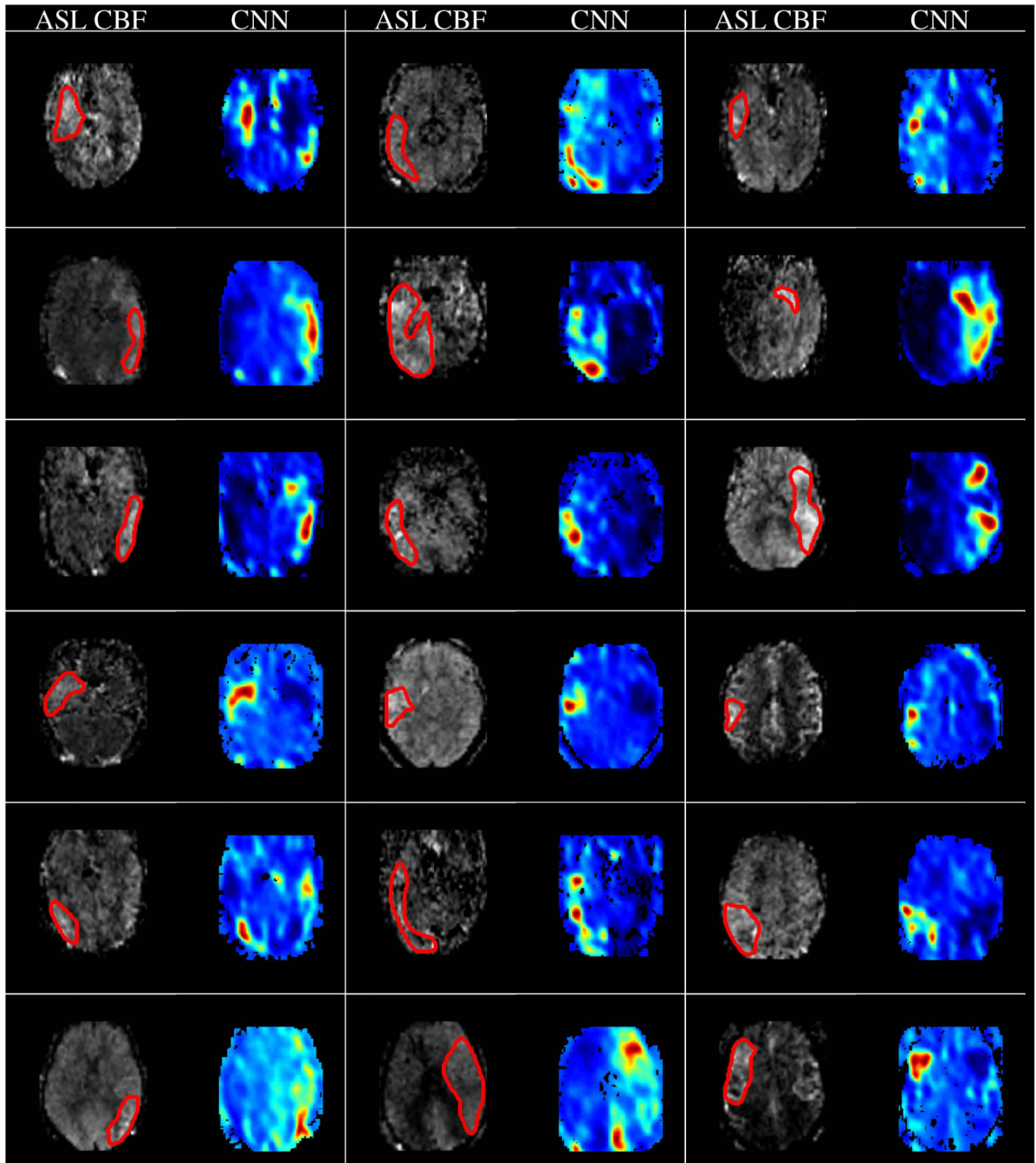
Fig. 2.
Illustration of the convolutional neural network we used. It shows the relationships between layers of the network. The gray rectangle in the input region represents the main sampling patch, while the white rectangle represents the contralateral patch.

Fig. 3.
Illustration of the source ASL CBF images, the corresponding manual annotation used as groundtruth (red contour), and the likelihood map produced by the CNN.