



Automated Multiclass Artifact Detection in Diffusion MRI Volumes via 3D Residual Squeeze-and-Excitation Convolutional Neural Networks

Nabil Ettehad^{1*}, Pratik Kashyap², Xuzhe Zhang¹, Yun Wang², David Semanek³, Karan Desai², Jia Guo^{3,4}, Jonathan Posner² and Andrew F. Laine¹

¹ Hefner Biomedical Imaging Laboratory, Department of Biomedical Engineering, Columbia University, New York, NY, United States, ² Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, United States, ³ Department of Psychiatry, Columbia University Medical Center, New York, NY, United States, ⁴ Zuckerman Institute, Columbia University, New York, NY, United States

OPEN ACCESS

Edited by:

Reza Lashgari,
Shahid Beheshti University, Iran

Reviewed by:

Alireza Hajibagheri,
Sonobi, United States
Mohammad Rasool Izadi,
University of Notre Dame,
United States

*Correspondence:

Nabil Ettehad
ne2289@columbia.edu

Specialty section:

This article was submitted to
Brain Imaging and Stimulation,
a section of the journal
Frontiers in Human Neuroscience

Received: 16 February 2022

Accepted: 07 March 2022

Published: 30 March 2022

Citation:

Ettehad N, Kashyap P, Zhang X,
Wang Y, Semanek D, Desai K, Guo J,
Posner J and Laine AF (2022)
Automated Multiclass Artifact
Detection in Diffusion MRI Volumes via
3D Residual Squeeze-and-Excitation
Convolutional Neural Networks.
Front. Hum. Neurosci. 16:877326.
doi: 10.3389/fnhum.2022.877326

Diffusion MRI (dMRI) is widely used to investigate neuronal and structural development of brain. dMRI data is often contaminated with various types of artifacts. Hence, artifact type identification in dMRI volumes is an essential pre-processing step prior to carrying out any further analysis. Manual artifact identification amongst a large pool of dMRI data is a highly labor-intensive task. Previous attempts at automating this process are often limited to a binary classification (“poor” vs. “good” quality) of the dMRI volumes or focus on detecting a single type of artifact (e.g., motion, Eddy currents, etc.). In this work, we propose a deep learning-based automated multiclass artifact classifier for dMRI volumes. Our proposed framework operates in 2 steps. In the first step, the model predicts labels associated with 3D mutually exclusive collectively exhaustive (MECE) sub-volumes or “slabs” extracted from whole dMRI volumes. In the second step, through a voting process, the model outputs the artifact class present in the whole volume under investigation. We used two different datasets for training and evaluating our model. Specifically, we utilized 2,494 poor-quality dMRI volumes from the Adolescent Brain Cognitive Development (ABCD) and 4,226 from the Healthy Brain Network (HBN) dataset. Our results demonstrate accurate multiclass volume-level main artifact type prediction with 96.61 and 97.52% average accuracies on the ABCD and HBN test sets, respectively. Finally, in order to demonstrate the effectiveness of the proposed framework in dMRI pre-processing pipelines, we conducted a proof-of-concept dMRI analysis exploring the relationship between whole-brain fractional anisotropy (FA) and participant age, to test whether the use of our model improves the brain-age association.

Keywords: diffusion MRI, artifacts, quality control, convolutional neural networks, deep learning, medical imaging

INTRODUCTION

Diffusion weighted imaging (DWI; Stejskal and Tanner, 1965; Le Bihan et al., 1986; Huisman, 2003; Baliyan et al., 2016), as well as diffusion tensor imaging (DTI; Basser and Jones, 2002; Alexander et al., 2007), are widely used these days in brain research and clinical neuroimaging (Huisman, 2010). Using dMRI one is able to gain insight into white matter's development based on the different diffusion rates of water across different brain tissues (Hüppi and Dubois, 2006; Ladouceur et al., 2012; Simmonds et al., 2014). Moreover, dMRI provides means for investigation of other abnormal white matter developments such as Schizophrenia (Roalf et al., 2015; Tønnesen et al., 2018) and Alzheimer's disease (Hoy et al., 2017; Lo Buono et al., 2020). dMRI data can also be used for different brain biomarker measurements in large population studies such as the Human Connectome Project (Van Essen et al., 2012). However, dMRI data is often contaminated by various sources of artifacts such as motion, Eddy currents, low signal to noise ratio (SNR), gradient distortions, chemical shift, susceptibility, Gibbs ringing, etc., (Le Bihan et al., 2006; Krupa and Bekiesińska-Figatowska, 2015). Existence of such artifacts in dMRI volumes without any exclusion or correction could bias the results of any subsequent analysis and make their interpretation unreliable (Bammer et al., 2003; Van Dijk et al., 2012; Reuter et al., 2015). Hence quality control and artifact identification in dMRI data is an essential pre-processing step before conducting any analysis.

Current approaches for quality control and artifact identification in dMRI data are mostly performed manually by visual inspection of all volumes (sometimes even all slices) by an expert(s). This process is extremely labor-intensive and suffers from being subjective in nature. Hence there is a need for automated ways of artifact identification that are fast and reliable.

Computerized approaches for quality control and artifact identification of dMRI data are able to alleviate the challenge of manual inspection. Throughout the years, several tools for automated quality control of dMRI data have been proposed such as FSL (Jenkinson et al., 2012; Andersson et al., 2016; Bastiani et al., 2019), DTIPrep (Oguz et al., 2014), DTI studio (Jiang et al., 2006), and TORTOISE (Pierpaoli et al., 2010). More recently, various statistical (Roalf et al., 2016) or Artificial Intelligence (AI) approaches for quality control and artifact detection of dMRI data have been introduced (Iglesias et al., 2017; Kelly et al., 2017; Alfaro-Almagro et al., 2018; Fantini et al., 2018; Graham et al., 2018; Samani et al., 2020; Ahmad et al., 2021; Ettehad et al., 2021). However, most of these approaches either only operate on a binary-level (i.e., distinction of "poor-quality" data from "good-quality" without identifying the specific artifact type) (Samani et al., 2020; Ahmad et al., 2021; Ettehad et al., 2021), or have been designed to only detect a single specific type of artifact (e.g., motion) (Iglesias et al., 2017; Kelly et al., 2017; Fantini et al., 2018). A detailed report on the performance of such tools can be found in Liu et al. (2015), Haddad et al. (2019).

In this work, we propose a deep learning-based framework for automatic multiclass artifact classification in poor-quality dMRI volumes. Unlike previous work, our framework is not customized for a single specific artifact type and takes into account a wider

range of artifacts to classify. In particular, our method classifies four classes of artifacts namely: motion, out of field of view (FOV), low signal to noise ratio (SNR), and MRI miscellaneous artifacts. The MRI miscellaneous artifacts category serves as a control group for the classifier which includes other MRI (as well as dMRI) artifacts that do not necessarily belong to any of the other three classes.

The proposed method operates in two steps. First, the dMRI volumes are partitioned into MECE slabs and fed to a designed convolutional neural network (CNN). The designed CNN then outputs the artifact class labels of the slabs. Second, through a voting process, the slab-level predicted labels are utilized to decide on the final label for the whole dMRI volume (i.e., volume-level artifact label). We tested our method on two separate datasets. Our results demonstrate that the proposed framework can be utilized for fast automatic classification of the four categories of artifacts considered here. Moreover, the extended validation analysis calibrating FA-age correlation demonstrates an improvement using the models' predicted artifact labels. These results together suggest that the model can be utilized in dMRI pre-processing pipelines to improve the results of subsequent analyses.

MATERIALS AND METHODS

As mentioned in the introduction section, the proposed framework employs a 2-step approach to classify poor-quality dMRI volumes into four categories of artifacts. In the first step, the prominent artifact class labels of individual MECE 3D slabs (extracted from dMRI whole volumes) are predicted through design and training of a residual squeeze and excitation (SE) CNN. In the second step, the most consistent label amongst the predicted slabs' labels is chosen *via* a voting system as the prominent artifact type present in the whole dMRI volume. **Figure 1** shows an overview of the multiclass major artifact detection framework. As a proof of concept, we utilize the labels generated by the framework to run an FA-age correlation analysis in order to test the model's efficacy. In this section, details of our 2-step approach, the conducted FA-age analysis, and the datasets used in this work are discussed.

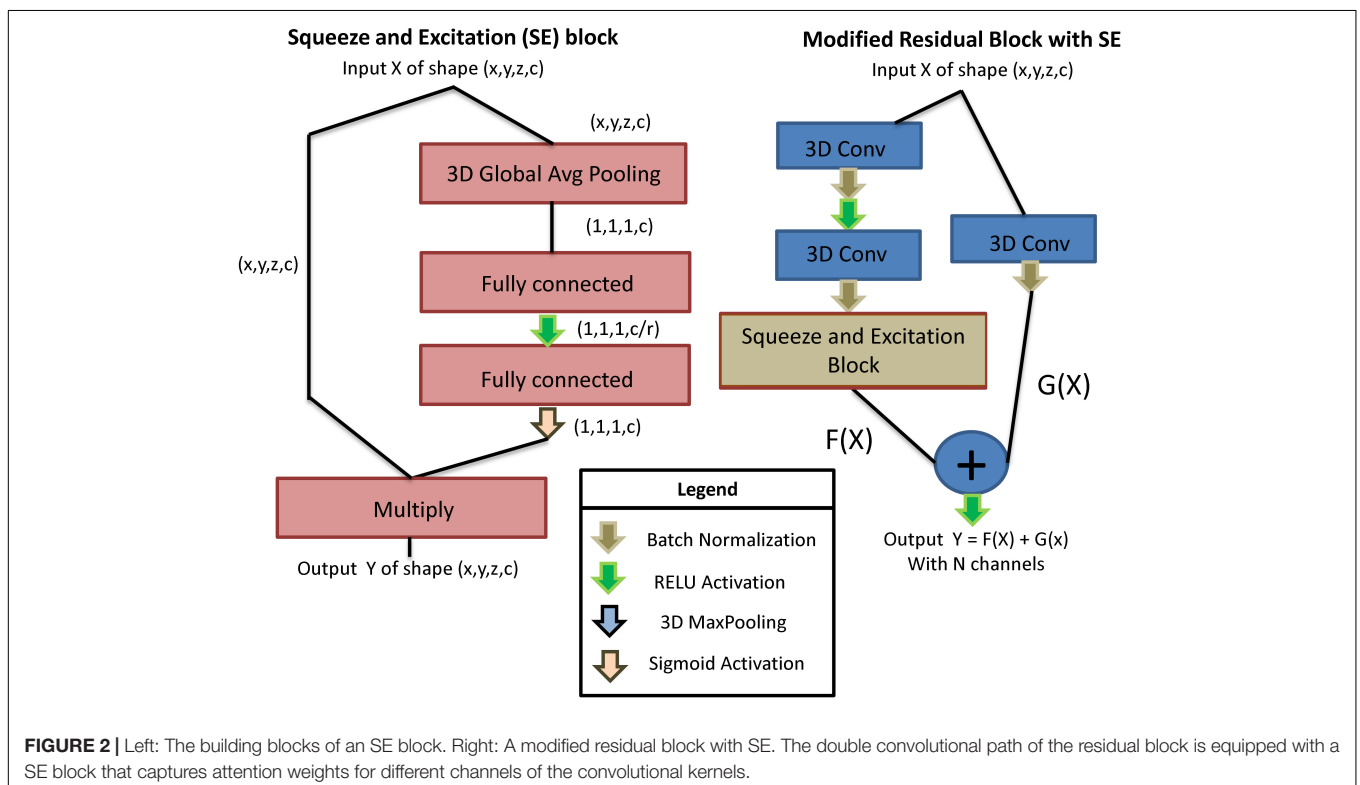
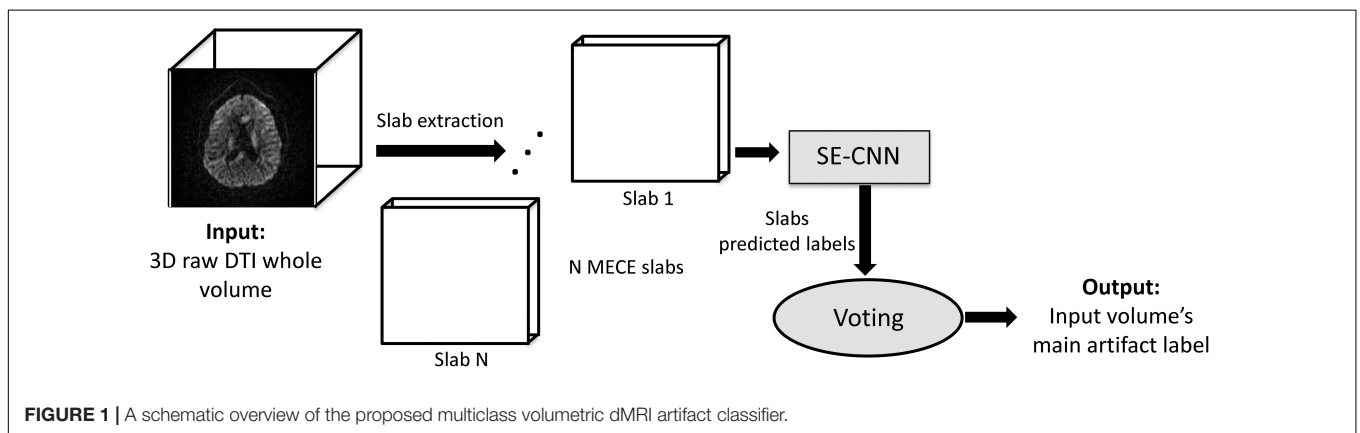
Slab-Level Artifact Classification Using Residual SE-CNN

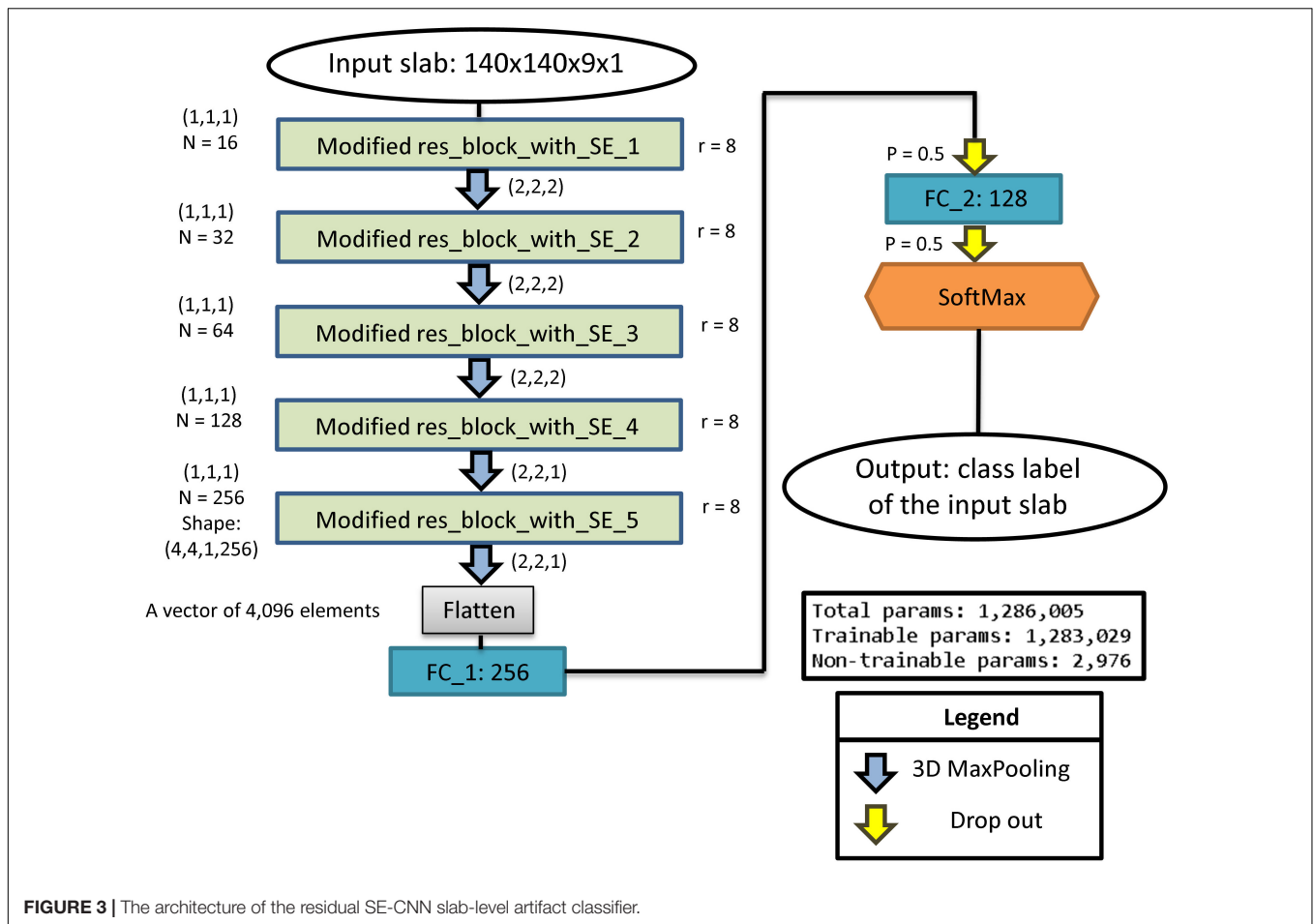
Due to memory limitations of the currently available typical GPUs, it is often not feasible to feed a whole 3D dMRI volume into a CNN. Hence, images need to be partitioned prior to feeding them to a GPU. Therefore, to address this issue while capturing artifact patterns over the entire 3D image without missing any region, we partition the dMRI volumes into a number of MECE slabs spanning the entire volume at hand. In addition to addressing GPU memory limitations, the MECE slab-based classifier approach has two benefits: (a) it allows the proposed residual SE-CNN model to capture local information, and to achieve global consistency, by sweeping the entire volume through 3D MECE regions, and (b) it allows utilization of a voting approach to predict the artifact type of a whole dMRI

volume, thereby making the final prediction more robust to slab-level misclassifications.

In order to classify the main artifact type in each slab associated with a dMRI volume, we use a custom-designed residual SE-CNN architecture. Our custom residual SE-CNN architecture consists of several cascaded modified versions of original residual blocks (He et al., 2016) equipped with squeeze and excitation components (Hu et al., 2018). The SE block is shown in **Figure 2** (left), and the modified residual block with SE is depicted in **Figure 2** (right). The SE block is a self-attention unit that aims at modeling the interdependencies between different channels within the layers (Hu et al., 2018). Essentially, SE blocks learn how to filter information across channels, in order to focus on the most relevant features for the classification task. This

is achieved through learning the relative importance weights of different channels. The process of learning the channel weights is done *via* a global average pooling along the image spatial dimensions (i.e., x, y, and z) followed by two fully connected layers with non-linear activation functions operating on the channel dimension. The key hyperparameter of a SE block is the squeeze parameter r (in the first fully connected layer) that reduces the number of channels by a ratio of r to find the most important channel representation for the classification task. After learning the channel weights, they are projected onto the original feature tensor by element-wise multiplication (see **Figure 2**, left). For more information on this topic, readers are directed to Hu et al. (2018). The SE block is then placed in the double convolutional path of the modified residual block to learn the





channels weights of the two preceding 3D convolutional kernels (see **Figure 2**, right). This block is called modified residual block with SE and is used sequentially in our classifier’s architecture.

The architecture of the proposed residual SE-CNN artifact classifier is shown in **Figure 3**. As depicted in **Figure 3**, our residual SE-CNN model is built with five stacked modified residual blocks with SE. After each block, a 3D Maxpooling layer is used to reduce the image dimensions. Throughout the network, the image spatial dimension is reduced while the number of channels is increased. For each layer, the kernel sizes and the number of channels are presented in **Figure 3**. After the 5th modified residual block with SE, the features vector (consisting of 4,096 elements) is unrolled and fed through two consecutive fully connected layers. Finally, a SoftMax classifier with 4 possible classes (motion, out of FOV, low SNR, and MRI miscellaneous artifacts) is used to perform the slab-level main artifact type classification. To regularize the model and minimize overfitting effects dropout units are placed after the fully connected layers.

Whole-Volume Artifact Prediction

After predicting the slab-level labels *via* the residual SE-CNN model, we utilize these labels to predict the main artifact label associated with the whole 3D dMRI volume *via* the voting block (**Figure 1**). The voting block takes the predicted labels for the N

MECE slabs of a dMRI whole volume as inputs and outputs the main existing artifact type of that dMRI volume. This is achieved *via* maximum consensus on the predicted labels between the slabs as detailed by Algorithm 1 presented in **Table 1**.

Fractional Anisotropy-Age Analysis as a Proof of Concept to Demonstrate Efficacy of the Artifact Classifier

The 4,226 manually labeled poor-quality dMRI volumes of the HBN dataset (details of this dataset are presented in the data and

TABLE 1 | Detailed steps of Algorithm 1 for voting on the final labels of dMRI whole volumes.

Algorithm 1: Generating whole volume artifact labels

Input: N slabs’ predicted labels: $L = \{l_1, l_2, \dots, l_N\}$

Output: Whole volume’s predicted artifact label: L_{Vol}

Procedure:

- Let integers 0, 1, 2, and 3 denote the four possible values for each label l_j
- Let n_j denote the count for the j^{th} possible label:
 - Then for each vector L we have:

$$N = \{n_0, n_1, n_2, n_3\}$$
- The whole volume label is predicted as follows:
 - $L_{Vol} = \text{argmax} \{n_j\}$ for $j = 0, 1, 2, \text{ and } 3$

pre-processing section) were derived from 66 randomly selected subjects from a larger pool of 100 participants whose age varied from 5.57 to 21.89 years with mean and median age of 10.95 and 10.05 years, respectively. The HBN dataset was selected because the age range was appropriate for the desired age-based analysis, and because many of the MRI scans were known to be corrupted with artifacts based on our prior experience with this dataset (Luna et al., 2021). Prior research has demonstrated a linear relationship between participant age and FA (Rathee et al., 2016; Richie-Halford et al., 2021). Exploiting this fact, a proof-of-concept analysis examined the bivariate correlation between mean whole-brain FA and participant age.

Two tests (referred to as A and B) were conducted in which A: the age of 100 HBN participants was correlated with their mean whole-brain FA values and, B: the same analysis was re-run after removing artifact corrupted dMRI volumes (as labeled by the SE-CNN artifact classifier). Following similar methods described elsewhere [e.g., (Pujol et al., 2016; Chow and Chang, 2017)], scans had their corrupted volumes removed if the number of poor-quality dMRI volumes were between 1 and 26 (or <20% of the dMRI volumes). Of note, scans with >26 poor-quality dMRI volumes were not excluded from this analysis to ensure the same sample size across the A/B testing. Of the 66 scans, the dMRI data of 17 were affected by the A/B testing. An increase in the FA-age correlation will support the efficacy of the classifier.

Data and Pre-processing

Diffusion MRI data is highly heterogeneous due to various reasons such as scanner differences, diffusion directions (various

gradient directions), demographics (e.g., age, gender, etc.), etc., (Ahmad et al., 2021). This heterogeneity makes it almost impossible to train a CNN on a dataset and use the trained model (without changing the learned parameter values) as is to account for the differences that may be found while analyzing a different study (Ahmad et al., 2021). Hence, to demonstrate the feasibility of high accuracy automated artifact classification using the proposed residual SE-CNN architecture, we trained our model on two different datasets separately, and evaluated the results. Namely, we used the Adolescent Brain Cognitive Development (ABCD; Casey et al., 2018) and Healthy Brain Network (HBN; Alexander et al., 2017) datasets. By training the model separately for each dataset, the model's parameters are learned optimally according to the target distribution. In what follows we discuss the two datasets as well as the annotation and pre-processing steps.

Adolescent Brain Cognitive Development Dataset

The goal of the ABCD study is to track human brain development over time (childhood through adolescence) (Casey et al., 2018). For this purpose, the study hired more than 10,000 participants between the age of 9–10 years old. Institutional review boards at 21 different sites that were involved in this study approved the study protocols. The ABCD dataset is available at¹. In this work we utilized multi-shell ($b = 0, 500, 1000, 2000, 3000$ s/mm²) diffusion scans from 85 participants. All scans are isotropic [$1.7 \times 1.7 \times 1.7$ mm³ with matrix size of ($140 \times 140 \times 81$)] and have identical diffusion directions (96). Due to imaging across

¹<https://nda.nih.gov/abcd>

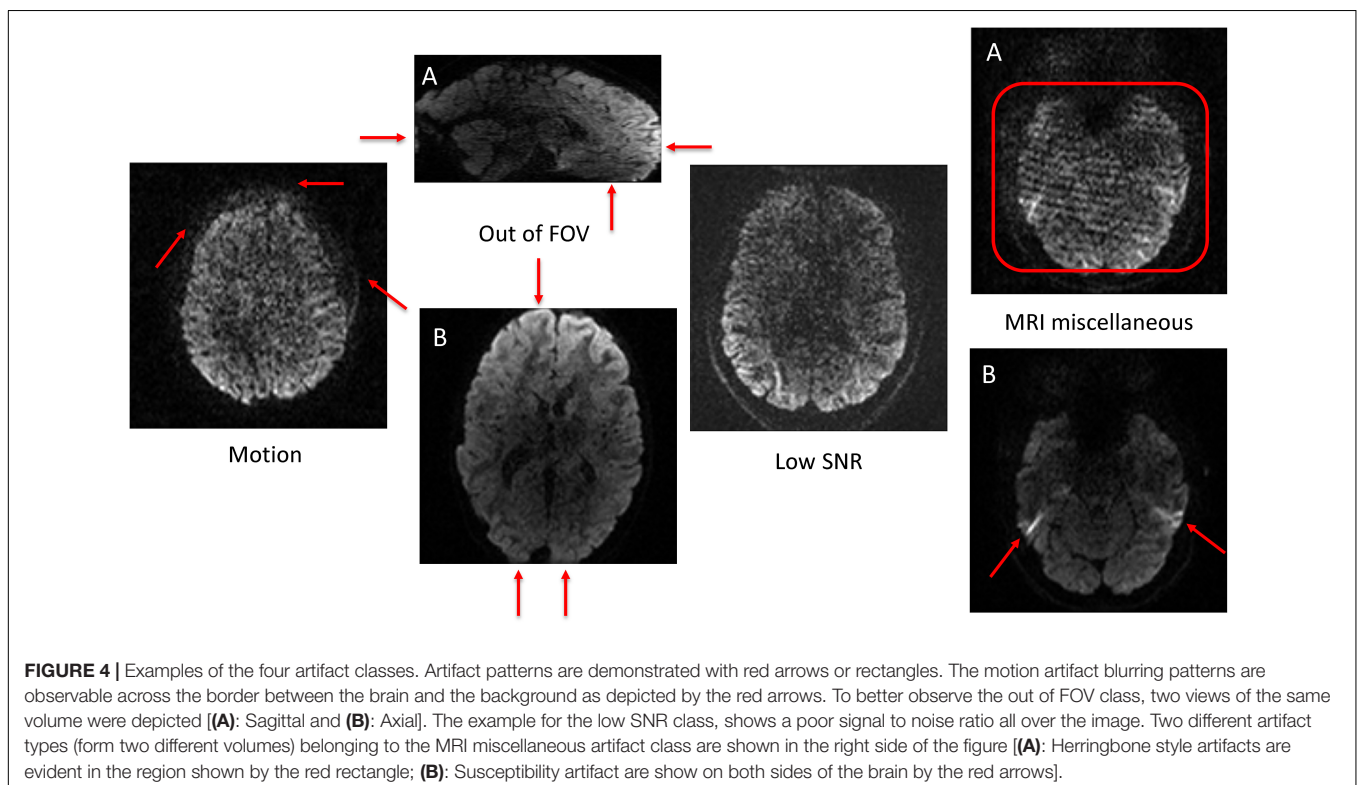


FIGURE 4 | Examples of the four artifact classes. Artifact patterns are demonstrated with red arrows or rectangles. The motion artifact blurring patterns are observable across the border between the brain and the background as depicted by the red arrows. To better observe the out of FOV class, two views of the same volume were depicted [(A): Sagittal and (B): Axial]. The example for the low SNR class, shows a poor signal to noise ratio all over the image. Two different artifact types (from two different volumes) belonging to the MRI miscellaneous artifact class are shown in the right side of the figure [(A): Herringbone style artifacts are evident in the region shown by the red rectangle; (B): Susceptibility artifacts are shown on both sides of the brain by the red arrows].

21 different sites, the acquisition parameters are slightly different which aids the model to learn robustness to heterogeneity of the acquisition parameters within the dataset.

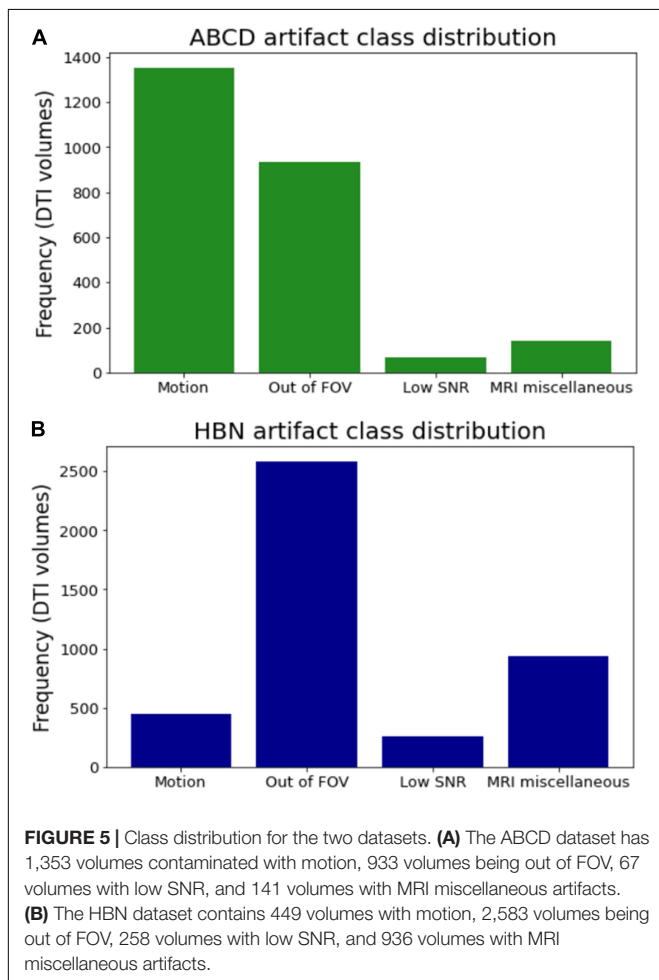
Healthy Brain Network Dataset

The Child Mind Institute² launched the HBN study in 2017 (Alexander et al., 2017) and made the data publicly available at³. This ongoing study focused on creating a large-scale dataset of 10,000 5–21 years old New York City area children and adolescents. This study utilized a community-referred recruitment strategy. The study design was approved by the Chesapeake Institutional Review Board⁴. In this work, we used multi-shell ($b = 0, 1000, 2000$ s/mm²) dMRI scans from 100 distinct subjects. All scans have isotropic resolution ($1.8 \times 1.8 \times 1.8$ mm³), with 72 slices, and identical diffusion directions (64). Since the in-plane matrix size of HBN dataset (mostly 104×104) is different from that of ABCD (i.e., 140×140) we resized the HBN images in the xy plane using the Bi-cubic interpolation implemented in the scikit-image Python

²<https://childmind.org/>

³http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/

⁴<https://www.chesapeakeirb.com/>



library (Van der Walt et al., 2014). The resized HBN volumes have a matrix size of ($140 \times 140 \times 72$).

Manual Annotation and Class-Distribution

All clean (no artifacts) volumes in both ABCD and HBN datasets were excluded from this work (except for the FA-age analysis). The remaining poor-quality volumes in both datasets (2,494 volumes in ABCD and 4,226 volumes in HBN) were manually annotated (at volume-level not slice-level, this reduces the labor intensity of manual annotation) into four main artifact classes (indexed as: 0: motion, 1: out of FOV, 2: low SNR, and 3: MRI miscellaneous artifacts) by an expert with 12 years of experience in MRI and DTI analysis. The MRI miscellaneous artifacts class refers to all other types of MRI (and dMRI) artifacts (such as Eddy currents, ghosting, etc.) that do not necessarily fall into the other three artifact categories considered here. One of the reasons for having this broad class is due to the difficulty of identifying a single major source of artifact in some dMRI volumes. Hence by having this class, the model can learn to distinguish between the other three prominent artifact types and classify a poor-quality volume into a 4th class if it doesn't fall into the other three categories. **Figure 4** illustrates examples of the artifact classes considered in this work. The volume-wise manually annotated class distributions in ABCD and HBN datasets are shown in **Figure 5**. The labeled volumes are randomly assigned to train, validation, and test sets using a split ratio of 6:2:2. The intensity of each volume is normalized to the range of (0, 1). To capture the visual patterns of the artifacts that manifest themselves better in the border between the brain and background (such as motion and out of FOV), no brain extraction or background removal was carried out.

Fractional Anisotropy-Age Pre-processing Pipeline

Standard dMRI pre-processing using MRtrix (Tournier et al., 2019) and FSL (Jenkinson et al., 2012) including denoising, EDDY (Andersson and Sotiropoulos, 2016), TOPUP (Andersson et al., 2003), and bias field correction (Zhang et al., 2001) was carried out for all 100 HBN participants followed by derivation of

TABLE 2 | The model and training's main hyperparameters.

Hyperparameter	Value/Algorithm
Kernel initialization technique	Glorot
Convolutional kernel size	[1, 1, 1] With zero paddings
Maxpooling kernel size	[2, 2, 2] (Last two layers: [2, 2, 1])
Squeeze parameter r	8
Activation function	RELU (Unless otherwise mentioned)
Dropout probability	0.5
Batch size	32
Optimizer	SGD with Nesterov Initial learning rate = 0.0001 Momentum = 0.6 Decay rate = 10^{-6}
Loss function	Categorical cross-entropy

whole-brain FA values (test A). The same pipeline was executed in addition to removing some of the noisy dMRI volumes (B). It is to be noted that the pre-processing can clean the dMRI data, but this cleaning is consistent across A/B.

RESULTS

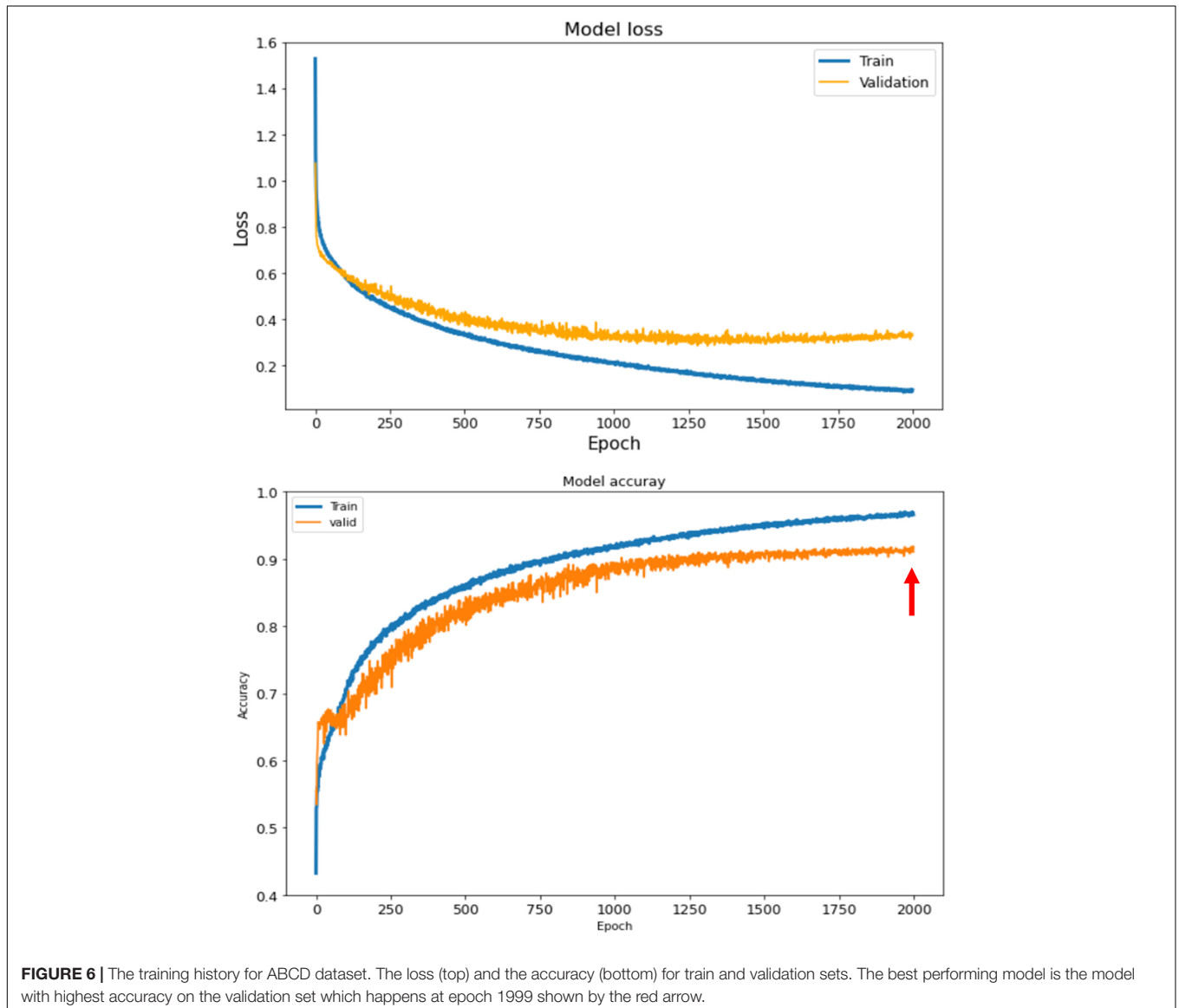
In this section we discuss our implementation and the results. In particular, we detail the slab extraction procedure, the architecture's hyperparameters, as well as the choice of hyperparameters for training the residual SE-CNN artifact classifier. Next, we present the slab-level classification results of the residual SE-CNN classifier on the ABCD and HBN test sets as well as the whole volume prediction accuracy using the voting procedure. Finally, we discuss the results of the FA-age analysis on the HBN dataset.

Extraction of Mutually Exclusive Collectively Exhaustive Slabs From Whole Volumes

As mentioned in the Materials and Methods section, we partition a whole dMRI volume into N MECE 3D slabs. Concatenation of these slabs along the z dimension forms the original dMRI volume without any loss of information. We chose a slab size of (140, 140, 9). Using this slab size, we end up with nine slabs per each dMRI volume in the ABCD and eight slabs for each dMRI volume in the HBN dataset. The slabs are then fed to the residual SE-CNN as inputs.

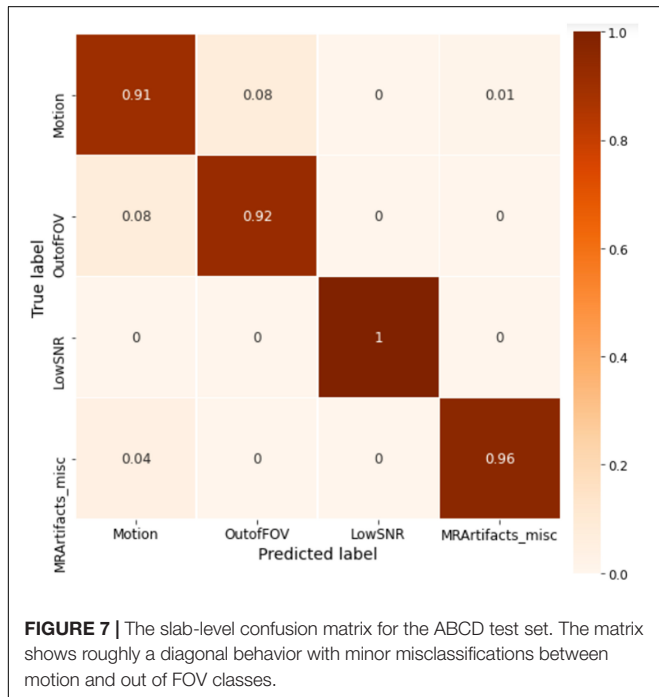
Model Hyperparameters and Training History

The architecture of our proposed residual SE-CNN was presented in **Figure 3**. This architecture with its set of chosen



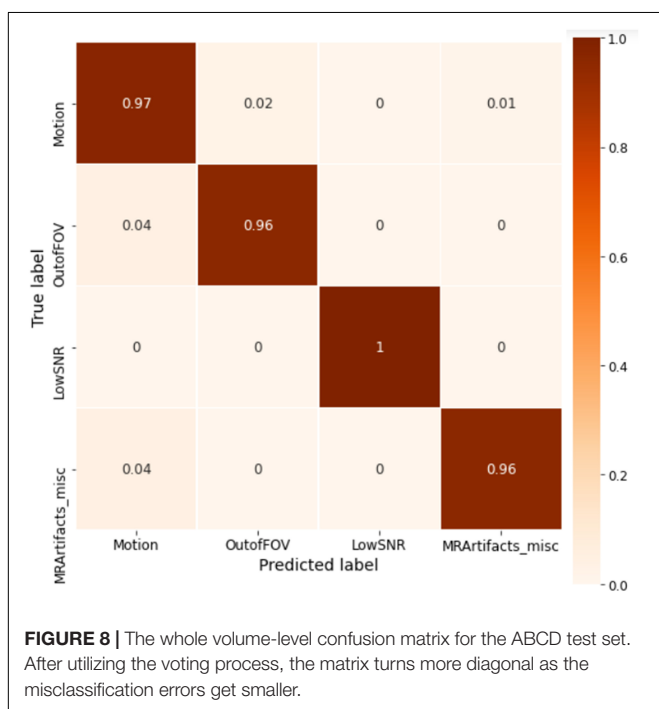
hyperparameters was used for both ABCD and HBN datasets. In the implementation phase, for all convolutional layers, RELU activation function was used (if any) as the choice of non-linearity unless explicitly noted. The kernels' parameters were initialized *via* the random Glorot initialization technique (Glorot and Bengio, 2010). All convolutional kernel sizes were set to (1, 1, 1) with zero paddings to keep the spatial dimensions same before and after employment of the convolutional filters. For

the Max-pooling layers, the first three consecutive blocks, used a pooling size of (2, 2, 2), and the pooling size of (2, 2, 1) was used for the last two blocks. The squeeze parameter r was set to 8. We used the RELU activation function for the last two fully connected layers (with 256 and 128 nodes, respectively). The probability for the two dropout units was set to 0.5. Finally, we used a SoftMax activation function for the SoftMax layer with four artifact classes: 0 (motion), 1 (out of FOV), 2 (low SNR), and 3 (MRI miscellaneous artifacts). The architecture was implemented in Python, using Keras with TensorFlow as backend. Except for the number of epochs, the training hyperparameters were also the same for both datasets. The model was trained for 2000 epochs on the ABCD dataset while the HBN dataset converged slightly faster and was trained for 1500 epochs. The batch size was set to 32. The model was trained, separately for ABCD and HBN, to minimize the categorical cross-entropy loss function using the manually labeled data in their respective training sets. For optimizing the cost function, SGD with Nesterov optimizer (Sutskever et al., 2013) (initial learning rate = 0.0001, Momentum = 0.6, and decay rate = 10^{-6}) was used. A summary of the hyperparameters and their values are presented in Table 2. In what follows we discuss the results on the two datasets separately.



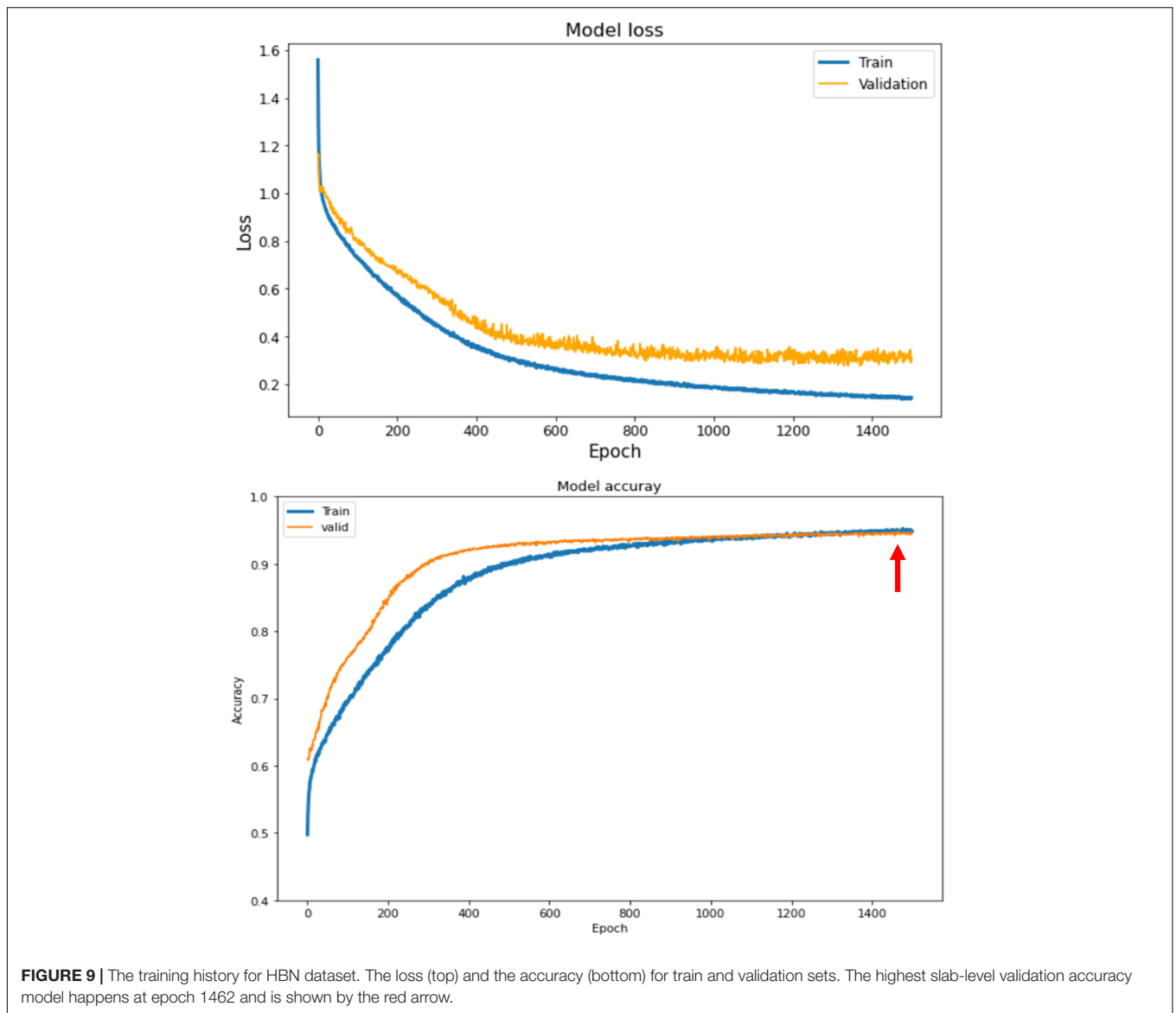
Evaluation on the Adolescent Brain Cognitive Development Dataset

The ABCD training history is depicted in Figure 6. As depicted, the training is stable and the overall accuracy is increasing over epochs. The criterion used for model selection was the highest overall classification accuracy on the validation set. The test set's slab-level confusion matrix for the best performing model (at epoch 1999) is presented in Figure 7. The slab-level classification accuracy of the model on the training, validation, and test sets were 96.80, 92.00, and 91.86%, respectively. Using the voting procedure discussed in Algorithm 1, we evaluated the performance of the model on whole volume major artifact classification. The model achieved a test set primary artifact classification average accuracy of 96.61%. The confusion matrix for whole volume predictions is depicted in Figure 8.



Evaluation on the Healthy Brain Network Dataset

The model was trained for 1500 epochs (using the exact same training hyperparameters as in ABCD) on the HBN dataset and the training history is shown in Figure 9. Similar to the ABCD, the training process also shows a stable behavior for the HBN dataset. The highest validation accuracy model (*i.e.*, the best model) was achieved at epoch 1462 resulting in slab-level classification accuracies of 95.00, 94.69, and 95.77% on the train, validation, and test sets, respectively. Figure 10 shows the slab-level artifact classification confusion matrix. Although the accuracy on the motion class is below 90%, through the voting process this accuracy increases as the misclassification errors become smaller by voting. To demonstrate this, Figure 11 shows the volume-level classification confusion matrix. The model archived an overall test set volume-level accuracy of 97.52% and the accuracy on the motion class improved significantly.



Fractional Anisotropy-Age Analysis Results

Results of the A/B testing demonstrate an improvement in FA-age correlation brought about by including the labels from the residual SE-CNN model. **Figures 12A,B** depict the linear regression fit with respective 95% confidence intervals for the FA-age relationship for the tests A and B, respectively. Test A demonstrates an R^2 of 0.116 whereas test B demonstrates an R^2 of 0.142 at the same significance level ($p < 0.001$).

DISCUSSION

Summary, Strengths, and Shortcomings

In this work, we developed an automatic 4-class major artifact classifier for 3D diffusion MRI volumes. We trained the exact

same architecture with the same exact choice of hyperparameters (both model's and training's hyperparameters) on two different datasets, namely ABCD and HBN. Our results demonstrate the capability of this architecture and the proposed hyperparameter choices, followed by the voting procedure to accurately classify four categories of artifacts in 3D dMRI volumes.

As mentioned earlier, diffusion MRI is highly diverse in nature. In particular, different diffusion directions across various studies is an important source of heterogeneity. This makes it highly challenging to propose and train an architecture that generalizes with an acceptable accuracy to unobserved and different distributions from the one it has been trained on. One main reason for this is due to the fact that the 3D convolutional kernels, when trained on a dataset with a specific set of diffusion directions, learn their spatial parameters' optimal values according to those specific directions. Hence, when the model faces a new dataset with different diffusion directions, the

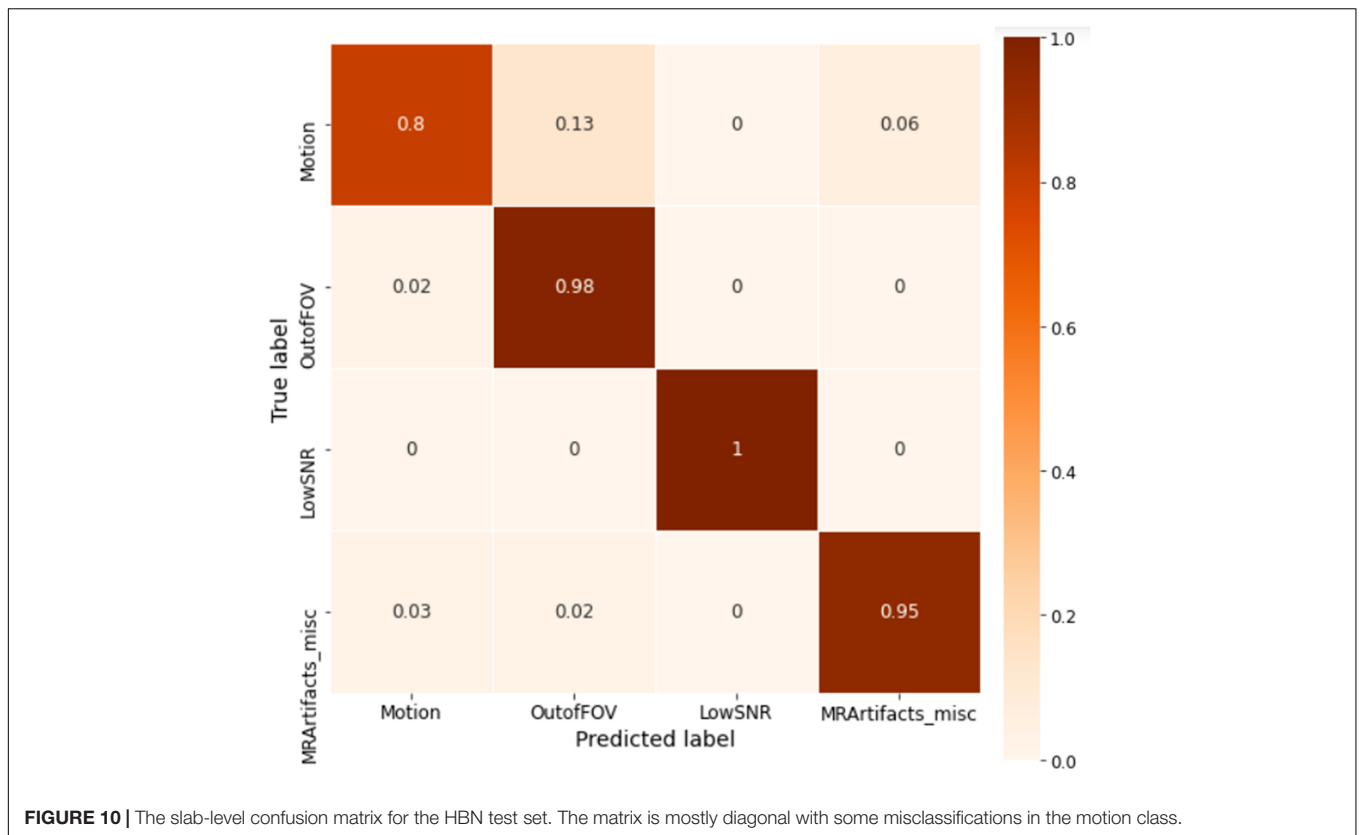


FIGURE 10 | The slab-level confusion matrix for the HBN test set. The matrix is mostly diagonal with some misclassifications in the motion class.

exact same learned values for the convolutional kernels are not optimal according to the new diffusion directions. Moreover, in the case of our architecture, the channel attention weights learned

in the SE blocks are also learned according to the diffusion directions of the original dataset they were trained on. These learned attention weights are not the optimal values for a dataset with different diffusion directions. Hence applying a model that is trained on a dataset, while fixing all of its learned parameters' values, to perform artifact classification on a different dataset is not a suitable solution.

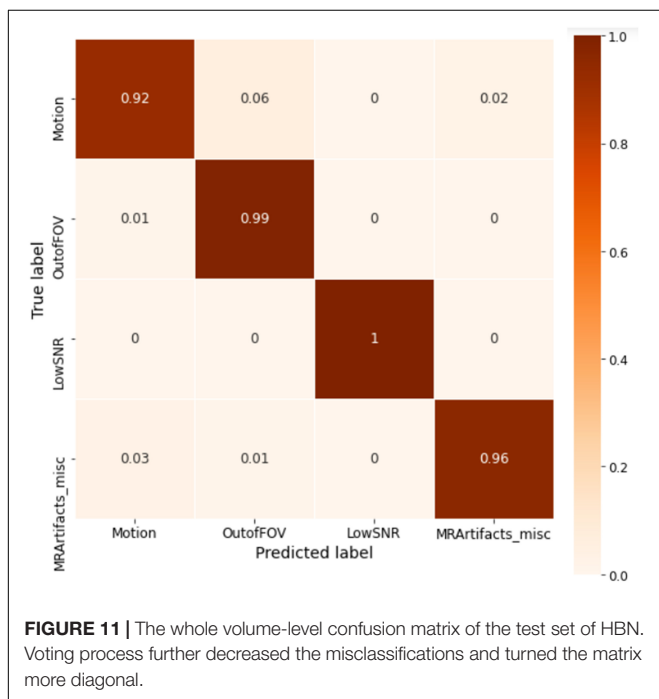
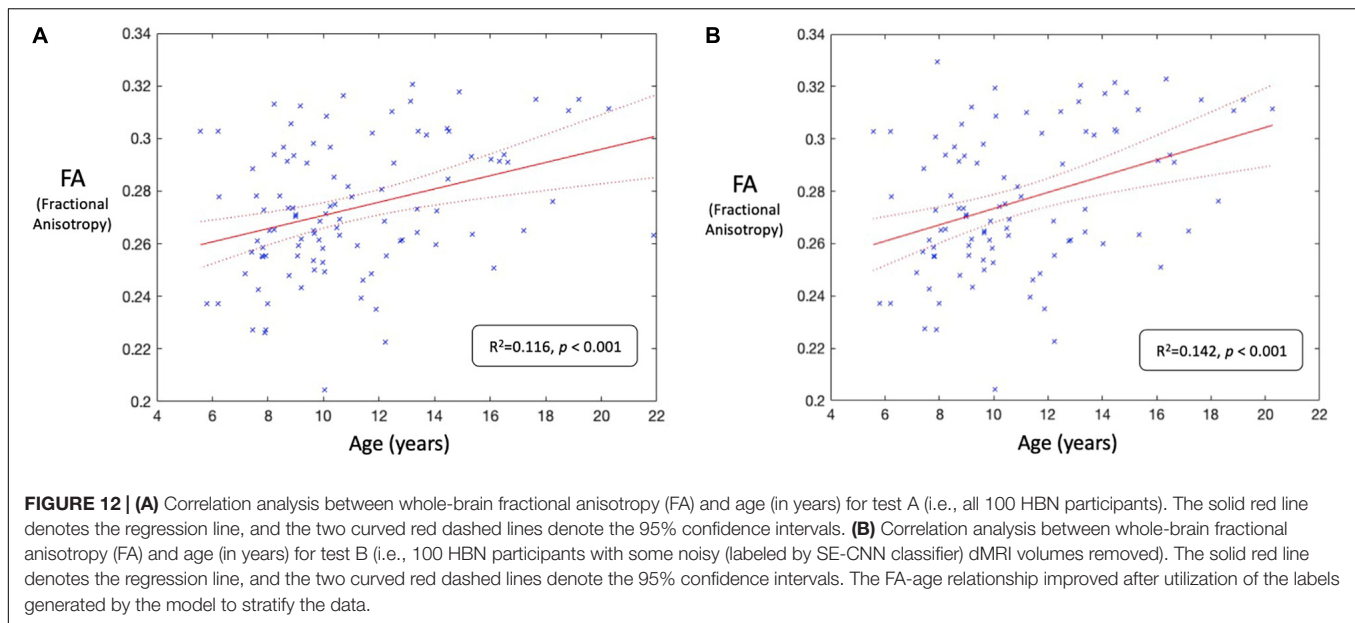


FIGURE 11 | The whole volume-level confusion matrix of the test set of HBN. Voting process further decreased the misclassifications and turned the matrix more diagonal.

To alleviate this issue, we proposed an architecture with a specific set of model's hyperparameters (not learnable parameters, it is important to note this distinction) as well as training's hyperparameters. We provided evidence that this architecture can be trained on a small subset of data when facing a new dataset to find the optimal kernel parameters values according to different heterogeneity sources, especially the diffusion directions. Previous works on binary ("poor" vs. "good") quality control of dMRI volumes such as (Ahmad et al., 2021), which is a much simpler task than a 4-class artifact classification, have also proposed to re-train the network on a small subset of manually annotated data to solve the generalization issue when facing a new diffusion MRI dataset.

The major challenges with training a neural network model are twofolds: (1) the labor-intensive task of manual annotation, and (2) model's as well as training's hyperparameters tuning process. Through our proposed framework we have tried to address these two challenges to make it more convenient for potential users to be able to train the model according to their target diffusion MRI dataset. First, since our framework utilizes a 3D architecture, manual annotation needs to be done only



at the volume level (not slice/voxel level) by assigning a single label to a whole volume. This significantly reduces the load of manual annotation task. For example, in our case, it only took our expert ~ 17 min to manually annotate one subject (~ 110 volumes) with four classes of artifacts. Second, we found a set of hyperparameters for the model's architecture as well as the training's hyperparameters, which gave us comparable results in two different datasets. Based on these results, we believe these sets of hyperparameters can be used to train the proposed architecture on a different dataset or at least serve as a starting point to reduce the search space.

As mentioned in the introduction section, most works on automated artifact detection or quality control of dMRI scans have been done on a binary ("poor" vs. "good") level. To the best of our knowledge, our proposed framework is the first automated and accurate method to consider a wider range of artifact classification (i.e., four classes of motion, out of FOV, low SNR, and MRI miscellaneous artifacts). Collecting dMRI scans is labor intensive and a costly process, especially when dealing with infants. Hence, we believe by teasing out the type of artifact in poor-quality volumes, one might be able to correct for the detected artifact (if possible) and restore the volumes to reduce the cost of acquiring dMRI scans.

The proposed model improved in the FA-age correlation when its predicted labels were incorporated to stratify the dataset. This improvement supports the utility of the model in a standard dMRI pre-processing application. The increase in FA-age correlation after removal of corrupted dMRI volumes underscores the applicability of this artifact classifier.

While we demonstrated the capability of our method to provide a convenient mean for accurate automatic classification of four categories of artifacts in poor-quality diffusion MRI volumes, we are aware of its shortcomings. The current method only addresses four categories of artifacts. In the future we plan to further improve upon the framework to be able to consider a

wider range of artifacts. In other words, we plan to further tease out the types of artifacts in the MRI miscellaneous category (the 4th class). Additionally, the current framework only identifies a single major artifact type for each volume. In the future, we plan to adapt our model to output multiple potential artifact types (if any) in cases where there might be several different artifact categories present in the volume under investigation. Moreover, the proposed method currently predicts artifact types on whole volume-level. While in most cases, the entire poor-quality volume is affected by the detected artifact and hence discarded (or corrected, if possible), in some cases where the size of the dataset is small, users are interested in only discarding (or correcting, if possible) the few slices that were affected by the detected artifact. Hence, in the future we plan to find ways and adapt our framework to further improve the resolution of our artifact classification from volume-level to slice-level.

CONCLUSION

Artifact type identification is an exhaustive but essential step in pre-processing of dMRI data to discard or correct (if possible) the contaminated volumes. Without artifact correction or removal of the contaminated volumes, one cannot guarantee the accuracy of any subsequent analysis and draw reliable conclusions. In this paper, we proposed a deep learning architecture, namely a 3D residual SE-CNN, followed by a voting procedure to automatically classify poor-quality volumes into 4 categories of artifacts (i.e., motion, out of FOV, low SNR, and MRI miscellaneous artifacts). Our results demonstrate the capability of the proposed framework in accurate multiclass artifact classification. Moreover, to take into account, the heterogeneity of the dMRI data, we found a set of hyperparameters for the model as well as training hyperparameters that were able to provide accurate classifications on two different datasets (i.e., 96.61 %

on ABCD's test set and 97.52% on HBN's). The provided sets of hyperparameters can be used to guide the potential users with training the proposed architecture according to their own dataset to compensate for the heterogeneity. This potentially enables the framework to be integrated in dMRI processing pipelines for fast automatic artifact type identification.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: ABCD dataset: <https://nda.nih.gov/abcd>, and HBN dataset: http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by for ABCD dataset: Institutional review boards at 21 different sites that were involved in this study, and for HBN dataset: The Chesapeake Institutional Review Board. Written

REFERENCES

- Ahmad, A., Parker, D., Samani, Z. R., and Verma, R. (2021). 3D-QCNet—a pipeline for automated artifact detection in diffusion MRI images. *arXiv [preprint]* arXiv:2103.05285.
- Alexander, A. L., Lee, J. E., Lazar, M., and Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics* 4, 316–329.
- Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., et al. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* 4, 1–26. doi: 10.1038/sdata.2017.181
- Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L., Griffanti, L., Douaud, G., et al. (2018). Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* 166, 400–424. doi: 10.1016/j.neuroimage.2017.10.034
- Andersson, J. L., Graham, M. S., Zsoldos, E., and Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage* 141, 556–572. doi: 10.1016/j.neuroimage.2016.06.058
- Andersson, J. L., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* 20, 870–888. doi: 10.1016/S1053-8119(03)00336-7
- Andersson, J. L., and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019
- Baliyan, V., Das, C. J., Sharma, R., and Gupta, A. K. (2016). Diffusion weighted imaging: technique and applications. *World J. Radiol.* 8:785. doi: 10.4329/wjr.v8.i9.785
- Bammer, R., Markl, M., Barnett, A., Acar, B., Alley, M. T., Pelc, N. J., et al. (2003). Analysis and generalized correction of the effect of spatial gradient field distortions in diffusion-weighted imaging. *Magn. Reson. Med.* 50, 560–569. doi: 10.1002/mrm.10545
- Basser, P. J., and Jones, D. K. (2002). Diffusion-tensor MRI: theory, experimental design and data analysis—a technical review. *NMR Biomed.* 15, 456–467. doi: 10.1002/nbm.783
- Bastiani, M., Cottaar, M., Fitzgibbon, S. P., Suri, S., Alfaro-Almagro, F., Sotiropoulos, S. N., et al. (2019). Automated quality control for within and between studies diffusion MRI data using a non-parametric framework for movement and distortion correction. *Neuroimage* 184, 801–812. doi: 10.1016/j.neuroimage.2018.09.073

informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

NE: methodology, implementation, visualization, and writing (original draft, review and editing). PK and KD: dMRI analysis, visualization, and writing (review and editing). XZ: implementation and writing (review and editing). YW and JG: writing (review and editing). DS: data visualization, manual labeling, and writing (review and editing). JP: conceptualization, dMRI supervision, and writing (review and editing). AL: conceptualization, writing (review and editing), and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by funding from the National Institute of Health (NIH) award numbers: UH3OD023328, R01MH121070, and R01MH119510.

- Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., et al. (2018). The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* 32, 43–54. doi: 10.1016/j.dcn.2018.03.001
- Chow, H. M., and Chang, S. E. (2017). White matter developmental trajectories associated with persistence and recovery of childhood stuttering. *Hum. Brain Mapp.* 38, 3345–3359. doi: 10.1002/hbm.23590
- Ettehad, N., Zhang, X., Wang, Y., Semanek, D., Guo, J., Posner, J., et al. (2021). “Automatic volumetric quality assessment of diffusion MR images via convolutional neural network classifiers,” in *Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (Piscataway, NJ: IEEE), 2756–2760. doi: 10.1109/EMBC46164.2021.9630834
- Fantini, I., Rittner, L., Yasuda, C., and Lotufo, R. (2018). “Automatic detection of motion artifacts on MRI using Deep CNN,” in *Proceedings of the 2018 International Workshop on Pattern Recognition in Neuroimaging (PRNI)* (Singapore: IEEE), 1–4.
- Glorot, X., and Bengio, Y. (2010). “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)*. JMLR Workshop and Conference, Sardinia, 249–256.
- Graham, M. S., Drobniak, I., and Zhang, H. (2018). A supervised learning approach for diffusion MRI quality control with minimal training data. *Neuroimage* 178, 668–676. doi: 10.1016/j.neuroimage.2018.05.077
- Haddad, S. M., Scott, C. J., Ozzoude, M., Holmes, M. F., Arnott, S. R., Nanayakkara, N. D., et al. (2019). Comparison of quality control methods for automated diffusion tensor imaging analysis pipelines. *PLoS One* 14:e0226715. doi: 10.1371/journal.pone.0226715
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, 770–778.
- Hoy, A. R., Ly, M., Carlsson, C. M., Okonkwo, O. C., Zetterberg, H., Blennow, K., et al. (2017). Microstructural white matter alterations in preclinical Alzheimer's disease detected using free water elimination diffusion tensor imaging. *PLoS One* 12:e0173982. doi: 10.1371/journal.pone.0173982
- Hu, J., Shen, L., and Sun, G. (2018). “Squeeze-and-excitation networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, 7132–7141.

- Huisman, T. A. (2003). Diffusion-weighted imaging: basic concepts and application in cerebral stroke and head trauma. *Eur. Radiol.* 13, 2283–2297. doi: 10.1007/s00330-003-1843-6
- Huisman, T. A. G. M. (2010). Diffusion-weighted and diffusion tensor imaging of the brain, made easy. *Cancer Imaging* 10:S163. doi: 10.1102/1470-7330.2010.9023
- Hüppi, P. S., and Dubois, J. (2006). Diffusion tensor imaging of brain development. *Semin. Fetal Neonatal Med.* 11, 489–497. doi: 10.1016/j.siny.2006.07.006
- Iglesias, J. E., Lerma-Usabiaga, G., Garcia-Peraza-Herrera, L. C., Martinez, S., and Paz-Alonso, P. M. (2017). “Retrospective head motion estimation in structural brain MRI with 3D CNNs,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention* (Cham: Springer), 314–322. doi: 10.1007/978-3-319-66185-8_36
- Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M. (2012). *Fsl. Neuroimage* 62, 782–790.
- Jiang, H., Van Zijl, P. C., Kim, J., Pearlson, G. D., and Mori, S. (2006). DtiStudio: resource program for diffusion tensor computation and fiber bundle tracking. *Comput. Methods Programs Biomed.* 81, 106–116.
- Kelly, C., Pietsch, M., Counsell, S., and Tournier, J. D. (2017). “Transfer learning and convolutional neural net fusion for motion artefact detection,” in *Proceedings of the Annual Meeting of the International Society for Magnetic Resonance in Medicine*, Vol. 3523, Honolulu, HI.
- Krupa, K., and Bekiesińska-Figatowska, M. (2015). Artifacts in magnetic resonance imaging. *Pol. J. Radiol.* 80:93.
- Ladouceur, C. D., Peper, J. S., Crone, E. A., and Dahl, R. E. (2012). White matter development in adolescence: the influence of puberty and implications for affective disorders. *Dev. Cogn. Neurosci.* 2, 36–54. doi: 10.1016/j.dcn.2011.06.002
- Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., and Laval-Jeantet, M. (1986). MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 161, 401–407. doi: 10.1148/radiology.161.2.3763909
- Le Bihan, D., Poupon, C., Amadon, A., and Lethimonnier, F. (2006). Artifacts and pitfalls in diffusion MRI. *J. Magn. Reson. Imaging* 24, 478–488. doi: 10.1002/jmri.20683
- Liu, B., Zhu, T., and Zhong, J. (2015). Comparison of quality control software tools for diffusion tensor imaging. *Magn. Reson. Imaging* 33, 276–285. doi: 10.1016/j.mri.2014.10.011
- Lo Buono, V., Palmeri, R., Corallo, F., Allone, C., Pria, D., Bramanti, P., et al. (2020). Diffusion tensor imaging of white matter degeneration in early stage of Alzheimer’s disease: a review. *Int. J. Neurosci.* 130, 243–250. doi: 10.1080/00207454.2019.1667798
- Luna, A., Bernanke, J., Kim, K., Aw, N., Dworkin, J. D., Cha, J., et al. (2021). Maturity of gray matter structures and white matter connectomes, and their relationship with psychiatric symptoms in youth. *Hum. Brain Mapp.* 42, 4568–4579. doi: 10.1002/hbm.25565
- Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., et al. (2014). DTIPrep: quality control of diffusion-weighted images. *Front. Neuroinform.* 8:4. doi: 10.3389/fninf.2014.00004
- Pierpaoli, C., Walker, L., Irfanoglu, M. O., Barnett, A., Basser, P., Chang, L. C., et al. (2010). “TORTOISE: an integrated software package for processing of diffusion MRI data,” in *Proceedings of the ISMRM 18th annual meeting*, Vol. 1597, Stockholm.
- Pujol, J., Martínez-Vilavella, G., Macià, D., Fenoll, R., Alvarez-Pedrerol, M., Rivas, I., et al. (2016). Traffic pollution exposure is associated with altered brain connectivity in school children. *Neuroimage* 129, 175–184. doi: 10.1016/j.neuroimage.2016.01.036
- Rathee, R., Rallabandi, V. S., and Roy, P. K. (2016). Age-related differences in white matter integrity in healthy human brain: evidence from structural MRI and diffusion tensor imaging. *Magn. Reson. Insights* 9, 9–20. doi: 10.4137/MRI.S39666
- Reuter, M., Tisdall, M. D., Qureshi, A., Buckner, R. L., van der Kouwe, A. J., and Fischl, B. (2015). Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *Neuroimage* 107, 107–115. doi: 10.1016/j.neuroimage.2014.12.006
- Richie-Halford, A., Yeatman, J. D., Simon, N., and Rokem, A. (2021). Multidimensional analysis and detection of informative features in human brain white matter. *PLoS Comput. Biol.* 17:e1009136. doi: 10.1371/journal.pcbi.1009136
- Roalf, D. R., Gur, R. E., Verma, R., Parker, W. A., Quarmley, M., Ruparel, K., et al. (2015). White matter microstructure in schizophrenia: associations to neurocognition and clinical symptomatology. *Schizophr. Res.* 161, 42–49. doi: 10.1016/j.schres.2014.09.026
- Roalf, D. R., Quarmley, M., Elliott, M. A., Satterthwaite, T. D., Vandekar, S. N., Ruparel, K., et al. (2016). The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. *Neuroimage* 125, 903–919. doi: 10.1016/j.neuroimage.2015.10.068
- Samani, Z. R., Alappatt, J. A., Parker, D., Ismail, A. A. O., and Verma, R. (2020). QC-Automator: deep learning-based automated quality control for diffusion mr images. *Front. Neurosci.* 13:1456. doi: 10.3389/fnins.2019.01456
- Simmonds, D. J., Hallquist, M. N., Asato, M., and Luna, B. (2014). Developmental stages and sex differences of white matter and behavioral development through adolescence: a longitudinal diffusion tensor imaging (DTI) study. *Neuroimage* 92, 356–368. doi: 10.1016/j.neuroimage.2013.12.044
- Stejskal, E. O., and Tanner, J. E. (1965). Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.* 42, 288–292. doi: 10.1063/1.1695690
- Sutskever, I., Martens, J., Dahl, G., and Hinton, G. (2013). “On the importance of initialization and momentum in deep learning,” in *Proceedings of the International Conference on Machine Learning* (Cambridge, MA: PMLR), 1139–1147. doi: 10.3390/brainsci10070427
- Tønnesen, S., Kaufmann, T., Doan, N. T., Alnæs, D., Córdova-Palomera, A., Meer, D. V. D., et al. (2018). White matter aberrations and age-related trajectories in patients with schizophrenia and bipolar disorder revealed by diffusion tensor imaging. *Sci. Rep.* 8:14129. doi: 10.1038/s41598-018-32355-9
- Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., et al. (2019). MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage* 202:116137. doi: 10.1016/j.neuroimage.2019.116137
- Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). scikit-image: image processing in Python. *PeerJ* 2:e453. doi: 10.7717/peerj.453
- Van Dijk, K. R., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 431–438. doi: 10.1016/j.neuroimage.2011.07.044
- Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., et al. (2012). The human connectome project: a data acquisition perspective. *Neuroimage* 62, 2222–2231. doi: 10.1016/j.neuroimage.2012.02.018
- Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57. doi: 10.1109/42.906424

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ettehadhi, Kashyap, Zhang, Wang, Semanek, Desai, Guo, Posner and Laine. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.