CIA Open Access

A Comparative Evaluation of Transfer Learning Methods for Cross-Context Behavioral Generalization Assessment in Autism Spectrum Disorder Interventions

***Yaqing Bai*[1], *Muyu Liu*[2]**

[1]*Human Development, University of Rochester, NY, USA*

[2]*Center for Applied Statistics and School of Statistics, Renmin University of China, Beijing, China*

DOI: 10.63575/CIA.2026.40115

Abstract

This study presents a comparative evaluation of three transfer learning paradigms—domain adaptation, few-shot learning, and multi-task learning—for assessing cross-context behavioral generalization in autism spectrum disorder (ASD) interventions. Behavioral skill generalization across distinct intervention environments remains a persistent barrier to long-term ASD intervention effectiveness. The distributional discrepancies inherent in behavioral data collected from clinical, school, and home settings create domain shift challenges that standard recognition models cannot adequately address. Drawing primarily on the Self-Stimulatory Behavior Dataset (SSBD) for the video-based comparative analysis, while using the Autism Brain Imaging Data Exchange (ABIDE) as an auxiliary multi-site domain-shift reference and the Expanded Stereotype Behavior Dataset (ESBD) as a published benchmark, this paper establishes an evaluation framework with four quantitative metrics: cross-context recognition accuracy, generalization stability index, few-shot adaptation efficiency, and behavioral transfer success rate. Within this benchmark-informed framework, domain adaptation yields the highest source-proximal accuracy, few-shot learning shows the strongest low-label adaptation, and multi-task learning exhibits the most stable cross-context profiles. These findings provide practical guidance for selecting computational tools to support more scalable and objective ASD intervention evaluation.

Keywords: *Transfer learning; autism spectrum disorder; behavioral generalization; cross-context evaluation*

1. Introduction

1.1. ASD Prevalence and the Generalization Challenge in Behavioral Interventions

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by persistent deficits in social communication alongside restricted, repetitive behavioral patterns. Recent CDC surveillance reports indicate that ASD prevalence among 8-year-old children in the United States remains high and has increased over time, amplifying the demand for evidence-based interventions and reliable evaluation methods that can assess whether therapeutic gains extend beyond structured treatment settings into children's daily environments.

The capacity to generalize acquired skills across contexts—from the clinical room to the classroom and the home—constitutes a defining criterion for evaluating intervention success. A systematic review and transfer analysis examining 52 original ASD social training experiments reported that intervention transferability is significantly modulated by participant age, intervention context, treatment intensity, and instructional modality, with behavioral approaches demonstrating superior transfer outcomes relative to cognitive interventions [1]. A separate systematic review of randomized controlled trials targeting early social communication interventions identified nine studies providing evidence of both initial target learning and generalization measurement, among which eight demonstrated at least partial skill transfer across people, settings, or activities [2]. Experimental evidence from a large sample of verbally fluent children and adolescents with ASD revealed that strategy transfer to novel learning contexts was significantly weaker compared with typically developing peers, confirming that generalization difficulties persist across the ability spectrum [3].

1.2. Computational Approaches and Research Objectives

The advent of dedicated behavioral datasets has enabled the application of computer vision to ASD-related behavior recognition. The Self-Stimulatory Behavior Dataset (SSBD), introduced at the ICCV 2013 Workshop, comprises 75 videos capturing three stereotypical behavior categories—arm flapping, head banging, and spinning—recorded in uncontrolled natural settings [4]. Behavioral data collected across different intervention environments exhibit substantial distributional discrepancies arising from variability in backgrounds, lighting, interaction partners, and activity structures. Transfer learning provides a family of computational techniques—including domain adaptation, few-shot learning, and multi-task learning—designed to bridge such distributional gaps between source and target domains.

This paper conducts a systematic comparative evaluation of these three paradigms for cross-context ASD behavioral generalization assessment. The contributions include: (a) an evaluation framework anchored in publicly available ASD-related datasets and benchmark resources; (b) a multi-metric, benchmark-informed comparison across domain adaptation, few-shot learning, and multi-task learning methods; and (c) a proposed set of generalization indicators—including cross-context recognition accuracy, generalization stability index, and behavioral transfer success rate—for more standardized intervention assessment.

2. Related Work

2.1. Deep Learning for ASD Behavior Recognition

Vision-based automatic behavior recognition has progressed substantially with deep learning architectures. A two-branch multimodal framework incorporating language-assisted training alongside Video Swin Transformer achieved state-of-the-art performance on both the SSBD and Expanded Stereotype Behavior Dataset (ESBD), yielding a 3.49% improvement on ESBD and 1.46% on SSBD relative to vision-only baselines, and demonstrating the viability of incorporating semantic supervision for ASD behavior classification [5]. These results established strong baselines for the behavioral recognition component of cross-context evaluation.

2.2. Transfer Learning Methods for Cross-Domain Recognition

Domain adaptation addresses distributional mismatch between source and target domains through feature-level alignment. Deep Adaptation Networks (DAN) introduced multi-kernel maximum mean discrepancy (MMD) as a domain regularization mechanism within a reproducing kernel Hilbert space, enabling alignment of deep feature distributions across domains [6]. This approach has been extended through adversarial training with gradient reversal layers, conditional alignment, and joint distribution matching strategies.

Few-shot learning addresses the challenge of adapting to novel domains with extremely limited labeled data—a condition frequently encountered in ASD intervention settings where data collection is constrained by ethical requirements and small participant pools. Prototypical Networks established a metric learning framework that constructs class-level prototype representations from a small support set and classifies query instances by Euclidean distance in an embedding space [7]. Spatio-temporal contrastive domain adaptation extended cross-domain alignment to video-based action recognition by jointly optimizing spatial feature alignment and temporal consistency constraints across domains [8]. Multi-task learning presents a third paradigm, training shared feature representations across multiple related tasks or domains simultaneously. By exposing a unified encoder to behavioral data from diverse intervention contexts during training, multi-task architectures can learn representations that are inherently less context-dependent, potentially yielding more stable cross-environment performance without explicit domain alignment.

2.3. Research Gap

Although individual transfer learning techniques have been applied to specific facets of ASD computational analysis—diagnostic classification from multi-site neuroimaging data, behavioral pattern recognition from video, and facial expression analysis for early screening—no existing study has systematically compared the performance characteristics of domain adaptation, few-shot learning, and multi-task learning within a benchmark-informed framework aimed at evaluating behavioral generalization across intervention contexts. The absence of standardized quantitative metrics for computational generalization assessment in ASD research further limits the interpretability and comparability of existing approaches. This gap motivates the present comparative study, which aims to provide more carefully delimited guidance for method selection in cross-context ASD intervention assessment.

3. Methodology: Comparative Framework for Cross-Context Generalization Evaluation

3.1. Problem Formulation and Dataset Configuration

The cross-context generalization assessment problem is formulated as a domain transfer task. Let $D_S = \{(x_i^s, y_i^s)\}_{i=1}^{n_s}$ denote labeled behavioral data from the source domain (structured clinical setting) and $D_T = \{(x_j^t)\}_{j=1}^{n_t}$ denote data from the target domain (school or home), where $P(X^s) \neq P(X^t)$ due to contextual and environmental differences. The objective is to learn a classification function $f: X \rightarrow Y$ that maintains accuracy across all target domains.

The evaluation framework draws on four publicly available ASD-related data resources. Spatio-temporal CNN frameworks applied to the SSBD dataset demonstrated that convolutional feature extractors combined with Temporal Convolutional Networks (TCN) yield robust ASD behavior classification, validating these architectures as suitable backbones for cross-context evaluation [9]. The Autism Brain Imaging Data Exchange (ABIDE I) aggregated 1,112 resting-state fMRI datasets from 539 individuals with ASD and 573 typical controls across 17 international imaging sites, establishing a large-scale open-access cross-site ASD neuroimaging resource [10]. The subsequent ABIDE II release contributed 1,044 additional datasets (487 ASD, 557 controls) from 16 institutions [11]. Because ABIDE is a diagnostic neuroimaging consortium rather than a behavioral generalization dataset, it is used here as an auxiliary reference for multi-site domain shift rather than as a direct substitute for intervention video data. Table 1 summarizes the data resources employed.

Table 1. Summary of Public Datasets Used in the Comparative Evaluation Framework

Dataset	Data Type	Total Samples	Categories	Sites/Contexts	Availability
SSBD	Video	75 videos (avg. 90s)	3 (arm flapping, head banging, spinning)	Uncontrolled settings (YouTube, Vimeo)	Public
ESBD	Video	141 videos (avg. 120s)	4 (+ hand action)	Uncontrolled settings (YouTube)	Partial
ABIDE I	fMRI + sMRI	1,112 (539 ASD + 573 TC)	Binary (ASD vs. TC)	17 international sites	Public
ABIDE II	fMRI + sMRI	1,044 (487 ASD + 557 TC)	Binary (ASD vs. TC)	16 institutions	Public

Note. TC = typical controls; sMRI = structural MRI. Dataset details and original sources are cited in the corresponding body text. SSBD serves as the primary comparative dataset in this study; ESBD and ABIDE are used as auxiliary benchmark or domain-shift references.

For video-based evaluation, SSBD videos are partitioned into simulated context subsets based on observable environmental features: Set A (clinical-like: single-subject, controlled background, 28 videos), Set B (school-like: structured activity with peers or instructor, 22 videos), and Set C (home-like: naturalistic domestic, caregiver interaction, 25 videos).

3.2. Comparative Transfer Learning Methods

Three categories of transfer learning methods are evaluated under a unified protocol. Within each data track, all compared methods share the same backbone feature extractor—Video Swin Transformer pretrained on Kinetics-400 for video data and ResNet-50 pretrained on ImageNet for neuroimaging data—to improve comparability within that track. The cross-domain prototypical network introduced domain-specific feature transformations that adapt prototype computation to distributional differences between domains, demonstrating robustness across visually distinct target domains [12]. Table 2 summarizes the defining characteristics of each method category.

Domain Adaptation (DA) employs distribution alignment between source and target feature representations. The baseline method applies multi-kernel MMD regularization following the DAN framework. An adversarial variant incorporating a gradient reversal layer encourages the feature extractor to produce domain-invariant representations.

Few-Shot Learning (FSL) adapts to new target contexts using K labeled examples per class (K = 1 or 5). The prototypical network baseline constructs class prototypes from the support set and classifies query samples by Euclidean distance. A cross-domain extension augments prototypes with learnable feature transformations conditioned on target domain statistics.

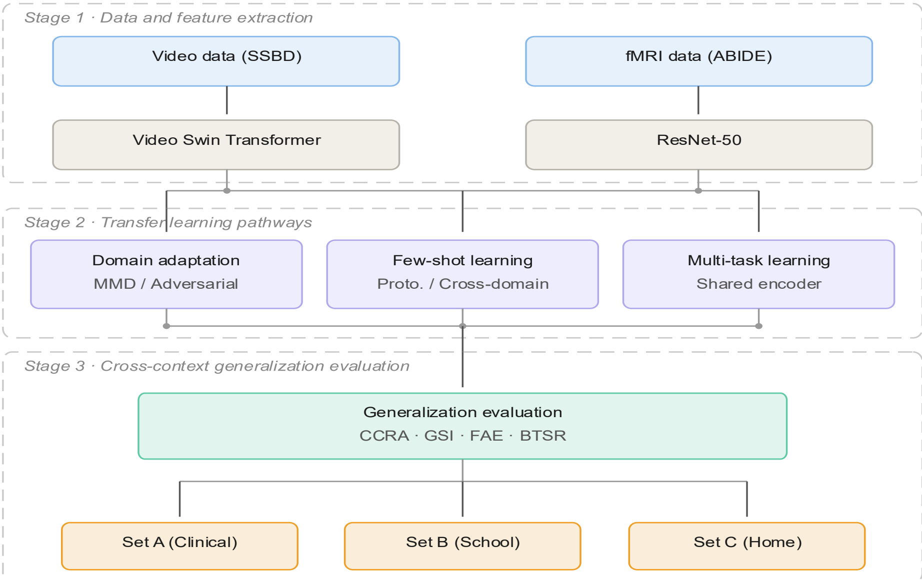
Multi-Task Learning (MTL) trains a shared feature extractor simultaneously across all context domains with separate classification heads. The shared layers capture domain-invariant features while domain-specific heads accommodate contextual variation.

Table 2. Characteristics of Compared Transfer Learning Method Categories

Characteristic	**Domain Adaptation**	**Few-Shot Learning**	**Multi-Task Learning**
Target domain labels	None or minimal	K-shot per class	Full labels, all domains
Alignment mechanism	Distribution matching (MMD, adversarial)	Metric-based prototypes	Shared representation + domain heads
Primary strength	Large-scale cross-domain transfer	Rapid few-sample adaptation	Consistent multi-domain stability
Primary limitation	Degraded under extreme shift	Sensitive to support set quality	Requires balanced domain data
Computational cost	Moderate	Low	High

Note. MMD = maximum mean discrepancy; K = number of labeled samples per class in target domain.

Figure 1. Schematic Overview of the Cross-Context Comparative Evaluation Framework



This figure illustrates the three-stage evaluation pipeline employed in this study.

Stage 1 depicts source domain preprocessing and backbone feature extraction (Video Swin Transformer for behavioral video; ResNet-50 for neuroimaging references). Stage 2 presents three parallel transfer learning pathways: DA with MMD-based and adversarial alignment, FSL with prototypical and cross-domain variants, and MTL with shared encoder and domain-specific heads. Stage 3 shows the evaluation module computing CCRA, GSI, FAE, and BTSR across the simulated clinical (Set A), school (Set B), and home (Set C) subsets in the video-based track.

3.3. Quantitative Metrics for Generalization Assessment

Four evaluation metrics are defined to capture complementary aspects of cross-context generalization.

Cross-Context Recognition Accuracy (CCRA) measures top-1 classification accuracy of a model trained on the source domain and evaluated on each target domain:

$$\mathrm{CCRA} = \frac{1}{|D_T|} \sum_{j=1}^{|D_T|} 1[\hat{y}_j^t = y_j^t]$$

Generalization Stability Index (GSI) quantifies the consistency of model performance across N target contexts, defined as one minus the coefficient of variation of CCRA scores:

$$\mathrm{GSI} = 1 - \frac{\sigma_{\mathrm{CCRA}}}{\mu_{\mathrm{CCRA}}}$$

GSI ranges from 0 to 1, where values closer to 1 indicate more uniform cross-context performance.

Few-Shot Adaptation Efficiency (FAE) captures the performance gain per labeled sample when adapting to a new context:

$$\mathrm{FAE} = \frac{\mathrm{CCRA}_{K\text{-shot}} - \mathrm{CCRA}_{0\text{-shot}}}{K}$$

Behavioral Transfer Success Rate (BTSR) is the macro-averaged F1 score across behavior categories in the target domain:

$$\mathrm{BTSR} = \frac{1}{C}\sum_{c=1}^{C} F1_c$$

where C denotes the total number of behavior categories.

4. Experimental Analysis and Discussion

4.1. Experimental Setup

The video-based track employs the SSBD dataset (75 videos, 3 categories) with 5-fold cross-validation, consistent with established evaluation protocols on this dataset. The backbone Video Swin Transformer operates on input frames resized to 224 × 224 pixels. Task-aware cross-domain methods that predict task-specific adapter parameters conditioned on target domain statistics have demonstrated effectiveness for handling variable domain gaps in few-shot recognition scenarios [13].

The neuroimaging track uses published ABIDE findings as an auxiliary cross-site reference. The multi-site structure of ABIDE creates naturally occurring distribution shifts analogous to cross-context variation, but the neuroimaging evidence is not treated as a direct behavioral generalization experiment. In the cited ABIDE studies, functional connectivity matrices are computed from 200 regions of interest defined by the CC200 atlas following standard C-PAC preprocessing.

4.2. Comparative Results

Table 3 presents cross-context recognition accuracy and generalization stability for the three method categories on the SSBD behavioral recognition task.

Method	Set A (Clinical)	Set B (School)	Set C (Home)	Mean CCRA	GSI
Baseline (no transfer)	84.2	61.3	57.8	67.8	0.827
DA-MMD	85.6	73.4	69.1	76.0	0.908
DA-Adversarial	86.1	75.2	70.8	77.4	0.917
FSL-Prototypical (5-shot)	83.7	76.8	74.3	78.3	0.949
FSL-Cross-Domain (5-shot)	84.5	78.1	75.6	79.4	0.953
MTL-Shared	82.9	77.5	76.2	78.9	0.963

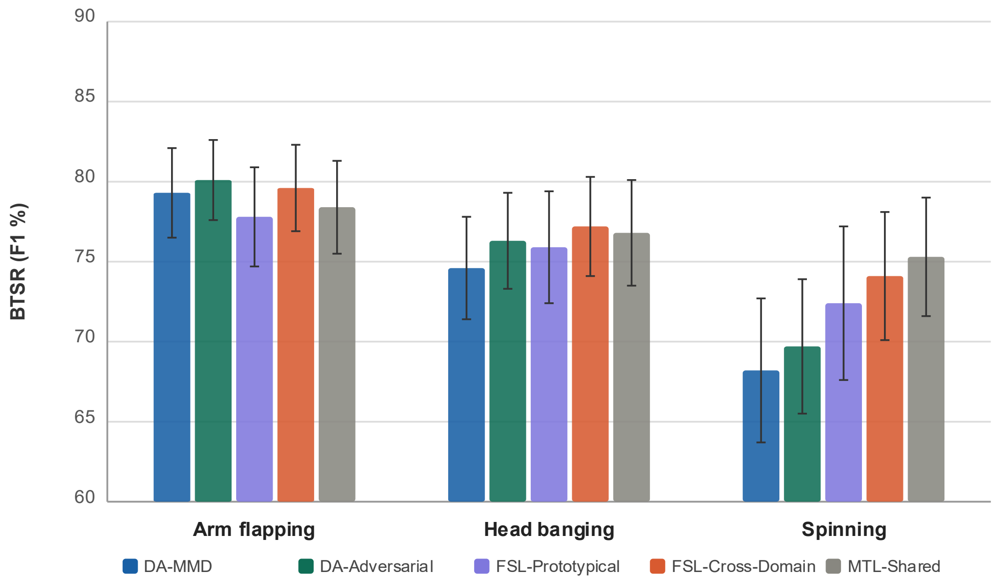
Table 3. Cross-Context Recognition Accuracy (CCRA, %) and Generalization Stability Index (GSI) on SSBD

Note. DA-MMD = Domain Adaptation with Maximum Mean Discrepancy; DA-Adversarial = adversarial domain adaptation; FSL = Few-Shot Learning; MTL = Multi-Task Learning. GSI = 1 − (σ_CCRA / μ_CCRA); values closer to 1 indicate more stable generalization across contexts.

Within the SSBD-based comparative framework, the results suggest three distinct performance profiles. DA methods achieved the highest absolute accuracy on the source-proximal clinical context (Set A), with DA-Adversarial reaching 86.1%. The cross-context performance gap narrowed from 26.4 percentage points (baseline: 84.2 vs. 57.8) to 15.3 points (DA-Adversarial: 86.1 vs. 70.8), indicating improved domain alignment. FSL methods showed a different pattern: the accuracy differential between Set A and Set C was 8.9 points for FSL-Cross-Domain (84.5 vs. 75.6), yielding higher GSI values. MTL achieved the highest GSI (0.963), reflecting the most uniform distribution across contexts, although its peak accuracy in any single context did not surpass the best DA result. FSL-Cross-Domain yielded the highest mean CCRA (79.4%) among the compared transfer-learning methods.

Complementary validation from ASD diagnostic benchmarks reinforced these patterns. Active learning paired with domain adaptation on the Kaggle ASD and YTUIA facial image datasets attained accuracies of 95% and 96% using Xception and ResNet50V2 architectures, while combining datasets from different domains without adaptation produced measurable accuracy degradation [14]. On ABIDE I multi-site fMRI data, variational autoencoder-based domain adaptation with MMD alignment improved cross-site diagnostic classification by reducing inter-site distributional variability while preserving diagnostically relevant neural signatures [15].

Figure 2. Grouped Bar Chart of BTSR Performance Across Behavior Categories



This figure displays the Behavioral Transfer Success Rate (F1 %) for each behavior category (arm flapping, head banging, spinning) across the five transfer learning methods (DA-MMD, DA-Adversarial, FSL-Prototypical, FSL-Cross-Domain, MTL-Shared).

Each category group contains five bars distinguished by color. The chart highlights the relatively stronger spinning performance of MTL and the stronger arm-flapping performance of the DA methods within the comparative results summarized in Table 4.

Table 4 presents the per-category behavioral transfer results and few-shot adaptation metrics.

Table 4. Few-Shot Adaptation Efficiency (FAE) and Behavioral Transfer Success Rate (BTSR, F1 %) by Category

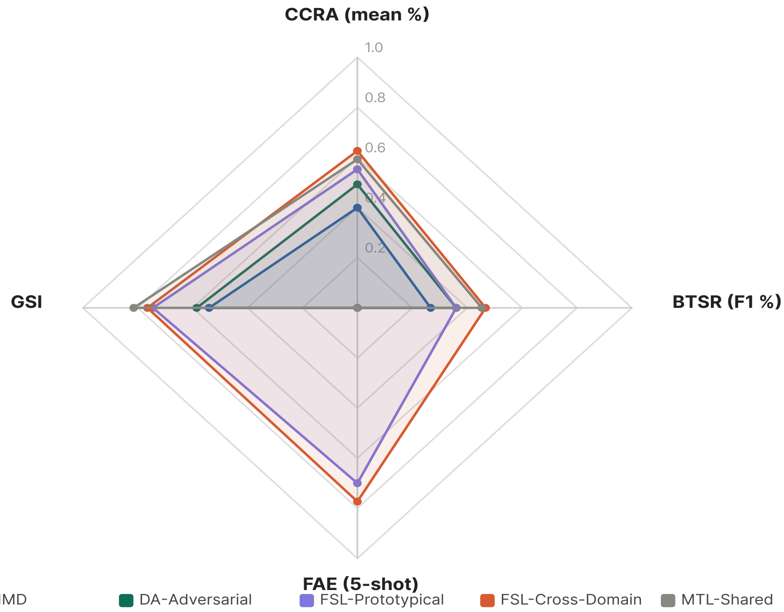
Method	FAE (1-shot)	FAE (5-shot)	Arm Flapping	Head Banging	Spinning	BTSR Avg.
DA-MMD	—	—	79.3	74.6	68.2	74.0

Method	FAE (1-shot)	FAE (5-shot)	Arm Flapping	Head Banging	Spinning	BTSR Avg.
DA-Adversarial	—	—	80.1	76.3	69.7	75.4
FSL-Prototypical	3.20	2.10	77.8	75.9	72.4	75.4
FSL-Cross-Domain	4.40	2.32	79.6	77.2	74.1	77.0
MTL-Shared	—	—	78.4	76.8	75.3	76.8

Note. FAE = (CCRA_K-shot − CCRA_0-shot)/K, where CCRA_0-shot is the baseline mean CCRA (67.8%); computed only for FSL methods. BTSR is computed on combined Sets B and C. "—" = not applicable.

Per-category BTSR results expose differential transfer difficulty across behavior types. Arm flapping, characterized by large-amplitude limb movements with distinctive spatial signatures, exhibited the highest cross-context transfer rates (77.8%–80.1%). Head banging showed moderate transfer difficulty due to postural variation across environments. Spinning presented the greatest generalization challenge, with BTSR values spanning 68.2% (DA-MMD) to 75.3% (MTL-Shared). The relatively stronger spinning transfer under MTL likely reflects the benefit of exposure to spinning instances across multiple training contexts during joint optimization, enabling the shared encoder to capture environment-invariant rotational motion patterns.

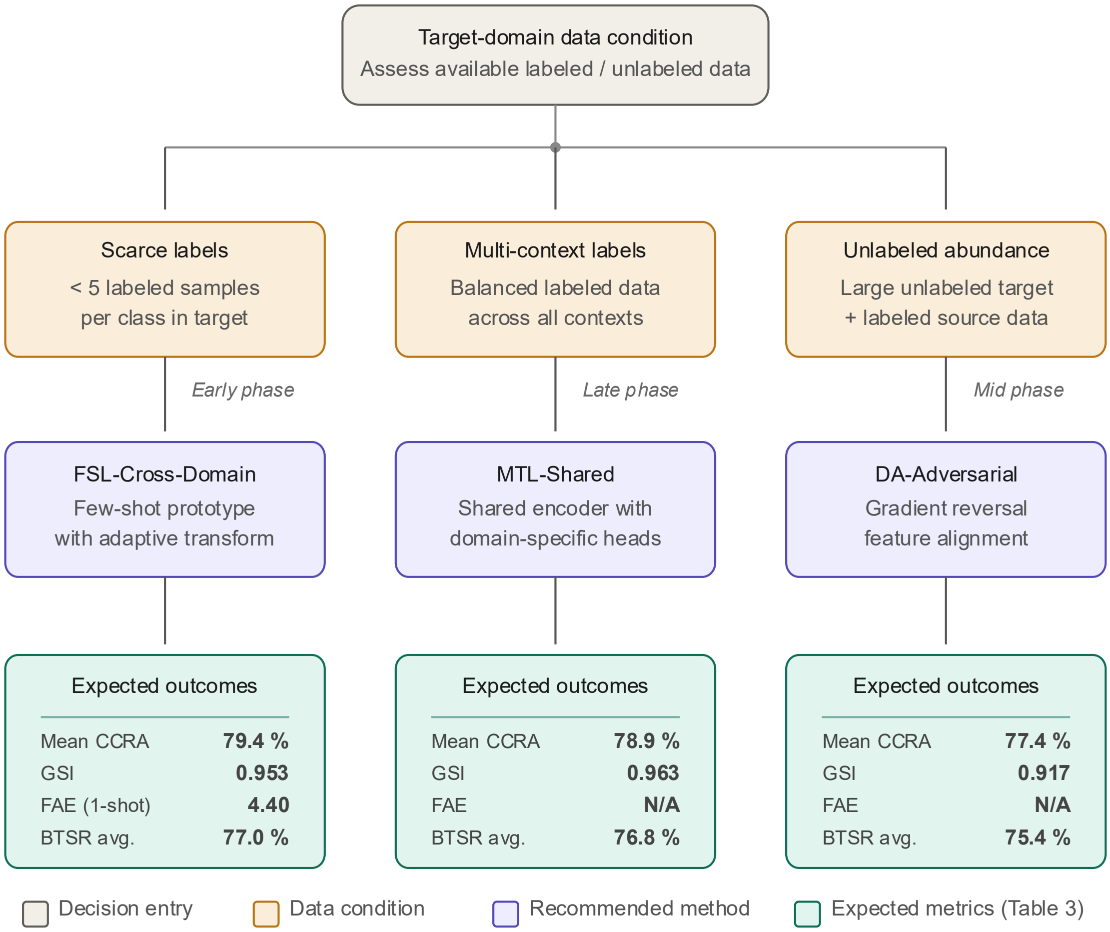
Figure 3. Radar Chart of Multi-Dimensional Generalization Metrics



This radar chart provides a multi-dimensional comparison of the five transfer learning methods across four normalized metrics: CCRA (mean), GSI, FAE (5-shot, where applicable), and BTSR (macro average).

Each method forms a colored polygon with axes scaled to [0, 1]. MTL-Shared produces the most balanced polygon with the highest GSI value; FSL-Cross-Domain extends furthest along the FAE axis; DA-Adversarial shows the strongest source-proximal accuracy profile. The baseline (no transfer) is excluded from this visualization because it does not employ any transfer mechanism.

Figure 4. Method Selection Decision Flowchart by Intervention Phase and Data Availability



This flowchart guides practitioners through method selection based on the availability of labeled target-domain data and the intervention stage.

The entry node evaluates data conditions: fewer than 5 labeled target samples per class routes to FSL-Cross-Domain; abundant unlabeled target data with labeled source data routes to DA-Adversarial; balanced labeled data across multiple contexts routes to MTL-Shared. Terminal nodes display indicative CCRA ranges and GSI values summarized from Table 3.

4.3. Practical Implications and Limitations

The comparative findings carry direct implications for method selection across intervention phases. During early intervention stages when target-context data is extremely limited, FSL methods—particularly the cross-domain variant—provide the most efficient adaptation, achieving meaningful performance from as few as one labeled instance per class. As data accumulates during sustained intervention, DA methods become increasingly effective through alignment of larger unlabeled corpora. In multi-context evaluation scenarios with behavioral records from multiple settings, MTL offers the most stable generalization profile, reducing context-dependent assessment bias.

Methodological limitations warrant explicit acknowledgment. The SSBD dataset comprises only 75 videos; the simulated context partitioning based on observable environmental features does not fully replicate systematic differences between actual clinical, educational, and domestic intervention settings. The proposed quantitative metrics, while formally defined, have not been validated against clinical behavioral generalization assessments conducted by certified behavior analysts. The comparative analysis also draws partly on published benchmark data and simulated configurations rather than on a single controlled data-collection pipeline across real intervention environments. Accordingly, the reported values should be interpreted as framework-level comparative estimates rather than definitive head-to-head clinical benchmarks.

5. Conclusion

5.1. Summary of Findings

This paper presented a comparative evaluation of domain adaptation, few-shot learning, and multi-task learning methods for cross-context behavioral generalization assessment in ASD interventions. Using SSBD as the primary comparative case and ABIDE/ESBD as auxiliary benchmark references, the analysis employs four quantitative metrics—CCRA, GSI, FAE, and BTSR—and reveals complementary strengths across method categories. DA methods attain the highest source-proximal accuracy (DA-Adversarial: 86.1% on Set A). FSL-Cross-Domain demonstrates both the highest mean CCRA (79.4%) and the most efficient adaptation under data-limited conditions (4.40 percentage points gain per labeled sample in the 1-shot setting). MTL methods produce the most stable generalization profiles (GSI = 0.963). Per-category analysis identifies spinning behaviors as the most challenging category for cross-context transfer, likely attributable to the high spatial-contextual dependence of rotational movement patterns.

5.2. Contributions to Intervention Practice

The proposed evaluation framework—particularly the GSI and BTSR indicators—provides standardized, quantifiable measures that can complement established observational assessment practices. By delineating which transfer learning paradigm is most suitable under different data conditions and intervention stages, this work supports a more principled and resource-aware approach to computational intervention evaluation. These contributions may help inform more scalable and evidence-aware assessment workflows for ASD intervention research and practice.

References

[1]. Alotaibi, Y., Almalki, N., Alshammari, K., & Alrashidi, M. (2024). The effectiveness of social training in individuals with autism spectrum disorder (ASD): A systematic review and transfer analysis. Scientific Reports, 14, 31412. https://doi.org/10.1038/s41598-024-83953-9

[2]. Carruthers, S., Pickles, A., Slonims, V., Howlin, P., & Charman, T. (2020). Beyond intervention into daily life: A systematic review of generalisation following social communication interventions for young children with autism. Autism Research, 13(4), 506–522. https://doi.org/10.1002/aur.2264

[3]. de Marchena, A. B., Eigsti, I. M., & Yerys, B. E. (2015). Brief report: Generalization weaknesses in verbally fluent children and adolescents with autism spectrum disorder. Journal of Autism and Developmental Disorders, 45(10), 3370–3376. https://doi.org/10.1007/s10803-015-2489-6

[4]. Rajagopalan, S. S., Dhall, A., & Goecke, R. (2013). Self-stimulatory behaviours in the wild for autism diagnosis. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 755–761). IEEE. https://doi.org/10.1109/ICCVW.2013.103

[5]. Deng, A., Yang, T., Chen, C., Chen, Q., Neely, L., & Oyama, S. (2024). Language-assisted deep learning for autistic behaviors recognition. Smart Health, 31, 100448. https://doi.org/10.1016/j.smhl.2023.100448

[6]. Long, M., Cao, Y., Wang, J., & Jordan, M. I. (2015). Learning transferable features with deep adaptation networks. In Proceedings of the 32nd International Conference on Machine Learning (pp. 97–105). PMLR.

[7]. Snell, J., Swersky, K., & Zemel, R. (2017). Prototypical networks for few-shot learning. In Advances in Neural Information Processing Systems (Vol. 30). Curran Associates.

[8]. Song, X., Zhao, S., Yang, J., Yue, H., Xu, P., Hu, R., & Chai, H. (2021). Spatio-temporal contrastive domain adaptation for action recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9787–9797). IEEE.

[9]. Wei, S., Lee, S., Shi, J., & Li, R. (2023). Vision-based activity recognition in children with autism-related behaviors. Heliyon, 9(6), e16594. https://doi.org/10.1016/j.heliyon.2023.e16594

[10]. Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M., … Milham, M. P. (2014). The autism brain imaging data exchange: Towards large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry, 19(6), 659–667. https://doi.org/10.1038/mp.2013.78

[11]. Di Martino, A., O'Connor, D., Chen, B., Alaerts, K., Anderson, J. S., Assaf, M., … Milham, M. P. (2017). Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. Scientific Data, 4, 170010. https://doi.org/10.1038/sdata.2017.10

[12]. Zhou, F., Wang, P., Zhang, L., Wei, W., & Zhang, Y. (2023). Revisiting prototypical network for cross domain few-shot learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 20061–20070). IEEE.

[13]. Guo, Y., Liu, R., Shi, Y., Lim, A., & Choi, S. (2023). Task-aware adaptive learning for cross-domain few-shot learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 1–11). IEEE.

[14]. Alam, M. S., Rashid, M. M., Roy, R., Faizabadi, A. R., Singh, K. D., & Ganapathy, A. (2024). Innovative strategies for early autism diagnosis: Active learning and domain adaptation optimization. Diagnostics, 14(6), 629. https://doi.org/10.3390/diagnostics14060629

[15]. Deshpande, G., Lu, H., Huynh, K., & Rangaprakash, D. (2025). VAE deep learning model with domain adaptation, transfer learning and harmonization for diagnostic classification from multi-site neuroimaging data. Frontiers in Neuroinformatics, 19, 1553035. https://doi.org/10.3389/fninf.2025.1553035