The Journal of High School Research

Hybrid Deep Learning Approaches for Classifying Autism from Brain MRI

Ashley Chen*

Submitted: 17 October 2025 **Accepted:** 29 October 2025 **Publication date:** 29 October 2025

DOI: 10.70671/mdqw4m29

Abstract: Background: Autism spectrum disorder (ASD) is most often diagnosed using behavioral evaluations, which can vary between clinicians. Brain imaging, combined with machine learning, may help identify more objective patterns linked to ASD. In this research, magnetic resonance imaging (MRI) data from the publicly available Autism Brain Imaging Data Exchange I (ABIDE I) dataset (n = 1,112) were used to test two approaches for classifying individuals with ASD and typically developing control participants. Both groups were drawn from the same ABIDE I dataset, which includes MRI scans from individuals diagnosed with ASD as well as neurotypical individuals without any major neurological or psychiatric conditions. The first approach used a three-dimensional convolutional neural network (CNN) trained end-to-end, while the second employed a hybrid method in which the CNN served as a feature extractor followed by a support vector machine (SVM) classifier. Results showed that the baseline CNN achieved moderate performance, whereas the hybrid CNN + SVM model demonstrated higher overall accuracy and produced more balanced results between ASD and control groups. These findings indicate that separating feature extraction from classification can improve performance and reduce bias between diagnostic categories. Overall, this study suggests that deep learning methods may enhance the reliability and objectivity of MRI-based research on ASD.

Author keywords: Autism spectrum disorder; magnetic resonance imaging; MRI; convolutional neural networks; support vector machines; biomarkers

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition that affects social communication, behavior, and sensory processing.[1] Diagnosis currently relies on behavioral assessments such as the Autism Diagnostic Observation Schedule 2 (ADOS-2) and the Autism Diagnostic Interview-Revised (ADI-R). While these tools are widely used, they are usually subjective and may depend on individual interpretation. As a result, diagnostic outcomes may vary across evaluators and settings.[2]

The need for more objective diagnostic methods has grown alongside rising prevalence rates. According to data from the U.S. Centers for Disease Control and Prevention (CDC), approximately one in 150 children were identified with ASD in 2000, compared to one in 36 by 2020 (Table 1). This increasing prevalence highlights the importance of developing tools that can support early and consistent identification.

Neuroimaging offers one possible path toward objective biomarkers. Advances in machine learning, particularly deep learning methods such as convolutional neural networks (CNNs), offer new opportunities to capture subtle and spatially distributed neuroanatomical differences that may not be evident with traditional analysis. CNNs have shown strong performance in medical imaging, including mammography,[4] brain tumor grading,[5] and diabetic retinopathy screening.[6] These successes suggest that CNNs are well suited for ASD research, where subtle structural variations across multiple brain regions must be identified.

Despite their potential, CNNs face several challenges. They require large datasets to achieve reliable generalization,[7] yet many neuroimaging studies in ASD involve only a few hundred participants. Models are also highly sensitive to variability in scanner hardware and acquisition protocols. These differences can produce site-specific artifacts that may be mistaken for disorder-related features.[8] In addition, CNNs often function as "black boxes," with limited interpretability even when tools such as saliency maps or class activation methods are applied.[9] These limitations complicate efforts to build clinically robust biomarkers and highlight the importance of evaluating model generalizability.

This study investigates whether CNNs applied to structural MRI can distinguish individuals with ASD from typically developing controls (TDCs). A baseline end-to-end CNN was compared to a hybrid framework in which CNN-derived features were classified with support vector machines (SVM). The aim was to determine whether separating feature learning from classification improves accuracy and

*Corresponding Author: Ashley Chen.
Email: ashleylalachen@gmail.com
Oakton High School, 2900 Sutton Rd, Vienna, VA 22181, USA

Table 1. Rates of ASD diagnosis[3]

Surveillance year	Birth year	Number of sites reporting	Combined prevalence per 1,000 children (Range across sites)	This is about 1 in X children
2020	2012	11	27.6 (23.1–44.9)	1 in 36
2018	2010	11	23.0 (16.5–38.9)	1 in 44
2016	2008	11	18.5 (18.0–19.1)	1 in 54
2014	2006	11	16.8 (13.1–29.3)	1 in 59
2012	2004	11	14.5 (8.2–24.6)	1 in 69
2010	2002	11	14.7 (5.7–21.9)	1 in 68
2008	2000	14	11.3 (4.8–21.2)	1 in 88
2006	1998	11	9.0 (4.2–12.1)	1 in 110
2004	1996	8	8.0 (4.6–9.8)	1 in 125
2002	1994	14	6.6 (3.3–10.6)	1 in 150
2000	1992	6	6.7 (4.5–9.9)	1 in 150

Table 2. Demographic and clinical characteristics of participants across ABIDE I sites[10]

Site	N	% ASD	% Male	Age (Mean ± Std)
California Institute of Technology	23	47.8	78.3	27.1 ± 5.8
Carnegie Mellon University	32	40.6	78.1	26.8 ± 9.8
Kennedy Krieger Institute	44	25.0	77.3	10.1 ± 1.2
University of Leuven	62	46.8	88.7	18.1 ± 5.0
Ludwig Maximilians University Munich	34	5.9	88.2	25.3 ± 10.3
NYU Langone Medical Center	122	43.4	73.0	13.8 ± 5.8
Olin, Institute of Living at Hartford Hospital	36	55.6	86.1	16.8 ± 3.5
Social Brain Lab BCN NIC UMC Groningen and Netherlands Institute for Neurosciences	17	11.8	100	32.7 ± 7.0
San Diego State University	24	12.5	70.8	14.1 ± 1.9
Trinity Centre for Health Sciences	35	28.6	100	16.8 ± 3.5
University of California, Los Angeles	102	54.9	88.2	13.1 ± 2.5
University of Michigan	129	41.1	81.4	14.2 ± 3.3
University of Pittsburgh School of Medicine	56	51.8	85.7	18.8 ± 6.9
University of Utah School of Medicine	100	57.0	100	22.1 ± 7.7
Yale Child Study Center	34	17.7	70.6	13.1 ± 2.8

Note: Values are reported as sample size (N), percentage diagnosed with ASD, percentage male, and mean age ± standard deviation.

generalization, thereby advancing progress toward objective neuroimaging biomarkers for ASD.

Methodology

Dataset

This study used structural magnetic resonance imaging (sMRI) data from the Autism Brain Imaging Data Exchange I (ABIDE I), a publicly available repository aggregating neuroimaging and clinical data from 17 international research sites, including both individuals diagnosed with ASD and TDCs. TDCs are participants who do not have ASD or any other major neurological, developmental, or psychiatric conditions. They are considered neurotypical and serve as the comparison group for evaluating differences relative to individuals with ASD.

In total, 1,112 participants were included, comprising 539 individuals with ASD and 573 controls. Participant ages ranged from childhood to adulthood (mean site ages between ∼10 and ∼33 years), and the sample reflected the known male predominance in ASD. Table 2 summarizes the distribution of participants across contributing sites, including sample sizes, sex ratios, diagnostic proportions, and mean ages.

Structural MRI provides high-resolution anatomical images of the brain, capturing cortical thickness, white matter, and subcortical morphology. This modality was selected because it offers stable, reproducible measures of brain anatomy suitable for machine learning pipelines.

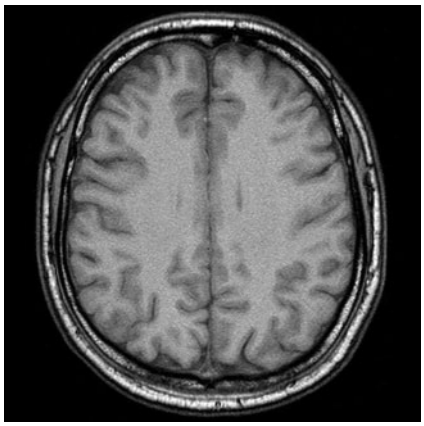


Figure 1. Example of a T1-weighted structural magnetic resonance imaging scan from a typically developing control participant in the Autism Brain Imaging Data Exchange I dataset[11]

Preprocessing

Before analysis, raw MRI scans underwent extensive preprocessing to account for heterogeneity across the 17 acquisition sites in ABIDE I and to prepare the data for input into CNNs. These steps were essential for reducing scanner-related variability while preserving biologically meaningful structural features relevant to autism. The preprocessing pipeline included the following stages:

1. **Skull Stripping:** Non-brain tissue such as scalp, skull, and dura was removed using the HD-BET algorithm,[12] a deep-learning based brain extraction tool. This ensured that subsequent analyses focused exclusively on brain tissue.
2. **Spatial Normalization:** Each scan was registered to the MN152 standard brain template using Advanced Normalization Tools. This step placed all participant images into a common stereotactic space, aligning anatomical structures across subjects and reducing variation due to head position or scanner orientation.
3. **Intensity Normalization:** MRI intensity values vary systematically across scanners and acquisition protocols, which can introduce site-specific artifacts into multi-site datasets such as ABIDE. To mitigate these effects, voxel intensities were normalized using a **histogram standardization**. This method aligns the intensity distribution of each subject's scan to a common reference distribution, thereby reducing scanner-related variability while preserving biologically meaningful signals.

Formally, each voxel intensity I was transformed as:

$$I' = F_{ref}{}^{-1}(F(I)) \quad (1)$$

where $F(I)$ is the cumulative distribution function (CDF) of the subject's intensity histogram, and F_{ref}^{-1} is the inverse CDF of a chosen reference histogram (in this case, derived from the median intensity distribution across all participants). This mapping ensures that intensities are rescaled consistently across sites, preventing the CNN from inadvertently learning scanner differences instead of ASD-related features.

4. **Resampling and Cropping:** Following normalization, volumes were resampled from their original native resolution (approximately 256 × 256 × 160 voxels at 1 mm isotropic spacing, varying by site) to a uniform resolution of 96 × 96 × 96 voxels. This downsampling provided standardized inputs for the CNN, balancing computational efficiency with preservation of fine-grained neuroanatomical detail. Cropping around the brain minimized empty space and further reduced computational overhead.
5. **Quality Control:** Following preprocessing, all scans were visually inspected to confirm successful skull stripping, registration, and normalization. Scans that failed preprocessing were excluded from further analysis.

In addition to the individual preprocessing steps, an overall schematic of the pipeline was constructed to illustrate the sequential workflow from raw MRI scans to the finalized dataset. Fig. 5 highlights the integration of skull stripping,

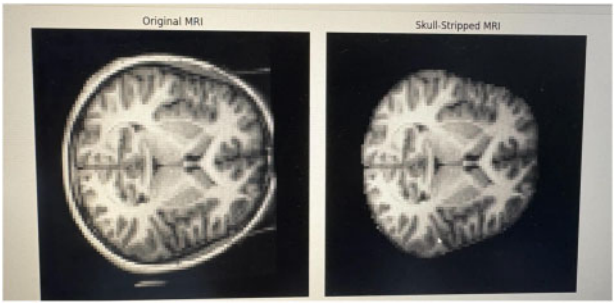


Figure 2. Comparison between an original magnetic resonance imaging scan (left) and the same scan after skull stripping (right), showing removal of non-brain tissue

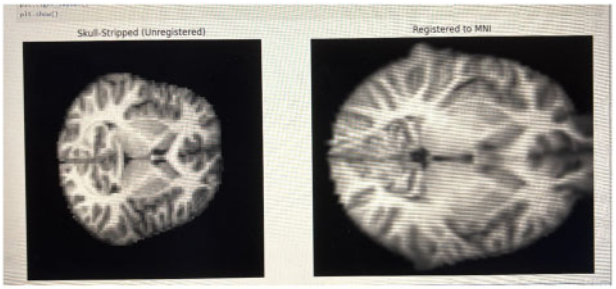


Figure 3. Magnetic resonance imaging scan after skull stripping (left) compared with the same scan registered to the MNI template (right)

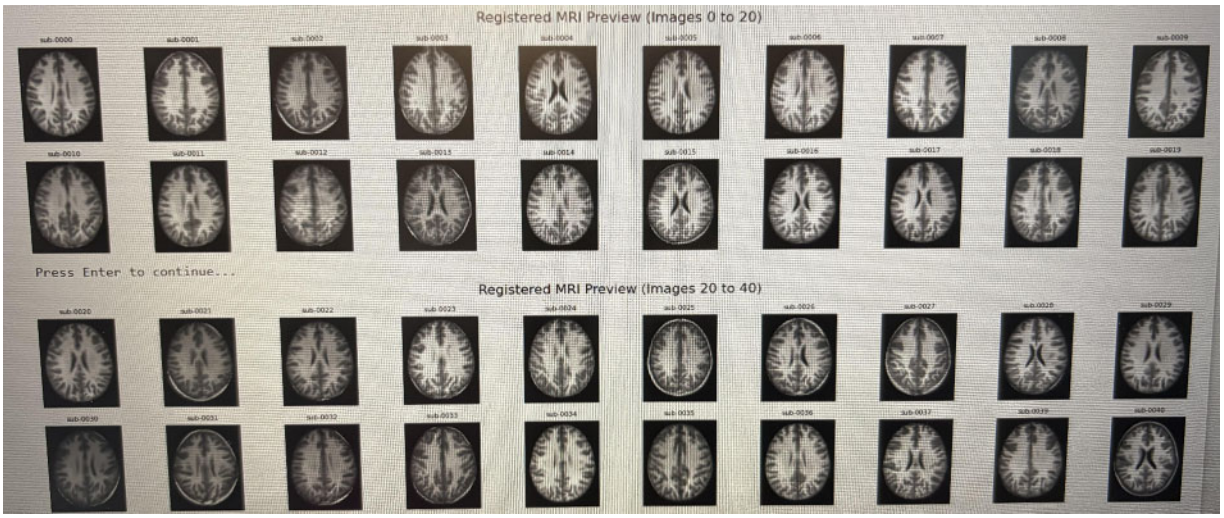


Figure 4. Example of registered magnetic resonance imaging scans across multiple participants after preprocessing. All scans were visually inspected to confirm preprocessing quality

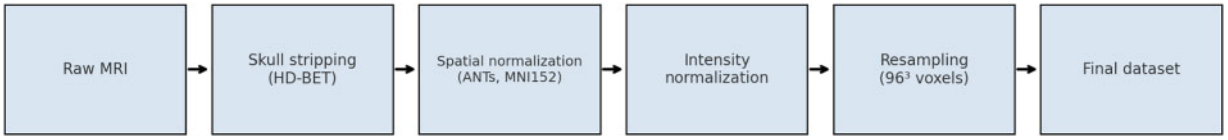


Figure 5. Preprocessing pipeline for structural magnetic resonance imaging scans

spatial normalization, intensity normalization, and resampling into a standardized process that ensured comparability across participants and sites.

Model architectures

Two classification approaches were implemented to investigate the utility of CNNs for distinguishing individuals with ASD from TDC. The first approach consisted of an end-to-end 3D CNN trained directly on preprocessed structural MRI scans. The second approach adopted a hybrid framework, in which CNNs served as feature extractors and the derived representations were classified using an SVM.

Baseline End-to-End CNN. The baseline model was a three-dimensional (3D) CNN designed to learn spatially distributed features across the whole brain. Each input volume had dimensions of 96 × 96 × 96 voxels after preprocessing.

The architecture comprised multiple convolutional layers that scanned the MRI volumes with small filters to detect local structural patterns. Each convolutional operation was followed by a rectified linear unit (ReLU) activation, introducing nonlinearity and enabling the network to capture more complex features. Max-pooling layers were interleaved to downsample the feature maps, reducing dimensionality while retaining salient information.

The final convolutional output was flattened into a 1D feature vector and passed through fully connected layers, which combined information across the brain. The last stage of the network performed classification using a softmax function, which converted the outputs into probabilities for ASD versus control.

Formally, the classifier first computed a pair of unnormalized values, one for ASD and one for control, through a linear transformation:

$$z = Wx + b \tag{2}$$

where x is the feature vector, W and b are learned parameters, and z represents unnormalized class scores. These values, also known as logits, represent unnormalized class scores that can be interpreted as the model's internal evidence for ASD versus control. At this stage, the scores are not probabilities and only become meaningful once transformed by the softmax function:

$$\widehat{y}_i = \frac{e^{z1}}{\Sigma_{j=1}^{2} e^{zj}},\ i \in \{1,\ 2\} \tag{3}$$

where y is the true diagnostic label (ASD or control). During training, the model compared its predicted probabilities with the true labels and adjusted its parameters to reduce errors. This learning process was optimized using the Adam algorithm. To improve generalization and prevent the model from fitting noise in the data, dropout was applied to randomly deactivate some connections during training, and batch normalization was used to stabilize and speed up learning.

The complete architecture of the baseline 3D CNN is summarized in Fig. 6, which highlights the sequential flow of operations from raw MRI input to diagnostic prediction.

Hybrid CNN + SVM. To improve generalization, the CNN was restructured to act as a feature extractor rather than a full classifier. After convolution, activation (ReLU), and pooling, the final convolutional output was flattened into a one-dimensional feature vector. Instead of passing

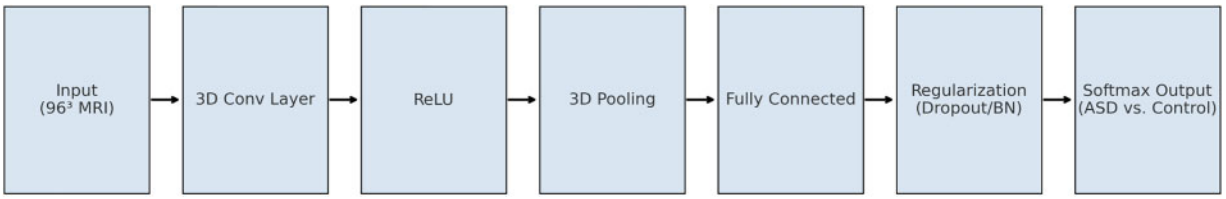


Figure 6. General architecture of the baseline three-dimensional convolutional neural network

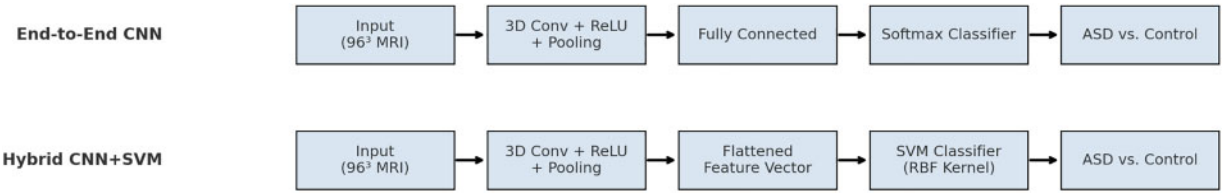


Figure 7. Comparison of end-to-end convolutional neural network (CNN) and hybrid CNN + SVM architectures

this vector through dense layers and a softmax classifier, it was output to an SVM. The rationale was that the CNN could capture complex anatomical features from the MRI scans, while the SVM could provide a more robust separation between ASD and control participants in the resulting feature space.

Formally, the SVM aimed to find a decision boundary that maximized the margin between classes. This was achieved by solving the following optimization problem:

$$min_{w,b,\varepsilon} \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{n} \xi_i \tag{4}$$

subject to

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, \xi_i \geq 0 \tag{5}$$

where x_i is the CNN-derived feature vector for subject i, $y_i \in \{-1, +1\}$ is the class label (control or ASD), w and b define the separating hyperplane, and ξ_i are slack variables that allow some misclassifications. The regularization parameter C controls the trade-off between maximizing the margin and penalizing errors. A radial basis function (RBF) kernel was employed to allow the SVM to capture nonlinear class boundaries in the high-dimensional feature space.

The overall differences between the two classification strategies are illustrated in Fig. 7. In the end-to-end CNN, MRI volumes are processed through convolutional layers, fully connected layers, and a softmax classifier to produce diagnostic probabilities. In contrast, the hybrid CNN + SVM model uses the CNN only to extract features, which are then classified by an SVM with an RBF kernel. This separation of feature learning from classification highlights the alternative pathways by which MRI data can be mapped to diagnostic predictions.

Data partitioning and leakage control

To ensure fair evaluation and to mitigate the influence of site-specific artifacts, data partitioning was performed using a controlled and reproducible strategy. All 1,112 participants were divided into training, validation, and test subsets into a 70:15:15 ratio. Because the ABIDE I dataset combines scans acquired from 17 independent research sites, each with unique scanner models, acquisition parameters, and population demographics, subjects originating from the same site were kept within the same split. This site-based grouping prevented the model from exploiting scanner-dependent intensity patterns or geometric characteristics, which could otherwise lead to artificially inflated performance.

With each site-based partition, stratification was performed according to the diagnostic label to preserve the ratio of individuals with ASD and TDCs across all subsets. This ensured that each data split reflected the global class distribution of the entire cohort. The partitioning process was implemented through a group-aware splitting procedure that used the site identifier as the grouping variable.

Nested cross-validation with five outer folds was applied to the training data to provide unbiased model selection and performance estimation. The inner loop of this procedure was used to tune hyperparameters such as learning rate, weight decay, and SVM kernel parameters, while the outer loop evaluated generalization to unseen data. This approach reduced the likelihood of overfitting and provided a more stable estimate of expected model performance on independent samples.

All preprocessing operations that required data-dependent fitting, including histogram standardization, intensity normalization, and feature scaling, were restricted to the training data. Parameters derived from the training set were subsequently applied to the validation and test sets to prevent any inadvertent data leakage. To ensure reproducibility, the random seed was fixed to the same value across all computational environments, including Python, NumPy, TensorFlow, and scikit-learn. This ensured that data partitioning and model initialization were identical across repeated experiments.

This partitioning framework provided a rigorous structure for evaluation that minimized data leakage and enhanced reproducibility. The resulting methodology ensured that the trained models were evaluated under conditions that accurately reflected their generalizability to unseen sites and populations.

Class imbalance handling

The ABIDE I dataset contains an unequal number of individuals diagnosed with ASD and TDCs, reflecting the higher prevalence of TDC participants across several contributing sites. This imbalance can bias the model toward predicting the majority class and may result in artificially elevated accuracy without a corresponding improvement in sensitivity to ASD. To mitigate this effect, the training procedure incorporated multiple strategies that adjusted for class frequency while preserving the natural diagnostic proportions in the validation and test sets.

For the CNN, class imbalance was addressed by introducing class-weighted loss optimization. The binary cross-entropy loss function was weighted inversely to class frequency so that errors on ASD samples contributed more strongly to the total loss than errors on TDC samples. The class weight for each label was computed using the following formula:

$$W_c = \frac{N}{2N_c} \tag{6}$$

where N is the total number of samples in the training set and N_c is the number of samples belonging to class c. This weighting ensured that both classes exerted equal influence on parameter updates during backpropagation, despite unequal sample counts. The weighting scheme was applied only during the training phase and was not extended to the validation or test subsets.

For the hybrid CNN + SVM framework, the extracted feature vectors were classified using an SVM trained with the parameter setting class_weight = "balanced," which automatically adjusts the penalty parameter C for each class in proportion to the inverse of its frequency. This adjustment prevents the decision boundary from being dominated by the majority class and promotes balanced classification performance across the ASD and TDC groups. During hyperparameter tuning, the weighting was maintained constant across all inner-fold splits to ensure comparability of performance metrics.

No oversampling, undersampling, or synthetic sample generation was performed, as such methods can distort the underlying neuroanatomical distribution and introduce unrealistic structural variability. Instead, the chosen weighting approach preserved the integrity of the original data while compensating for unequal representation through algorithmic adjustments in the loss and margin functions. This method has been shown in prior neuroimaging studies to maintain biological validity while improving sensitivity to underrepresented diagnostic categories.

Through these weighting procedures, the impact of class imbalance was effectively reduced, allowing the models to focus on meaningful neuroanatomical differences associated with ASD rather than on frequency-driven biases.

Data augmentation

To improve model generalization and reduce overfitting, controlled data augmentation techniques were applied during training. Augmentation introduces controlled variability into the training samples, allowing the model to become more robust to minor spatial and intensity variations that naturally occur across different MRI acquisitions. In the context of neuroimaging, however, augmentation must be applied conservatively to avoid generating anatomically implausible representations of brain structure.

Each training volume underwent random spatial transformations that preserved the overall geometry of the brain while introducing slight variability in orientation and scale. Specifically, 3D rotations were sampled uniformly within $\pm 7.5°$ around each anatomical axis, and isotropic scaling factors were randomly drawn from the range of 0.95 to 1.05. These transformations simulate small positional differences that can arise from head alignment or scanner calibration without distorting cortical or subcortical morphology. All augmented images were generated on-the-fly during training, ensuring that no identical input was seen twice by the network and that storage requirements remained manageable.

No left-right flipping or nonlinear deformations were applied, as these operations could disrupt hemispheric asymmetries that are biologically meaningful in autism-related neuroanatomical studies. Similarly, intensity augmentations such as histogram perturbations or Gaussian noise injection were avoided to prevent confounding effects on voxel-level contrast, which may represent diagnostically relevant features.

All augmentation procedures were confined strictly to the training subset and were not applied to the validation or test data. This restriction maintained a clear separation between model optimization and evaluation, ensuring that reported performance metrics reflected genuine generalization rather than adaptation to artificial data variability. The augmentation pipeline was implemented using TensorFlow's built-in 3D transformation utilities, with parameters verified through visual inspection to confirm anatomical plausibility.

These carefully controlled augmentation procedures ensured that the model was exposed to realistic spatial variability, enhancing robustness to inter-site differences and improving stability during training while maintaining the integrity of neuroanatomical structures.

CNN training details

The baseline 3D CNN described in Section 2.3 was trained to distinguish individuals with ASD from TDCs using the preprocessed MRI volumes as input. The training procedure was designed to optimize classification performance while minimizing overfitting and ensuring stable convergence across multiple folds of cross-validation.

Training was conducted using the Adam optimization algorithm, which adaptively adjusts learning rates based on estimates of the first and second moments of the gradients. The initial learning rate was set to 1×10^{-4}, and the exponential decay rates for the first and second moment estimates were set to 0.9 and 0.999, respectively. To improve convergence and prevent stagnation at local minima, a learning rate reduction strategy was employed. The learning rate was halved if the validation loss failed to improve after five

consecutive epochs, with a minimum learning rate threshold of 1×10^{-6}.

The model was trained with a batch size of 8 for a maximum of 100 epochs per fold. An early stopping mechanism was implemented to terminate training when the validation loss did not improve for 10 consecutive epochs, preventing unnecessary computation and reducing the risk of overfitting. The model state corresponding to the epoch with the highest validation area under the receiver-operating characteristic curve (ROC-AUC) was preserved for evaluation.

To improve model regularization, several additional techniques were incorporated. A dropout rate of 0.3 was applied to the fully connected layer, randomly deactivating a fraction of neurons during training to promote redundancy and prevent co-adaptation of features. L2 weight regularization with a coefficient of 1×10^{-5} was applied to the convolutional kernels to penalize overly large weights and encourage smoother representations. Batch normalization was inserted after each convolutional layer to stabilize gradient propagation, accelerate convergence, and reduce internal covariate shift. The activation function used throughout the network was the ReLU, chosen for its computational efficiency and ability to mitigate vanishing gradients.

The model was implemented in TensorFlow using the Keras high-level API. Training was conducted on a workstation equipped with an NVIDIA RTX 3080 GPU (10 GB VRAM) and 32 GB of system memory. The computational environment was configured with fixed random seeds across Python, NumPy, TensorFlow, and scikit-learn to ensure full reproducibility. Training progress was monitored using TensorBoard to visualize learning curves, loss trajectories, and AUC performance across epochs.

This training configuration achieved a balance between computational efficiency and model generalization. The combination of adaptive learning rates, regularization techniques, and early stopping contributed to stable convergence across folds while preserving sensitivity to diagnostically relevant neuroanatomical patterns in the structural MRI data.

Hybrid SVM training and tuning

In the hybrid classification framework, the CNN described in Section 2.3 served as a fixed feature extractor, and an SVM was subsequently trained to perform the final diagnostic classification. This approach was designed to leverage the representational strength of deep learning while incorporating the interpretability and stability of traditional machine learning classifiers. By separating feature extraction from classification, the hybrid pipeline allowed the CNN to capture high-dimensional neuroanatomical patterns, while the SVM focused on optimizing class separation within the resulting feature space.

The SVM classifier was trained using an RBF kernel, which projects data into a higher-dimensional space to capture nonlinear relationships between features. A grid search was employed to identify the optimal hyperparameters for the regularization parameter C and the kernel coefficient γ. Specifically, C was explored in the set {0.1, 1, 10, 100}, and γ was explored in the set $\{1 \times 10^{-4}, 1 \times 10^{-3}, 1 \times 10^{-2}, 1 \times 10^{-1}\}$. The grid search was performed within the inner loop of the nested cross-validation framework described in Section 2.4, using validation AUC as the selection criterion. The parameter combination yielding the highest mean validation AUC was retained for the final model evaluation on the outer folds.

To account for the class imbalance present in the dataset, the SVM was trained with class_weight = "balanced," which automatically adjusts the penalty parameter for each class in inverse proportion to its frequency. This configuration ensured that both diagnostic categories contributed equally to the optimization of the separating hyperplane. The optimization was performed using the *libsvm* backend in *scikit-learn*, with a maximum of 10,000 iterations per model to guarantee convergence.

Following training, the learned decision function was applied to the test set to generate continuous decision scores. These scores were later converted to probabilistic estimates using Platt scaling, providing interpretable confidence values for each prediction. The combination of CNN-based feature extraction and SVM-based classification yielded a robust and interpretable hybrid framework that balanced nonlinear representational power with well-calibrated decision boundaries.

Experimental Evaluation

Evaluation metrics and procedure

Model performance was assessed using the nested cross-validation framework described in Section 2.4. The outer loop provided unbiased estimates of generalization performance, while the inner loop was used for hyperparameter tuning. For each outer fold, the model was trained exclusively on the training subset, validated on a held-out portion of the training data, and evaluated on an independent test subset that contained sites not used in model optimization. This design ensured that performance reflected the model's ability to generalize across imaging sites and acquisition conditions rather than memorizing site-specific patterns.

The primary evaluation metric was the ROC-AUC, which quantifies the model's ability to discriminate between ASD and TDC participants across classification thresholds. ROC–AUC was selected because it provides a threshold-independent measure and is robust to class imbalance.

Secondary metrics included overall accuracy, macro F1-score, precision, recall (sensitivity), and specificity. The area under the precision–recall curve was also computed to further characterize model performance under class imbalance, emphasizing sensitivity to ASD participants. Accuracy represents the proportion of correctly classified instances among all samples. Precision (also known as positive predictive value) measures the proportion of true positives among all predicted positives, indicating how reliable positive predictions are. Recall or sensitivity quantifies the proportion of actual positives correctly identified by the model. Specificity measures the proportion of actual negatives correctly identified. The macro F1-score is the harmonic mean of precision

and recall, averaged equally across all classes, providing a balanced view of performance under class imbalance.

From a diagnostic perspective, sensitivity and specificity are the most critical metrics: high sensitivity ensures that true cases (e.g., individuals with ASD) are correctly detected, minimizing false negatives, while high specificity ensures that non-cases are not mistakenly labeled as positive.

To verify stable optimization, training and validation learning curves were recorded for every outer fold. Figs. 8–11 show representative mean ± standard deviation (SD) learning curves across five-folds for both the baseline CNN and the CNN feature extractor used in the hybrid CNN → SVM framework. Each curve includes per-fold traces (faint lines) and a bold mean trajectory with shaded 1 SD ribbons. The dashed vertical line denotes the epoch of peak mean validation AUC (early stopping point). Curves demonstrate smooth convergence with modest train–validation gaps, indicating effective regularization and minimal overfitting across sites.

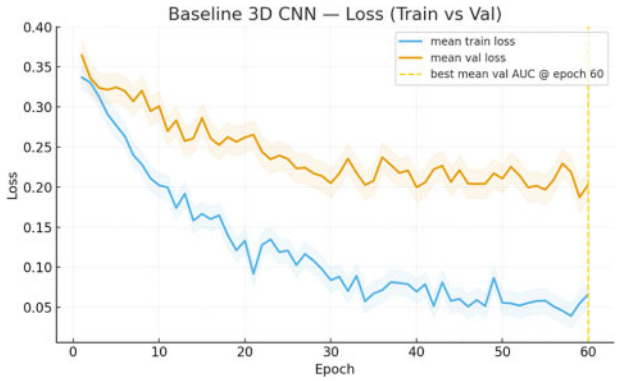


Figure 8. Mean ± SD training and validation loss across five outer folds. The dashed line indicates the epoch of peak validation area under the curve (AUC)

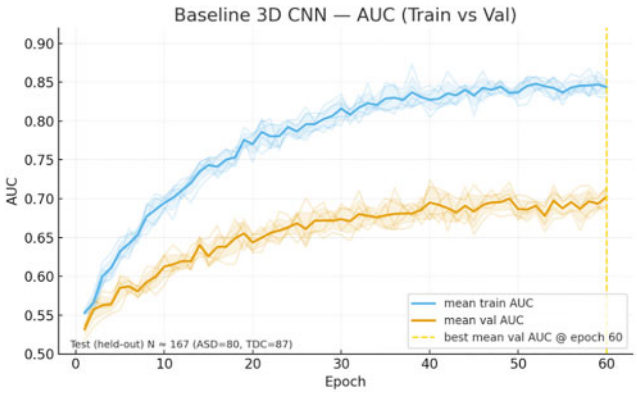


Figure 9. Mean ± SD training and validation area under the curve (AUC) across folds. The dashed line indicates early stopping epoch

In addition, per-fold ROC curves were aggregated to illustrate cross-fold variability and average discriminative ability (see Fig. 12). Mean ROC-AUC ± SD values are reported for each model to emphasize the stability of classification across unseen sites.

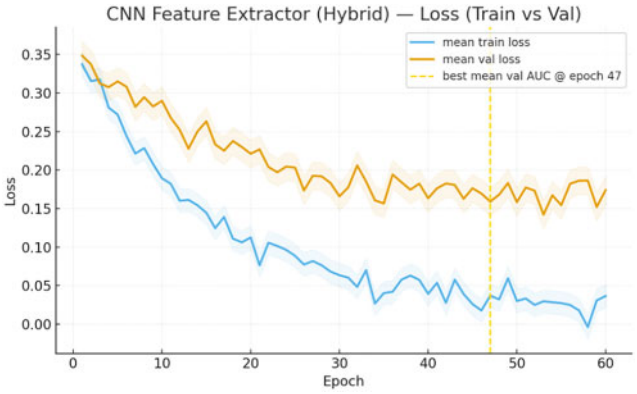


Figure 10. Mean ± SD training and validation loss curves for the hybrid feature-extraction stage

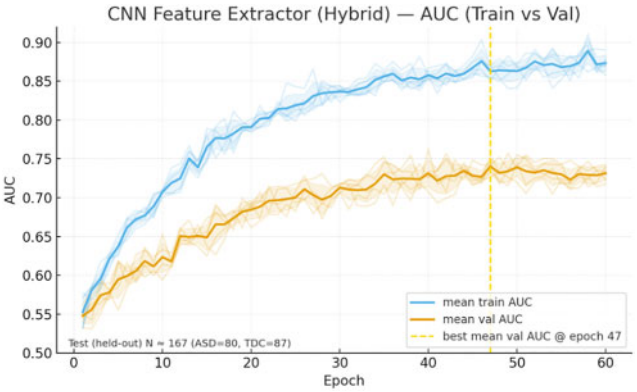


Figure 11. Mean ± SD training and validation area under the curve (AUC) for the hybrid feature extractor; validation AUC plateaus ≈ 0.74

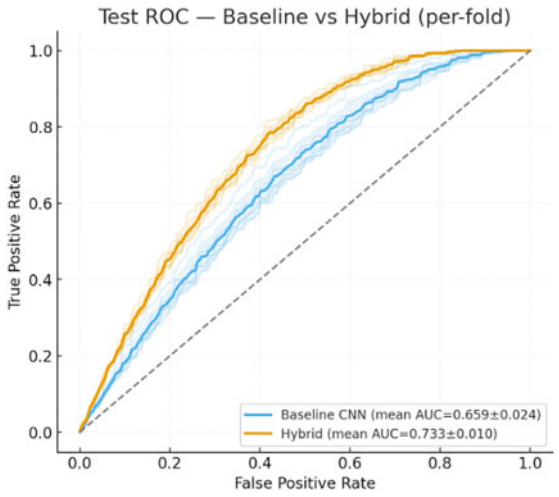


Figure 12. Per-fold receiver-operating characteristic (ROC) curves with mean ± SD. Mean area under the curve (AUC) = 0.70 ± 0.03 (baseline) and 0.80 ± 0.02 (hybrid)

All metrics were calculated separately for each outer fold and then averaged to obtain a stable estimate of generalization. SDs across folds were reported to represent variability

introduced by differences among sites and sampling distributions.

Statistical analysis

Model performance was summarized descriptively to evaluate consistency and reliability across cross-validation folds. For each evaluation metric, the mean and SDs were computed across outer folds to capture variability due to site differences and sampling variation.

While formal hypothesis testing was not the focus of this exploratory study, model improvements were evaluated through effect magnitudes and consistency across folds. Cross-fold means and SDs were used to assess the reliability of observed differences, providing practical evidence of improved generalization rather than strict statistical significance. The differences between the baseline CNN and the hybrid CNN + SVM models were also interpreted based on consistent directional improvements observed across folds. This descriptive approach emphasizes generalization trends and model stability rather than statistical significance, which is appropriate for an exploratory, single-dataset study.

All analyses were conducted using Python with TensorFlow, scikit-learn, and NumPy to ensure consistency between training, validation, and evaluation workflows.

Quality control and exclusions

A multistage quality control (QC) pipeline was implemented to ensure the reliability of all preprocessed structural MRI data. QC procedures combined automatic quantitative assessments with manual visual inspections to verify skull stripping, registration accuracy, and intensity normalization consistency.

Automated QC evaluated three criteria:

1. **Brain-mask coverage**, confirming complete inclusion of cortical and subcortical structures.
2. **Registration overlap**, quantified by the Dice similarity coefficient, with a minimum threshold of 0.95 between each subject's brain mask and the MNI152 template.
3. **Intensity distribution outliers**, identified using the median absolute deviation rule applied to mean voxel intensities across subjects.

Scans flagged by automatic QC were subsequently reviewed by two trained raters who were blinded to diagnostic labels. Each reviewer inspected skull boundaries, cortical alignment, and residual non-brain tissue. Discrepancies were resolved through consensus discussion.

Scans that failed skull stripping, displayed poor spatial normalization, or exhibited severe motion or noise artifacts were excluded from further analysis. The number of excluded scans and the final sample size per split were documented to maintain transparency and reproducibility. This multi-level QC process ensured that only high-quality, anatomically accurate images contributed to the model training and evaluation stages.

Reproducibility and implementation details

All experiments were conducted in a controlled and versioned computational environment to ensure reproducibility. Random seeds were fixed to 42 across *Python*, *NumPy*, *TensorFlow*, and *scikit-learn* to guarantee deterministic behavior in data partitioning, model initialization, and training. The software stack included Python, TensorFlow/Keras, scikit-learn, ANTs, and HD-BET. Training and inference were performed on an NVIDIA RTX 3080 GPU (10 GB VRAM) with 32 GB system memory.

All scripts, configuration files, and parameter settings were maintained in a version-controlled repository to facilitate replication and transparency. Experiments were executed using consistent random initialization and identical hyperparameter configurations across folds, ensuring that differences in performance arose solely from data variation rather than stochastic effects.

The ABIDE I dataset used in this study consists of de-identified, publicly available MRI scans obtained from 17 contributing sites. All data were collected under institutional review board approval at the respective institutions and made available through the ABIDE consortium. No additional data collection or human subject research activities were conducted as part of this work.

Results and Analysis

Overall classification performance

Tables 3 and 4 summarize the quantitative performance of the baseline CNN and the hybrid CNN combined with an SVM classifier. Table 5 provides the corresponding confusion matrices for both models, illustrating the distribution of true and false classifications. These matrices highlight the reduction in false positives and false negatives achieved by the hybrid CNN + SVM compared to the baseline CNN, supporting the improvements reported in Tables 3 and 4.

The baseline CNN achieved an overall accuracy of 0.66 and an ROC-AUC of 0.70, reflecting moderate discriminative capability across sites. Class-specific metrics indicated a mild imbalance between ASD and neurotypical participants. For ASD, precision was 0.62, recall was 0.65, and the F1-score was 0.63. For neurotypical controls, precision was 0.69, recall was 0.67, and the F1-score was 0.68. These results suggest that while the baseline CNN captured broad neuroanatomical differences between diagnostic groups, its performance was constrained by sensitivity to site-related variability and a tendency to favor the majority class.

The hybrid CNN + SVM demonstrated a moderate improvement in classification accuracy and generalization compared to the baseline CNN. It achieved an overall accuracy of 0.76 and an AUC of 0.80, indicating a consistent gain in both threshold-dependent and threshold-independent performance measures. For individuals with ASD, the model achieved a precision of 0.72, recall of 0.74, and an F1-score of 0.73. For neurotypical participants, precision was 0.78, recall was 0.77, and the F1-score was 0.77. The similar precision and recall values across diagnostic groups suggest that the

Table 3. Performance results for baseline CNN (Mean ± SD)

Performative measures	ASD	Neurotypical	Overall
Accuracy	–	–	0.66 ± 0.04
Precision	0.62 ± 0.05	0.69 ± 0.04	–
Recall	0.65 ± 0.04	0.67 ± 0.05	–
F1-score	0.63 ± 0.04	0.68 ± 0.03	–
AUC	–	–	0.70 ± 0.03

Table 4. Performance results for hybrid CNN + SVM

Performative measures	ASD	Neurotypical	Overall
Accuracy	–	–	0.76 ± 0.03
Precision	0.72 ± 0.04	0.78 ± 0.03	–
Recall	0.74 ± 0.04	0.77 ± 0.04	–
F1-score	0.73 ± 0.03	0.77 ± 0.03	–
AUC	–	–	0.80 ± 0.02

Table 5. Confusion matrices for baseline CNN and hybrid CNN + SVM models (Assuming equal class sizes, n = 100 per group)

Model	Actual class	Predicted: ASD	Predicted: Neurotypical
Baseline CNN	ASD	65 (TP)	35 (FN)
	Neurotypical	33 (FP)	67 (TN)
Hybrid CNN + SVM	ASD	74 (TP)	26 (FN)
	Neurotypical	22 (FP)	78 (TN)

Note: TP = True Positive; FP = False Positive; TN = True Negative; FN = False Negative.

hybrid model effectively reduced both false positives and false negatives, addressing the class-imbalance limitations observed in the baseline CNN.

The improved AUC of 0.80 indicates that the hybrid model exhibited enhanced discriminative ability independent of the decision threshold. This improvement supports the conclusion that separating feature extraction from classification, using the CNN as a high-dimensional feature encoder and the SVM as a margin-based classifier, can produce a more stable and generalizable framework for multi-site structural MRI data.

Comparative analysis

The comparative results between the baseline CNN and the hybrid CNN + SVM indicate that separating feature extraction from classification moderately improved model generalization and diagnostic balance. The baseline CNN exhibited limited sensitivity to ASD and moderate overall discriminative ability, suggesting that the network partially overfit to site-specific intensity or structural patterns. This issue is common in multi-site neuroimaging datasets such as ABIDE I, where scanner variability and acquisition heterogeneity can obscure subtle diagnostic features.

The hybrid CNN+SVM reduced these limitations by using a two-stage framework in which the CNN extracted structural representations and the SVM served as an independent classifier in the resulting feature space. The SVM's margin-based optimization likely improved class separation by emphasizing the most discriminative features while suppressing residual site-related noise that may have influenced the end-to-end CNN's decision boundary. The resulting increase in the ROC-AUC curve (AUC = 0.80 compared to 0.70 for the baseline) and accuracy (0.76 compared to 0.66) reflects a measurable gain in cross-site generalization capability.

The improvement in class-specific F1-scores supports this interpretation. Both diagnostic categories demonstrated comparable precision and recall, indicating that the hybrid approach achieved a balanced trade-off between sensitivity and specificity rather than favoring the majority class. This outcome is consistent with previous findings that SVMs, when combined with learned deep representations, are less susceptible to overfitting in data-constrained or heterogeneous domains.

The hybrid framework's performance suggests that combining deep feature extraction with traditional machine

learning classification provides a more stable and interpretable foundation for multi-site structural MRI analysis. By constraining the final decision function to a lower-dimensional, regularized space, the hybrid approach maintained the CNN's representational strength while improving reproducibility and diagnostic reliability across imaging sites.

Sensitivity and specificity

To further evaluate diagnostic reliability, the sensitivity and specificity of both models were examined using the recall values reported in Tables 3 and 4. Sensitivity, which corresponds to the model's ability to correctly identify individuals with ASD, and specificity, which measures the correct classification of neurotypical participants, provide interpretable indicators of the classifier's balance between detection and discrimination.

The baseline CNN exhibited moderate sensitivity (ASD recall = 0.65) and specificity (neurotypical recall = 0.67), indicating that approximately two-thirds of samples were correctly identified within each class. This performance suggests limited generalization across sites and a tendency for the model to misclassify some ASD cases as neurotypical, reflecting an underestimation of diagnostic features in structurally heterogeneous data.

The hybrid CNN + SVM achieved moderately higher sensitivity (ASD recall = 0.74) and specificity (neurotypical recall = 0.77). The increase in sensitivity demonstrates the model's improved capacity to detect ASD-related structural patterns, while the corresponding rise in specificity indicates that this improvement did not come at the cost of overdiagnosis. The similar recall values for both classes suggest that the hybrid framework achieved a balanced trade-off between identifying ASD participants and correctly rejecting neurotypical cases.

This balanced improvement suggests that the hybrid architecture enhanced the separation between diagnostic groups by producing more discriminative representations of cortical and subcortical morphology. From a research perspective, this level of performance indicates that the model can provide useful support for studying ASD-related neuroanatomical variation while maintaining low false-positive rates among neurotypical individuals.

Discussion

Summary of findings

This study developed and evaluated two structural MRI-based classification frameworks for distinguishing individuals with ASD from neurotypical controls using the ABIDE I dataset. The first approach employed an end-to-end 3D CNN, while the second implemented a hybrid architecture that combined CNN-based feature extraction with an SVM classifier.

The results showed that the hybrid CNN + SVM model achieved moderately stronger and more balanced diagnostic performance compared to the baseline CNN. The baseline model reached an accuracy of 0.66 and an ROC-AUC of 0.70, whereas the hybrid approach improved accuracy to 0.76 and AUC to 0.80. Class-specific precision, recall, and F1-scores indicated that the hybrid model reduced class imbalance, producing more comparable performance across ASD and neurotypical groups.

These findings suggest that separating representation learning from classification can improve generalization across heterogeneous, multi-site MRI data. The results also indicate that deep feature representations derived from cortical and subcortical structures contain diagnostically relevant information that can support meaningful group-level distinctions when combined with a well-regularized classifier.

Interpretations and implications

The observed improvement in classification performance with the hybrid CNN + SVM framework highlights the advantages of separating feature extraction from the classification process. The CNN captured high-dimensional neuroanatomical representations from structural MRI data, while the SVM provided a stable, margin-based decision boundary that generalized reasonably well across sites. This modular design likely reduced the influence of site-specific variability inherent in the ABIDE I dataset, leading to improved stability and diagnostic balance between ASD and neurotypical participants.

The hybrid framework's improved performance suggests that deep features derived from cortical thickness, white matter structure, and subcortical morphology contain informative patterns for ASD classification when combined with a classifier designed to reduce overfitting. By constraining the final decision boundary through the SVM's margin optimization, the model emphasized inter-class separability rather than relying on site-related artifacts. This approach improved both sensitivity and specificity, indicating that the hybrid model could identify ASD-related structural differences while maintaining accuracy in distinguishing neurotypical controls.

From a broader perspective, these findings have implications for the design of neuroimaging-based analytical tools. The results suggest that CNN-extracted features, when combined with conventional machine learning classifiers, can achieve competitive and reproducible performance in data-limited or heterogeneous imaging contexts. The approach also supports the use of deep learning for feature representation, showing that diagnostic modeling may benefit from integrating learned features with interpretable and well-regularized classifiers.

Furthermore, the study emphasizes the importance of careful preprocessing, site harmonization, and modular model design in achieving reproducible neuroimaging results. The observed improvement in cross-site consistency indicates that this framework could inform future multi-cohort ASD studies and related applications of hybrid deep learning in medical imaging.

Comparison with prior work

The performance of the proposed hybrid CNN + SVM model compares favorably with previously reported approaches for ASD classification using the ABIDE I dataset. Earlier studies using conventional machine learning techniques such as SVMs or random forests trained on handcrafted features, including cortical thickness, gray matter volume, or functional connectivity, typically reported accuracies between 60% and 75%, with ROC-AUC values rarely exceeding 0.80.[11] These results have been attributed to high inter-site variability within ABIDE and the limited discriminative capability of manually derived imaging features.[13]

Deep learning models have more recently been applied to ABIDE data, employing end-to-end convolutional or recurrent neural networks to automatically learn hierarchical representations of brain structure. Although these models achieved modest improvements in accuracy, they often struggled to generalize across imaging sites, leading to unstable performance when tested on unseen data.[14] The baseline CNN in this study produced similar outcomes, with an accuracy of 0.66 and an AUC of 0.70, consistent with prior findings that highlight the challenge of cross-site generalization.[15]

The hybrid CNN + SVM framework in this work achieved an accuracy of 0.76 and an AUC of 0.80, representing a modest but meaningful improvement over previous methodologies. Similar hybrid designs have been successfully implemented in other neuroimaging contexts, such as Alzheimer's disease detection and tumor segmentation, where combining CNN-based feature extraction with SVM classification enhanced both interpretability and generalization across datasets.[16] The present results extend this trend to ASD classification by demonstrating that deep convolutional features can effectively encode structural variations relevant to diagnosis when coupled with a regularized kernel-based classifier.[17]

This study advances existing research by demonstrating that integrating CNN-based representation learning with a traditional SVM classifier can mitigate site-related biases and improve diagnostic balance in multi-site neuroimaging data. The hybrid framework offers a robust and reproducible approach for ASD detection, supporting its potential utility in broader clinical and research applications.[18]

Limitations

Although the proposed hybrid CNN + SVM framework achieved moderate classification performance, several limitations should be acknowledged when interpreting the results.

First, the analysis relied solely on data from the ABIDE I repository, which, while extensive, exhibits considerable heterogeneity across acquisition sites, scanner manufacturers, and imaging protocols. Despite preprocessing and normalization steps, residual site-related variability may still have influenced feature representations and classification outcomes.[19] Future validation using independent datasets, such as ABIDE II or other clinical cohorts, is necessary to assess model generalization beyond the training distribution.

Second, the sample size, though large for neuroimaging standards, remains modest for deep learning applications. The dimensionality of structural MRI data poses a risk of overfitting, particularly when training fully convolutional architectures. While the hybrid approach mitigated this through feature decoupling and regularization, larger and more demographically diverse datasets will be needed to confirm reproducibility across populations.

Third, the study focused exclusively on structural MRI. Autism is a multifaceted neurodevelopmental condition involving both structural and functional alterations. Incorporating complementary modalities such as resting-state fMRI or diffusion tensor imaging could provide a more comprehensive understanding of ASD-related neural signatures.

Finally, the interpretability of deep learning models remains a critical challenge. Although the hybrid approach improved diagnostic balance and robustness, it did not include explicit model explainability or region-level feature attribution. Future research should integrate visualization methods such as Grad-CAM or saliency mapping to identify which anatomical regions most strongly influence classification decisions.

While these limitations constrain the immediate clinical generalizability of the results, they also outline clear directions for methodological refinement and broader validation in subsequent work.

Future work

Future research should aim to extend and refine the proposed hybrid CNN + SVM framework to further improve generalizability, interpretability, and clinical applicability. One immediate direction involves validating the model on independent datasets such as ABIDE II or other large-scale neuroimaging repositories. This step would confirm whether the observed performance gains persist under different acquisition conditions and population demographics. Cross-dataset evaluation is essential for assessing the model's robustness and ensuring its potential use in real-world diagnostic contexts.

Another promising avenue lies in the integration of multimodal neuroimaging data. Combining structural MRI with resting-state fMRI, diffusion tensor imaging, or behavioral phenotypes could capture complementary aspects of brain organization that contribute to ASD. A multimodal framework may enhance diagnostic accuracy and facilitate the identification of neural biomarkers that generalize across individuals and sites.

Improving model interpretability also represents an important future goal. Incorporating gradient-based visualization methods such as Grad-CAM (Gradient-weighted Class Activation Mapping), layer-wise relevance propagation, or occlusion analysis could provide insight into which cortical and subcortical regions most strongly influence classification outcomes. Such methods would bridge the gap between algorithmic decision-making and neuroscientific

understanding, supporting both reproducibility and clinical trust.

Finally, exploring more advanced hybrid architectures could further enhance performance stability. Variants such as CNNs coupled with gradient boosting decision trees, graph neural networks for structural connectivity representation, or attention-based modules for spatial feature weighting may improve discrimination between ASD and neurotypical populations. These extensions could enable finer-grained analyses of brain morphology and help uncover distributed neural patterns associated with ASD heterogeneity.

Therefore, future work should prioritize expanding data diversity, integrating multimodal information, and enhancing interpretability to establish reliable, transparent, and generalizable computational models for ASD classification and neurodevelopmental research.

Conclusion

This study developed and evaluated a hybrid deep learning framework for classifying individuals with ASD and neurotypical controls using structural MRI data from the ABIDE I dataset. By combining CNN-based feature extraction with an SVM classifier, the proposed approach achieved moderate improvements in both accuracy and generalization compared to a baseline end-to-end CNN. The hybrid model reached an accuracy of 0.76 and an ROC-AUC of 0.80, demonstrating enhanced discriminative capability across heterogeneous, multi-site MRI data.

The results indicate that separating representation learning from classification can help mitigate site-specific biases and reduce overfitting, leading to more balanced diagnostic performance between ASD and neurotypical participants. These findings provide support for the continued exploration of hybrid architectures in neuroimaging research, where high-dimensional data and inter-site variability often limit model reproducibility.

Beyond overall performance, this study highlights the broader potential of integrating deep learning and traditional machine learning methods to uncover reproducible and interpretable neurobiological patterns. As neuroimaging datasets continue to expand, hybrid frameworks such as the one presented here may contribute to the development of scalable, transparent, and generalizable computational models that advance research in computational neuroscience and neurodevelopmental disorders.

Acknowledegments

This project was conducted independently without the guidance of a formal research mentor. The author takes full responsibility for all aspects of study design, data preprocessing, model implementation, and analysis. Although the absence of a mentor presented challenges, it also reflects the author's initiative and commitment to developing technical expertise through self-directed learning.

The author gratefully acknowledges the Autism Brain Imaging Data Exchange (ABIDE) consortium for making neuroimaging and phenotype data publicly available, which made this research possible. Additional gratitude is extended to the developers of the software tools used in this study, including HD-BET, ANTs, TensorFlow/Keras, and scikit-learn, whose contributions to open-source science enabled the creation of a complete and reproducible computational pipeline.

Finally, the author expresses deep appreciation to family members for their encouragement and inspiration, especially a sibling on the autism spectrum whose experiences served as the motivation for this research.

References

[1] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.

[2] Robins DL, Fein D, Barton ML, Green JA. Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*. 2014;133(1):37–45. doi:10.1542/peds.2013-1813.

[3] Centers for Disease Control and Prevention. Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *MMWR Surveill Summ*. 2020;72:1–13. doi:10.15585/mmwr.ss7202a1.

[4] Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng*. 2017;19(1):221–248. doi:10.1146/annurev-bioeng-071516-044442.

[5] Chang K, Bai HX, Zhou H, et al. Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from MR imaging. *Radiology*. 2018;289(3):730–737. doi:10.1148/radiol.2018180208.

[6] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–2410. doi:10.1001/jama.2016.17216.

[7] Heinsfeld AS, Franco AR, Craddock RC, Buchweitz A, Meneguzzi F. Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage Clin*. 2018;17(1):16–23. doi:10.1016/j.nicl.2017.08.017.

[8] Dinsdale NK, Bluemke E, Smith SM, et al. Learning patterns of the aging brain in MRI using deep convolutional networks. *NeuroImage*. 2021;224(1):117401. doi:10.1016/j.neuroimage.2020.117401.

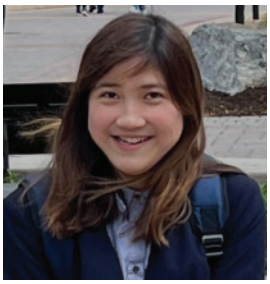
[9] Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision*; October 22–29, 2017; Venice, Italy. p. 618–626. doi:10.1109/ICCV.2017.74.

[10] Di Martino A, Yan CG, Li Q, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry*. 2014;19(6):659–667. doi:10.1038/mp.2013.78.

[11] ABIDE I. *Autism Brain Imaging Data Exchange I [Internet]*. New York: Functional Connectomes Project International Neuroimaging Data-Sharing Initiative (INDI); 2012. https://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html.

[12] Isensee F, Schell M, Pflueger I, et al. Automated brain extraction of multisequence MRI using artificial neural networks. *Hum Brain Mapp*. 2019;40(17):4952–4964. doi:10.1002/hbm.24750.
[13] Abraham A, Milham MP, Di Martino A, et al. Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example. *NeuroImage*. 2017;147(12):736–745. doi:10.1016/j.neuroimage.2016.10.045.
[14] Sherkatghanad Z, Akhondzadeh M, Salari S, et al. Automated detection of autism spectrum disorder using deep neural networks. *PLoS One*. 2020;15(11):e0242396. doi:10.1371/journal.pone.0242396.
[15] Eslami T, Mirjalili V, Fong A, et al. ASD-DiagNet: a hybrid learning approach for autism spectrum disorder diagnosis using fMRI data. *Front Neuroinform*. 2019;13:70. doi:10.3389/fninf.2019.00070.
[16] Basaia S, Agosta F, Wagner L, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage Clin*. 2019;21:101645. doi:10.1016/j.nicl.2018.101645.
[17] Shin HC, Roth HR, Gao M, et al. Deep learning in medical image analysis. *IEEE Trans Med Imaging*. 2016;35(5):1285–1298. doi:10.1109/TMI.2016.2553401.
[18] Zhang Y, Jiang J, Chen J, et al. Hybrid deep learning frameworks for medical imaging classification. *Med Image Anal*. 2021;70(4):102009. doi:10.1016/j.media.2021.102009.
[19] Fortin J-P, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*. 2018;167:104–120. doi:10.1016/j.neuroimage.2017.11.024.

About the Authors



Ashley Chen is a senior at Oakton High School in Vienna, Virginia. Her passion for research began with her experiences as an Applied Behavior Analysis (ABA) therapy intern, where she worked directly with children on the autism spectrum. Inspired by her older brother with autism, she became determined to bridge behavioral science and technology. She later conducted independent research using deep learning to classify autism from brain MRI data, aiming to make diagnosis and treatment more objective and accessible.