QEI-Net: A Deep learning-based automated quality evaluation index for ASL CBF Maps

Xavier Beltran Urbano[1,2], Manuel Taso[3], Ilya M. Nasrallah[1], Ze Wang[4], John A. Detre[1,2], Sudipto Dolui[1]

[1]: *Department of Radiology,* [2]: *Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania, USA*
[3]: *Siemens Medical Solutions USA, Malvern, USA*
[4]: *Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine, USA*

Target Audience

Anyone interested in evaluating the quality of Arterial Spin Labeling (ASL) derived cerebral blood flow (CBF) maps, particularly in large research studies.

Purpose

Arterial Spin Labeling (ASL) MRI is widely used to quantify regional cerebral blood flow (CBF)[1,2] because it is noninvasive and can be acquired as part of a multimodal MRI protocol, especially in large-scale research studies. However, ASL has a relatively low signal to noise ratio (SNR) and is prone to artifacts, necessitating robust quality control (QC). Manual QC methods are laborious and subject to bias. We previously proposed an automated quality evaluation index[3] (hereafter referred to as QEI_{basic}), which was based on fitting three pre-determined image derived features on a small set of data obtained with non-background suppressed 2D protocols. While QEI_{basic} showed reliable agreement with manual rating on average, there were disagreements in individual cases and a lack of generalization across different imaging protocols. Here, we proposed QEI-Net, a deep learning (DL) model to derive a QEI (hereafter referred to as QEI_{DL}), which has the potential to provide more generalizable results. We further aimed to use QEI_{DL} to identify locations of artifacts in the ASL CBF map. Compared to QEI_{basic}, the QEI_{DL} was trained on a wider variety of scanning protocols and leveraged the entire image instead of predetermined features.

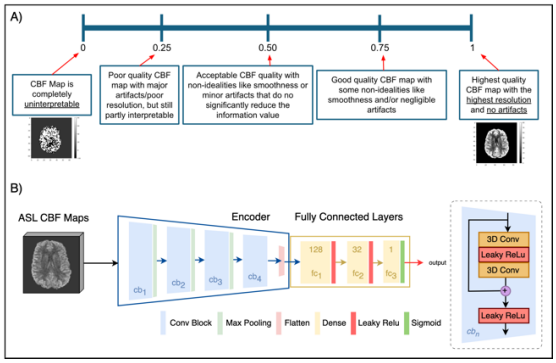

Figure 1: A) Rating strategy guidelines provided to the expert raters to annotate the dataset. B) Overview of the QEI-Net architecture which consists of different convolutional blocks followed by three fully connected layers. The block in the right shows the detailed schematic of each convolutional block, which utilizes residual connections. cb_{1-4} denotes the corresponding convolutional block numbers.

Method

Dataset: QEI_{DL} provides a numerical value between 0 and 1, where a higher value indicates better quality. We utilized N=150 ASL scans acquired with different protocols on Siemens 3T scanners. The data distribution is as follows: N=43 with 2D pulsed ASL (PASL), N=50 with 2D pseudo-continuous ASL (PCASL), N=28 with PCASL labeling and 3D background suppressed (BS) GRASE acquisition, and N=29 with PCASL labeling and 3D BS Stack-of-Spirals acquisition. Three raters with extensive experience working with ASL data visually rated each dataset on a scale between 0 and 1 following some specific guidelines (Figure 1A). Thereafter we divided the data randomly for training and validation (N=120) and testing (N=30). To assess the robustness of the method, an independent fourth rater also rated the test set. *Preprocessing:* The preprocessing pipeline involved calculating CBF from the raw ASL data, registering to the MNI space, down-sampling the images to 64x64x64 voxels, intensity clipping to a range of [-80, 80], and scaling to [-1,1] range. *Deep learning method:* We used a convolutional neural network (CNN) consisting of four convolutional blocks, each with residual connections (Figure 1B). Max pooling layers with a size of 2 were applied after the first three blocks. The network ends with three fully connected layers followed by a sigmoid-activated output neuron. A dropout rate of 20% was applied after the fully connected layers. Training used the Adam optimizer (initial learning rate $1e^{-4}$) with a batch size of 16, and early stopping (patience of 30 epochs) to prevent overfitting. Mean Squared Error (MSE) was used as the loss function of this approach. Artifact location maps were empirically defined as the GradCAM[4] heatmap from the second convolutional layer of the first convolutional block. *Evaluation:* The model was trained using a 5-fold cross-validation strategy and the model-predictions were averaged to provide a more robust prediction on the test set. QEI_{DL} was validated and compared with QEI_{basic} by comparing them with the average manual ratings using Pearson correlation coefficient and the squared errors as the performance indices. Next, we binarized the ratings as unacceptable and acceptable using a threshold of 0.25 and computed the area under the receiver operating characteristic (ROC-AUC) curve of this classification. We also computed the Youden index corresponding to the ROC that can be used as a threshold to discard unacceptable quality CBF maps.

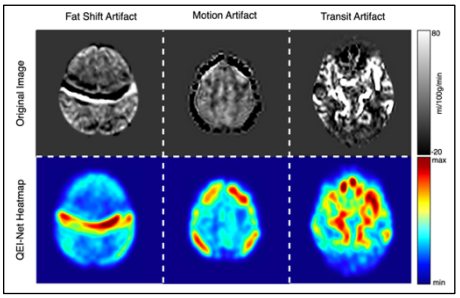

Figure 2: ASL CBF maps contaminated with different types of artifacts and the corresponding heatmaps generated from the QEI-Net showing the location of the artifacts.

Results

The correlation between the ratings of raters 1 and 2 was 0.88, between 1 and 3 was 0.80 and between 2 and 3 was 0.79 ($p<0.0001$ in each case). The correlation between QEI_{DL} and the average manual ratings (R=0.92) was comparable to the inter-rater agreement and was significantly higher (p=0.02) than the correlation between QEI_{basic} and the average manual rating (R=0.81). QEI_{DL} correlation with the fourth rater (R=0.93), whose rating was not used to train the model, was comparable to QEI_{basic} (R=0.88). The squared errors obtained with QEI_{basic} were significantly higher (p=0.002) than those for QEI_{DL}, with mean ± standard deviation values of 0.034 ± 0.039 and 0.009 ± 0.011 for the QEI_{basic} and QEI_{DL}, respectively. The ROC-AUC of QEI_{DL} (0.88) was comparable to that of QEI_{basic} (0.85). Examples of artifactual images, along with the corresponding heatmaps highlighting artifact regions, are shown in Figure 2. Additionally, examples where both QEI_{basic} and QEI_{DL} provided similar values and agreed with the manual ratings (A and B), as well as cases where QEI_{DL} provided better agreement with the manual ratings than QEI_{basic} (C and D), are shown in Figure 3.

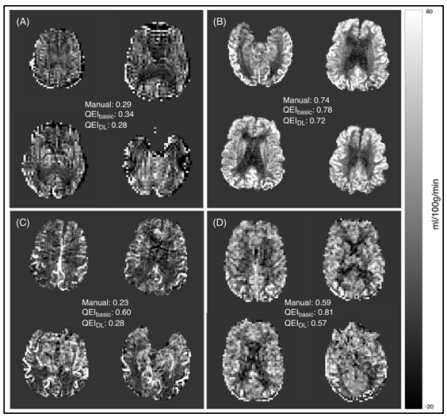

Figure 3: Examples of CBF maps with the corresponding manual ratings and the automated quality evaluation indices derived using QEI_{basic} (Dolui et al.[3]) and QEI_{DL} (proposed). (A) and (B) show examples where both the automated methods agreed with manual ratings while (C) and (D) show examples where the QEI_{DL} provided better agreement than QEI_{basic}.

Discussion and Conclusion

QEI_{DL} demonstrated superior performance compared to our current automated method[3] for predicting the quality of ASL CBF maps, achieving significantly lower squared error metrics and stronger correlations with expert ratings. While QEI_{basic} can suggest the presence of artifacts through low QEI values, it does not specify the location of these artifacts. In contrast, QEI_{DL} provides additional insights by generating heatmaps from its CNN's convolutional layers, offering a visual representation of artifact regions. This capability enhances the interpretability of the results and enables targeted region-of-interest (ROI) analysis, allowing for the identification and exclusion of unreliable CBF regions from statistical analyses. Future work will focus on expanding QEI_{DL}'s training dataset to include ASL scans from a broader range of protocols and scanner vendors to improve generalizability.

References

[1] Detre JA, Leigh JS, Williams DS, Koretsky AP. Perfusion imaging. Magnetic Resonance in Medicine 1992;23(1):37-45. [2] Alsop DC, Detre JA, Golay X, et al. Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia. Magnetic Resonance in Medicine 2015;73(1):102-116. [3] Dolui S, Wang Z, Wolf RL, et al. Automated Quality Evaluation Index for Arterial Spin Labeling Derived Cerebral Blood Flow Maps. Journal of magnetic resonance imaging: JMRI 2024. [4] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. IEEE International Conference on Computer Vision (ICCV). Venice, Italy; 2017. p. 618-626.