Abstract #03986 ISMRM ABSTRACT

ART-Net: A deep learning framework for artifact source classification in ASL CBF Maps

Primary: Contrast Mechanisms - Arterial Spin Labelling 〉 **Secondary:** Analysis Methods - Classification and Prediction 〉 **Keywords:** DEEP LEARNING ARTERIAL SPIN LABELLING (ASL) CEREBRAL BLOOD FLOW (CBF) IMAGE ARTIFACTS ARTIFACT CLASSIFICATION

Xavier Beltran Urbano [1], **Manuel Taso**[2], **Katie R Jobson**[3], **Paul A Yushkevich**[4], **Ilya M Nasrallah**[4], **John Detre**[3,4], **Ze Wang**[5], **Sudipto Dolui**[4]

[1]Department of Bioengineering, University of Pennsylvania, Philadelphia, United States of America
[2]Siemens Medical Solutions USA Inc, Malvern, United States of America
[3]Department of Neurology, University of Pennsylvania, Philadelphia, United States of America
[4]Department of Radiology, University of Pennsylvania, Philadelphia, United States of America
[5]Diagnostic Radiology and Nuclear Medicine, University of Maryland, Baltimore, Baltimore, United States of America

Presenting Author: *Xavier Beltran Urbano (xurbano@seas.upenn.edu)*

Impact

We propose ART-Net to identify sources of ASL artifacts, enabling actionable, source-aware quality control. This framework can enhance scan reliability, reduce data loss in multi-site studies, and provide a generalizable template for artifact classification across imaging modalities.

Synopsis

Motivation: Arterial spin labeling (ASL) MRI quantifies cerebral blood flow (CBF) noninvasively but is susceptible to artifacts. Accurately identifying the artifact source could enable immediate remediation during acquisition and early detection of protocol issues across sites, thereby increasing the proportion of usable data and improving overall quality and reproducibility.

Goals: Develop a deep learning model (ART-Net) to identify the dominant artifact source in ASL CBF maps among four categories: motion, arterial transit-time effects, fat-shift, and z-blurring.

Approach: We trained ART-Net on expert-annotated, multi-protocol ASL scans and evaluated against expert assessments on a held-out test set.

Results: ART-Net achieved excellent classification performance across all artifact types.

Introduction

Arterial spin labeling (ASL) MRI provides noninvasive quantification of cerebral blood flow (CBF)[1,2], a disease-relevant imaging marker linked to functional activity as well as cerebrovascular and neurodegenerative pathophysiology[3,4]. However, its utility can be limited by artifacts that degrade maps in distinct ways, inflating variance and obscuring true effects[5]. Existing automated quality control (QC) methods[6-9], including the recently reported QEI-Net[7], generate global quality scores indicating overall scan quality, but do not identify the causes of degradation, thereby limiting their value for guiding corrective action. Identifying the artifact source during acquisition enables immediate remediation when feasible (e.g., head stabilization or parameter tuning), and supports targeted protocol review when systematic issues arise. In multi-site studies, such feedback can detect configuration errors early and prevent large-scale data loss. In this work, we introduce ART-Net, a deep-learning model that predicts the dominant artifact sources for ASL data.

Methods

We utilized 578 ASL scans acquired using varying protocols on Siemens 3T scanners. Each scan was visually evaluated by an expert rater to determine the dominant artifact source among four common categories: motion, arterial transit-time (ATT), fat-shift, and z-blurring artifacts. The rater also had the option to assign scans to a no-artifact category, but as some level of artifact is typically present in ASL imaging, all scans were ultimately assigned to one artifact class, with artifact severity reflected by a QC score. The no-artifact category was therefore excluded from this study. A summary of dataset sources, acquisition protocols, and expert annotations can be found in Table 1 (Figure 1).

The model was trained and evaluated using stratified 5-fold cross-validation, with class proportions approximately maintained across the training set (N = 467; ~80%), which was divided into training and validation folds, and the held-out test set (N = 111; ~20%). All scans were preprocessed using a uniform pipeline: CBF map derivation from raw ASL data, registration to MNI space, resampling to 64 × 64 × 64 resolution, intensity clipping to [–80, 80], and normalization to [–1, 1] by division by 80.

The proposed ART-Net architecture (Figure 2) comprises four convolutional blocks with residual connections and 2×2×2 max-pooling layers after the first three blocks. The network ends with three fully connected layers with 4 neurons in the last layer, one for each class included in this study. A dropout rate of 20% was applied after the fully connected layers. Training used the Adam optimizer (learning rate = 0.0001, batch size = 32), early stopping (patience = 50 epochs), and learning-rate reduction on plateau (patience = 15 epochs). A weighted categorical cross-entropy loss, with weights inversely proportional to class frequency, was used as the loss function of this approach. Predictions on the held-out test set from the five fold models were averaged for reporting model performance.

For interpretability, Grad-CAM[10] activation maps from the fifth convolutional layer were used to highlight spatial regions contributing most strongly to artifact classification. Performance was evaluated using confusion matrices, one-vs-rest receiver operating characteristic (ROC) curves with area under the curve (AUC), and precision-recall curves with average precision (AP) for each artifact class. Confidence intervals (95%) for ROC-AUC and AP were estimated by bootstrap resampling (n = 10,000). Additional performance metrics included recall, precision, accuracy, F1-score, and specificity for each artifact category.

Results

The confusion matrix (Figure 3A) showed strong diagonal dominance, indicating robust classification across all artifact types. Motion artifacts exhibited the most diverse misclassification pattern (79.1% accuracy), most often being confused with z-blurring (11.6%) or transit-time (9.3%) artifacts. Misclassifications were infrequent and primarily involved cases with multiple co-occurring artifacts where the model's predicted dominant source differed from expert assessment. Representative test cases with correct (left) and incorrect (right) classifications are illustrated in Figure 4.

The one-vs-rest ROC-AUC values further demonstrated excellent discriminative performance: 0.94 (95% CI: 0.89–0.98) for motion, 0.99 (95% CI: 0.97–

1.00) for z-blurring, 0.97 (95% CI: 0.92–0.99) for transit-time, and 0.99 (95% CI: 0.98–1.00) for fat-shift artifacts (Figure 3B).

Grad-CAM–derived heatmaps (Figure 5) effectively localized artifact-related regions. Additional performance metrics, including accuracy, precision, recall, specificity, and F1-score for each artifact class, are summarized in Table 2 (Figure 1).

Discussion and Conclusion

ART-Net is the first deep learning framework to identify the dominant sources of ASL artifacts, achieving performance comparable to expert assessment. By providing source-level classification and Grad-CAM–based localization, ART-Net transforms quality control from a global rating into actionable guidance for acquisition and analysis, with potential for real-time implementation at the scanner console. Future work will expand training across vendors, protocols, and artifact sources, incorporate multi-label classification, and integrate ART-Net with QEI-Net for comprehensive ASL quality evaluation.

Acknowledgements

NIH grants R33AG080518, R01AG081693 and R01EB031080.

References

1. Detre JA, Leigh JS, Williams DS, Koretsky AP. Perfusion imaging. Magn Reson Med. Jan 1992;23(1):37-45. doi:10.1002/mrm.1910230106

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. 2015/01/01 2015;73(1):102-116. doi:https://doi.org/10.1002/mrm.25197

3. Wolk DA, Detre JA. Arterial spin labeling MRI: an emerging biomarker for Alzheimer's disease and other neurodegenerative conditions. Curr Opin Neurol. Aug 2012;25(4):421-8. doi:10.1097/WCO.0b013e328354ff0a

4. Wierenga CE, Hays CC, Zlatar ZZ. Cerebral blood flow measured by arterial spin labeling MRI as a preclinical marker of Alzheimer's disease. J Alzheimers Dis. 2014;42 Suppl 4(Suppl 4):S411-9. doi:10.3233/jad-141467

5. Jaganmohan D, Pan S, Kesavadas C, Thomas B. A pictorial review of brain arterial spin labelling artefacts and their potential remedies in clinical studies. Neuroradiol J. Jun 2021;34(3):154-168. doi:10.1177/1971400920977031

6. Dolui S, Wang Z, Shinohara RT, Wolk DA, Detre JA, for the Alzheimer's Disease Neuroimaging I. Structural Correlation-based Outlier Rejection (SCORE) algorithm for arterial spin labeling time series. Journal of Magnetic Resonance Imaging. 2017/06/01 2017;45(6):1786-1797. doi:https://doi.org/10.1002/jmri.25436

7. Dolui S, Wolk DA, Detre JA. SCRUB: a structural correlation and empirical robust bayesian method for ASL data.

8. Li Y, Dolui S, Xie DF, Wang Z. Priors-guided slice-wise adaptive outlier cleaning for arterial spin labeling perfusion MRI. J Neurosci Methods. Sep 1 2018;307:248-253. doi:10.1016/j.jneumeth.2018.06.007

9. Beltran-Urbano X, Taso M, Nasrallah I, Detre JA, Wang Z, Dolui S. QEI-Net: A Deep learning-based automated quality evaluation index for ASL CBF Maps. Proceedings of the 2025 ISMRM & ISMRT Annual Meeting & Exhibition. 2025:0730.

10. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017:618-626.

11. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. Neurology. Jan 19 2010;74(3):201-9. doi:10.1212/WNL.0b013e3181cb3e25

12. Austin TR, Nasrallah IM, Erus G, et al. Association of Brain Volumes and White Matter Injury With Race, Ethnicity, and Cardiovascular Risk Factors: The Multi-Ethnic Study of Atherosclerosis. J Am Heart Assoc. Apr 5 2022;11(7):e023159. doi:10.1161/jaha.121.023159

13. Dolui S, Detre JA, Gaussoin SA, et al. Association of Intensive vs Standard Blood Pressure Control With Cerebral Blood Flow: Secondary Analysis of the SPRINT MIND Randomized Clinical Trial. JAMA Neurol. Apr 1 2022;79(4):380-389. doi:10.1001/jamaneurol.2022.0074

14. Dolui S, Wang Z, Wang DJJ, et al. Comparison of non-invasive MRI measurements of cerebral blood flow in a large multisite cohort. J Cereb Blood Flow Metab. Jul 2016;36(7):1244-56. doi:10.1177/0271678x16646124

15. Dolui S, Tisdall D, Vidorreta M, et al. Characterizing a perfusion-based periventricular small vessel region of interest. Neuroimage Clin. 2019;23:101897. doi:10.1016/j.nicl.2019.101897

16. Sadaghiani S, Tackett W, Tisdall MD, Detre JA, Dolui S. Reliability of Periventricular White Matter Cerebral Blood Flow using Different ASL protocols.

Figures and Tables

Table1. Information of the different datasets used in this work.

Dataset	Protocol	Sample Size	Motion Artifact	Z-Blurring Artifact	ATT Artifact	Fat Shift Artifact
Alzheimer's Disease Neuroimaging Initiative (ADNI)[11]	2D PASL	70	56	0	14	0
Multi-Ethnic Study of Atherosclerosis (MESA)[12]	3D BS PCASL	122	93	13	4	124
Systolic Pressure Intervention (SPRINT)[13] *trial*	2D PCASL	48	30	0	18	0
Coronary Artery Risk Development in Young Adults (CARDIA)[14]	2D PCASL	36	32	0	4	0
National Alzheimer's Coordinating Center (NACC)[15]	3D BS PCASL	112	22	55	55	0
Vascular Contributions to Cognitive Impairment and Dementia (VCID)[16]	3D BS PCASL	58	2	55	1	0
Total	-	578	235	123	98	124

Abbreviations: PASL: Pulsed ASL, PCASL: Pseudo-Continuous ASL, BS: Background Suppression.

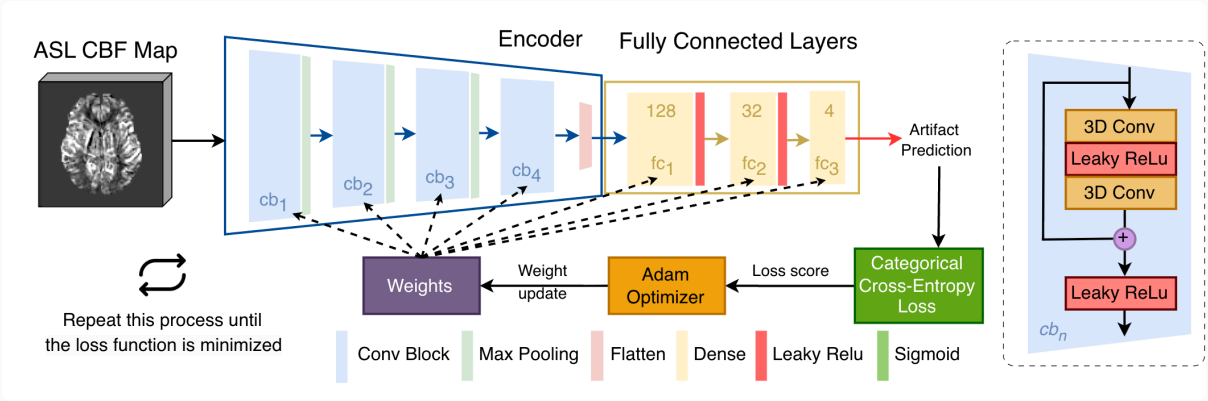
Table 2. Classification performance metrics for each artifact source.

Artifact Source	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Specificity (%)
Motion	88.29	89.47	79.07	83.95	94.12
Z-Blurring	94.59	81.48	95.65	88.00	94.32
ATT	92.79	77.27	85.00	80.95	94.51
Fat Shift	99.10	100.00	96.00	97.96	100.00

Figure 1. Table 1 summarizes the datasets used in this study, including the source, acquisition protocol, and the sample size corresponding to each artifact class. Table 2 shows the performance of ART-Net on the held-out test set (n=111).

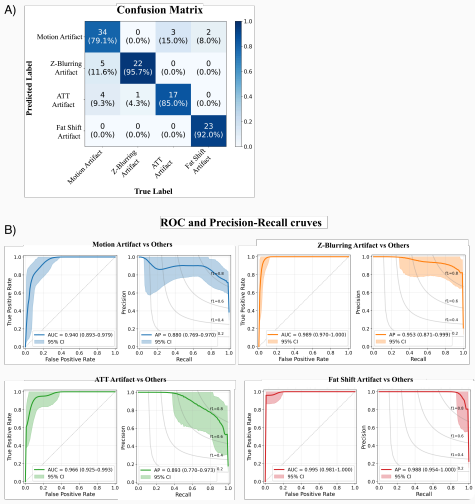
Scan for high-resolution version

Figure 1: *Table 1 summarizes the datasets used in this study, including the source, acquisition protocol, and the sample size corresponding to each artifact class. Table 2 shows the performance of ART-Net on the held-out test set (n=111).*



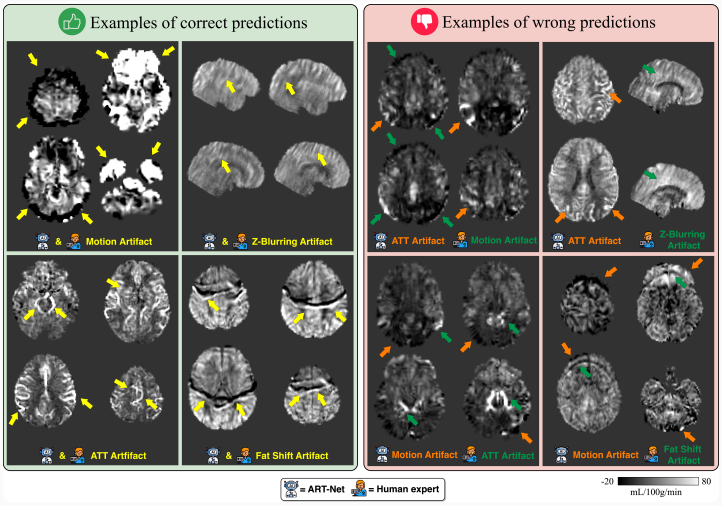
Scan for high-resolution version

Figure 2: *Schematic of the artifact classification approach: The network consists of four convolutional blocks followed by three fully connected layers. The block on the right shows the detailed schematic of each convolutional block, which incorporates residual connections; $cb_{1\text{-}4}$ denote the corresponding convolutional blocks.*



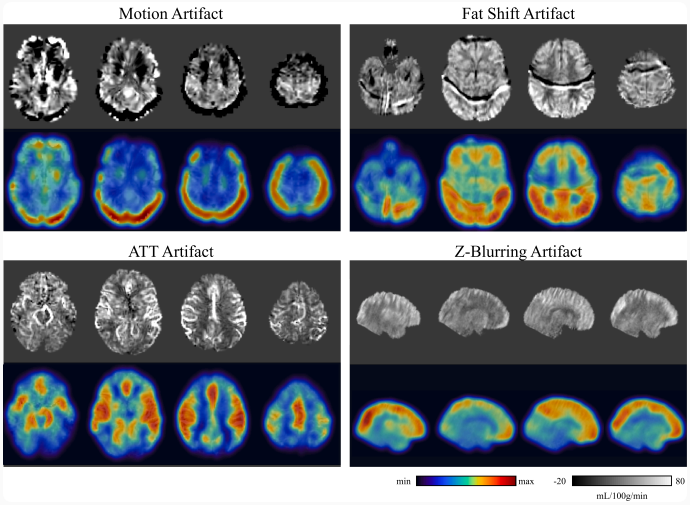
Scan for high-resolution version

Figure 3: *ART-Net performance on the held-out test set (n = 111): A) Confusion matrix showing artifact classification results. True and predicted labels are shown in x and y axes respectively. Cells report counts with column-normalized percentages. B) One-vs-rest classification performance for motion, z-blurring, transit-time, and fat-shift artifacts. For each class, ROC (left) and precision–recall (right) curves are shown with bootstrapped 95% confidence bands, reporting AUC and average precision (AP).*



Scan for high-resolution version

Figure 4: *Examples of ASL CBF maps with correct (left) and incorrect (right) ART-Net predictions from the held-out test set. Yellow arrows indicate artifact-affected regions where ART-Net and human expert agreed on the dominant artifact source. In cases of misclassification, orange arrows highlight regions exhibiting the artifact source predicted by ART-Net, while green arrows indicate regions of the artifact source identified by the expert rater.*



Scan for high-resolution version

Figure 5: *Representative examples from the held-out test set showing ASL CBF maps of each artifact source (top row) and corresponding heatmaps generated from ART-Net (bottom row). The heatmaps highlight regions that contributed most to the artifact classification. The examples demonstrate that these regions spatially overlap with the actual artifact locations.*