COMPARATIVE STUDY OF DATASET FOR CLASSROOM ATTENTION ANALYSIS

Nimitha Mohan[1], V Vadivu[2]

Abstract

The quality and appropriateness of the datasets used for training and assessment have a significant impact on how well artificial intelligence and deep learning models perform in classroom attention analysis. With an emphasis on kids and educational environments, this study offers a comparative examination of publicly accessible and research-focused datasets used for classroom attention and behavior analysis. The review examines datasets according to their application to attention state recognition, subject demographics, annotation techniques, data modalities, and ethical considerations. This study attempts to assist academics in choosing suitable datasets without disclosing secret algorithms by methodically analyzing datasets published in recent publications. Additionally, the review draws attention to current shortcomings and difficulties, encouraging the creation of new datasets for reliable classroom video analysis.

Keywords

Classroom attention analysis, video datasets, deep learning, student engagement, educational data mining

I. INTRODUCTION

One of the main challenges in educational research is comprehending students' attention in classroom settings. Automated attention state recognition using classroom video data has been made possible by recent developments in deep learning and artificial intelligence.

However, the availability of ethically produced and well-annotated datasets is crucial to the effectiveness of such systems. This work avoids algorithmic disclosure by concentrating only on dataset features documented in the literature and offers a structured comparison specifically designed for children's attention analysis. In recent years, increasing class sizes, diverse learning needs, and the rapid integration of digital technologies have made it challenging for educators to continuously monitor student attention using traditional observational methods. Attention plays a pivotal role in learning outcomes, influencing comprehension, retention, and overall academic performance. Consequently, there is growing interest in automated systems capable of analysing student attention objectively and continuously within classroom environments.

Artificial intelligence-driven classroom attention analysis has emerged as a promising solution, enabling large-scale observation through visual data such as classroom videos and webcam streams. These systems rely heavily on supervised learning approaches, where model reliability and generalizability are fundamentally constrained by the quality of training datasets. Factors such as demographic diversity, classroom realism, annotation granularity, and ethical compliance directly impact the effectiveness of attention recognition systems.

Despite notable progress in attention and engagement modelling, existing studies often overlook the dataset dimension, treating datasets as secondary to model architecture and performance metrics. This oversight has resulted in fragmented research outcomes, limited reproducibility, and difficulty in benchmarking across studies. In child-centric educational contexts, these challenges are further intensified by privacy concerns, consent requirements, and ethical restrictions on data sharing.

Therefore, a systematic examination of datasets used for classroom attention analysis is essential.

Rather than proposing new algorithms, this study emphasizes dataset characteristics as the foundation for reliable and ethical attention modelling. By critically comparing datasets reported in recent literature, this work provides a structured reference that supports informed dataset selection, highlights limitations affecting generalization, and identifies directions for future dataset development tailored to real-world classroom environments.

II. RELATED WORK

Autism Spectrum Disorder (ASD) is a neurological condition that requires early detection to reduce its impact. Significant advancements in this domain have been driven by

Department of Computer Applications[1],
Sree Narayana Gurukulam College of Engineering, Kadayiruppu.[1]
nimithamohan18@gmail.com.[1]

Department of Computer Science[2],
Karpagam Academy of Higher Education Coimbatore-641 021.[2]
vadivu.vijayan@kahedu.edu.in[2]

* Corresponding Author

progress in machine learning techniques. Researchers have explored a wide range of approaches for early identification of ASD, utilizing diverse data sources such as MRI images, EEG signals, video data, and facial expressions. Various deep learning models, including Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and Recurrent Neural Networks (RNN), have been widely applied in these studies.

In work Studies [1] on classroom attention assessment have experienced substantial growth due to the progress in artificial intelligence and deep learning methods. A significant amount of current research centers on identifying student attention, engagement, and behavior through visual data collected from classroom settings. Nonetheless, instead of focusing on algorithmic advancements, multiple studies underscore the essential importance of datasets in influencing the performance and applicability of attention recognition systems.

Initial efforts in classroom behavior analysis mainly depended on specially gathered video datasets recorded in real classroom environments. These datasets usually consist of school-aged children and are marked with broad labels such as attention and inattention. Though these datasets provide authentic classroom environments, their small size and limited access resulting from privacy issues diminish their reusability and comparability across studies. Nonetheless, they establish the fundamental groundwork for attention modeling in practical educational settings.

In research focused on special education,[2] especially studies related to children with autism spectrum disorder (ASD), datasets are frequently gathered in structured learning environments. These datasets include video recordings enhanced with gaze direction, head position, or response-based annotations to obtain detailed attention signals. While particularly important for comprehending attention mechanisms, these datasets typically face issues with limited sample sizes and restricted variability, which hinders their applicability to mainstream classroom environments.

As digital education platforms have expanded,[3] numerous studies have examined e-class and online interaction datasets, analyzing webcam video streams to determine levels of student engagement. These datasets primarily cater to university students and emphasize engagement rather than specific attention states. Although they provide valuable methodological insights, their relevance to analyzing child-focused classroom attention is restricted because of variations in learning behaviors, contextual environments, and developmental influences.

Other similar studies explore datasets on student activity monitoring gathered in smart classrooms or surveillance-type settings.[4] These datasets typically focus on activity categorization or action recognition instead of attention modeling.

Even though they offer extensive data and varied visuals. This comparative perspective complements existing attention recognition studies[5] by emphasizing dataset suitability and limitations, thereby supporting reproducibility and ethical research practices in classroom video analysis.

III. PROPOSED WORK

This research presents a methodical comparative structure for examining datasets utilized in studies of classroom attention and engagement, particularly focusing on [6] child-centered learning environments. In contrast to current studies that mainly emphasize model architecture or performance indicators, the proposed approach prioritizes systematic dataset assessment as the basis for dependable attention state recognition.

The suggested framework starts with the identification and gathering of representative datasets found in recent literature [7] concerning classroom attention analysis, student engagement assessment, and behavior recognition. Both traditional classrooms and online learning platforms are regarded as essential for achieving thorough representation of actual educational situations. Datasets are chosen according to their relevance to visual attention indicators such as gaze direction, head position, facial expressions, and body posture, while avoiding any disclosure of proprietary data sources.

Every dataset [8] is subsequently assessed using a consistent evaluation standard that encompasses environmental context, demographic information of participants, data modality, annotation approach, accessibility, and adherence to ethical standards. This uniform assessment allows for significant comparison among datasets that were initially created for various purposes. Attention is paid to datasets that include children, where maintaining privacy, obtaining consent, and using non-intrusive data collection techniques are essential.

The suggested research [9] further classifies datasets according to their appropriateness for recognizing attention states, differentiating between datasets focused on attention, those oriented towards engagement, and activity-related datasets that have indirect connections to attention evaluation. This categorization aids researchers in grasping the advantages and drawbacks of current resources, allowing them to choose suitable datasets for their specific research objectives.

Moreover, the framework emphasizes annotation detail and temporal consistency, which are vital elements for representing attention dynamics in classroom videos. [10]

The proposed work highlights gaps [11] that could influence model generalization and robustness by examining annotation levels across different datasets, especially in extended classroom recordings. Ultimately, the suggested comparative analysis seeks to uncover research deficiencies and dataset constraints that impede scalable and replicable attention evaluation in education.

IV. DATASET REQUIREMENTS FOR CLASSROOM ATTENTION ANALYSIS

Datasets for classroom attention analysis must capture realistic learning scenarios, [12] a variety of student behaviors, and various visual indicators including gaze direction, head stance, facial expression, and body posture must all be captured in datasets for classroom attention analysis. High-resolution video, temporal continuity, accurate attention state annotation, and adherence to ethical and privacy standards, especially when involving children- are important prerequisites.

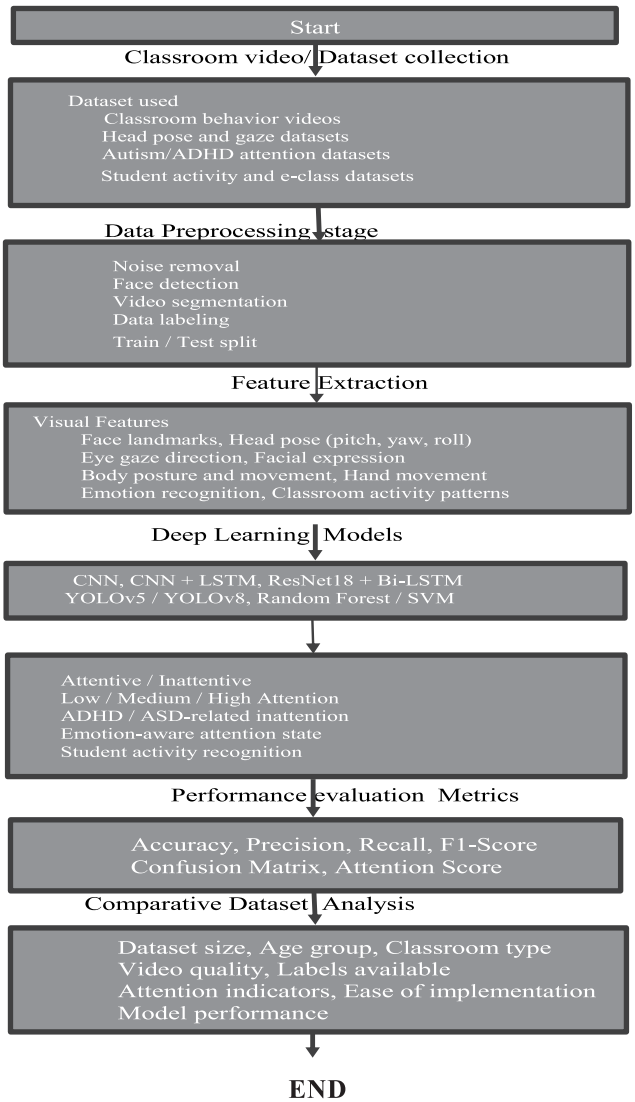


Figure1: Framework for Classroom Attention Analysis

V. OVERVIEW OF REPRESENTATIVE DATASETS

The datasets used in recent research on e-class engagement detection, [13] autism-related learning analysis, attention estimation, and classroom behavior monitoring are compiled in this area. A few datasets are partially available for academic use, but the majority are custom-collected because of privacy concerns.

VI. COMPARITIVE ANALYSIS OF DATA

Table 1: Comparative study of Data

Dataset Source	Environment	Participants	Data Modality	Annotations	Limitations
Classroom Behavior Dataset	Physical Classroom	School Children	RGB Video	Attention / Inattention Labels	Limited public access
Autism Attention Dataset	Controlled Learning Setup	Children with ASD	Video + Gaze	Attention Levels	Small sample size
E-Class Engagement Dataset	Online Classroom	University Students	Webcam Video	Engagement States	Not child-focused
Student Activity Monitoring Dataset	Smart Classroom	Mixed Age Group	Video + Facial Cues	Activity Categories	Activity-focused, not attention-specific

Figure illustrates the distribution of datasets used in classroom attention analysis across different application domains. The chart indicates that a majority of the datasets focus on general classroom attention analysis, reflecting the primary research emphasis on understanding student attentiveness in traditional learning environments. [18] Datasets related to engagement analysis in e-class or online learning settings and activity monitoring in smart classrooms appear in moderate proportions, demonstrating growing interest in technology-assisted educational environments.

In contrast, [20] datasets specifically developed for autism and special education contexts remain limited. This scarcity can be attributed to ethical constraints, privacy considerations, and challenges in large-scale data collection involving children, highlighting a significant research gap in child-centric attention analysis.[14]

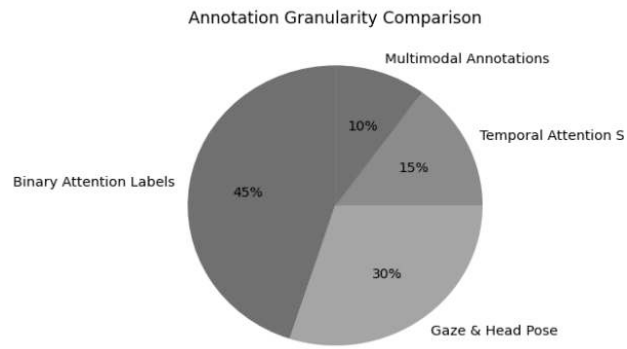


Figure 2: Dataset Annotation Used in Attention Analysis

VII. RESULT AND DISCUSSION

The comparative study reveals notable trends and constraints within current datasets utilized for analyzing classroom attention. A prevailing observation is that the majority of datasets are specific to tasks, designed to assist in narrowly defined research goals like activity recognition, engagement estimation, or behavior monitoring, rather than comprehensive attention modeling. Consequently, these datasets frequently lack the variety and contextual depth necessary for creating generalized attention recognition systems suitable for various classroom settings.

A significant discovery [15] is that most datasets are gathered in controlled or semi-controlled environments. Although these environments promote stable data collection and dependable labeling, they fail to capture the full complexity and variability found in actual classrooms. Elements such as different lighting situations, obstructions, varied seating configurations, and unpredictable student actions are frequently overlooked. This constraint hinders the strength and practical use of models developed with these datasets.

In the work [16] Datasets related to children, especially within educational or special education settings, prioritize ethical clearance, safeguarding privacy, and using non-intrusive data collection techniques. While these practices are crucial, they often result in limited access to datasets, reduced sample sizes, or total unavailability for public access. As a result, reproducibility and comparative analysis among studies become difficult, hindering advancement in this research field.

The analysis [20] also shows significant differences in annotation approaches and detail levels. Certain datasets depend on broad binary labels such as attention and inattention, whereas others offer detailed annotations including gaze orientation, head position, facial expressions, or temporal attention metrics. This inconsistency impacts model training and assessment, as variances in annotation depth and definitions of attention impede cross-dataset generalization and evaluation.

There is a considerable absence of extensive, publicly accessible datasets explicitly created for analyzing classroom attention in children. Current extensive datasets usually concentrate on adult learners, online learning, or general activity detection instead of attention in child-centered classrooms.

Overall, [17] the findings indicate that while current datasets provide valuable insights and serve as important steppingstones, they fall short of supporting comprehensive and generalized classroom attention analysis. Addressing these limitations is crucial for advancing AI-driven educational technologies and enabling reliable deployment in real-world learning environments.

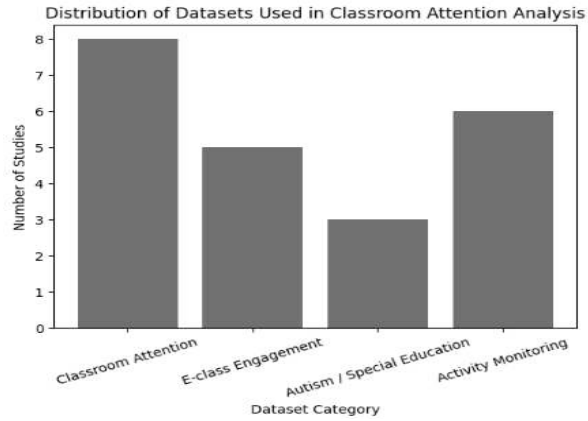


Figure 3: Dataset Categories Used for Attention Analysis

VIII. RESEARCH GAP AND FUTURE DIRECTION

Despite growing interest in AI-driven classroom attention analysis, several critical research gaps remain primarily related to dataset availability, design, and standardization. One of the most significant gaps is the absence of standardized, large-scale datasets specifically tailored for classroom attention analysis in children. Existing datasets are often custom collected, limited in size, and designed for isolated tasks, making it difficult to compare results across studies or validate model robustness.

Future research should prioritize the development of ethically compliant datasets that adhere to strict privacy, consent, and data protection guidelines, particularly when involving children. Transparent ethical protocols, anonymization techniques, and controlled data-sharing frameworks can help balance participant safety with research accessibility. Establishing clear ethical standards for dataset creation would also encourage wider collaboration and data reuse within the research community.

Another major gap lies in the lack of diversity in classroom settings and participant demographics. Many current datasets represent single classrooms, specific age groups, or controlled learning environments. Future datasets should encompass varied educational contexts, including different school types, teaching styles, class sizes, cultural backgrounds, and age ranges. Such diversity is essential for developing attention models that generalize effectively across real- world educational scenarios.

The limited use of multimodal data further constrains current attention modeling approaches. While most datasets rely primarily on RGB video, attention is inherently multimodal, influenced by audio cues (e.g., teacher speech,

student responses), temporal context, and classroom interactions. Future dataset initiatives should integrate video, audio, depth, and contextual metadata such as lesson type or seating layout to enable more comprehensive and accurate attention state recognition.

Additionally, there is a clear need for benchmark datasets with standardized annotation schemes and evaluation protocols. Inconsistent definitions of attention and varying annotation granularity hinder fair comparison of methods across studies. Establishing benchmark datasets with well-defined attention categories, temporal annotations, and publicly available evaluation metrics would significantly advance reproducibility and comparative evaluation in this domain.

Overall, addressing these research gaps through coordinated dataset development efforts will be crucial for advancing reliable, scalable, and ethically responsible classroom attention analysis systems. Such progress would not only enhance technical performance but also support meaningful deployment of AI technologies in educational environments

IX. FUTURE WORK

Future work will focus on the design and development of a child-centric classroom attention dataset that captures realistic learning environments while maintaining strict ethical and privacy compliance. Data collection will aim to include multiple classroom settings, age groups, and teaching styles to improve diversity and ecological validity. Special attention will be given to non-intrusive recording methods and anonymization techniques to ensure participant's safety.

In addition, future efforts will explore the incorporation of multimodal information, including synchronized video and audio streams, to better capture subtle indicators of attention such as verbal responses, classroom interactions, and teacher–student dynamics. Temporal annotation of attention states will be emphasized to support continuous attention modeling rather than static frame-based analysis.

Another direction involves the standardization of attention definitions and annotation protocols. Developing clear labeling guidelines and inter-annotator agreement measures will improve annotation reliability and enable meaningful comparison across studies. The creation of a benchmark subset with predefined evaluation metrics is also planned to facilitate reproducibility and comparative performance assessment. Learning environments, thereby supporting broader adoption of AI-driven attention monitoring systems in education.

Finally, future work will investigate the scalability and adaptability of attention analysis frameworks across different educational contexts. This includes evaluating dataset applicability to inclusive classrooms, special education settings, and technology-assisted learning environments, thereby supporting broader adoption of AI-driven attention monitoring systems in education.

Finally, future work will investigate the scalability and adaptability of attention analysis frameworks across different educational contexts. This includes evaluating dataset applicability to inclusive classrooms, special education settings, and technology-assisted learning environments, thereby supporting broader adoption of AI-driven attention monitoring systems in education.

X. CONCLUSION

This paper presented a comprehensive comparative review of datasets used for classroom attention analysis, with a particular focus on child- centric educational environments. By systematically examining dataset characteristics such as data modality, annotation strategies, participant demographics, ethical considerations, and accessibility, the study highlighted both the strengths and limitations of existing datasets provide valuable foundations for attention modeling, they are often constrained by limited scale, restricted availability, and narrow task-specific design reported in the literature.

The findings further reveal a critical gap in the availability of large- scale, publicly accessible datasets specifically designed for classroom attention analysis in children. Ethical and privacy requirements, although essential, have contributed to fragmented and inaccessible data resources, thereby limiting comparative evaluation and generalization of attention recognition models. Addressing these challenges is essential for enabling robust, scalable, and real-world deployable AI systems in educational contexts.

Overall, this review underscores the urgent need for the development of open, ethically compliant, and standardized datasets that reflect the diversity and complexity of real classroom environments. Such datasets will play a pivotal role in advancing AI-based attention state recognition, fostering fair benchmarking, and ultimately supporting more effective and intelligent educational technologies.

XI. ENHANCED LITERATURE TO DATASET MAPPING

Table 2 : Literature to Dataset Mapping

Reference Study	Application Domain	Target Age Group	Dataset Origin	Annotation Type	Dataset Size	Contribution to Attention Analysis
Attention Assessment through Classroom Behavior Detection	Physical Classroom	School Children	Custom-Collected Classroom Videos	Attention vs. Inattention Labels	Moderate	Direct support for classroom attention modeling
Attention Level Evaluation in Children with Autism	Special Education	Children with ASD	Controlled Video Dataset	Head Pose and Gaze Indicators	Limited	High relevance for fine-grained attention cues
Response-to-Name Detection for Autism Screening	Early Behavioral Screening	Young Children	Structured Interaction Videos	Visual and Auditory Response Events	Limited	Indirect relevance through attention response behavior
Real-Time Student Activity Detection and Monitoring	Smart Classroom	Mixed Age Group	Surveillance-Style Video Data	Activity-Based Labels	Large	Moderate relevance, activity-oriented focus
Student Recognition and Activity Monitoring in E-Classes	Online Learning	Higher Education	Webcam Video Streams	Engagement Indicators	Moderate	Limited relevance to child-centric attention analysis

REFERENCES

[1] D. S. Hasin, C. P. O'brien, M. Auriacombe, G. Borges, K. Bucholz, A. Budney, W. M. Compton, T. Crowley, W. Ling, N. M. Petry,et al., "Dsm-5 criteria for substance use disorders: recommendations and rationale," American Journal of Psychiatry, vol. 170, no. 8, pp. 834–851,2013.

[2] C. Lord, M. Elsabbagh, G. Baird, and J. Veenstra-Vanderweele, "Autism spectrum disorder," The lancet, vol. 392, no. 10146, pp. 508–520, 2018.

[3] R. M. Joseph and H. Tager-Flusberg, "An investigation of attention and affect in children with autism and down syndrome," Journal of autism and developmental disorders, vol. 27, no. 4, pp. 385–396, 1997.

[4] A. Knopf, "Autism prevalence increases from 1 in 60 to 1 in 54: Cdc," The Brown University Child and Adolescent Behavior Letter, vol. 36, no. 6, pp. 4-4, 2020.

[5] M. Juneja and S. Sairam, "Autism spectrum disorder—an Indian perspective," Recent Advances in Autism, 2018.

[6] B. Geetha, C. Sukumar, E. Dhivya Deepa, J. K. Reddy, and V. Bal-Chandar, "Autism in India: a case–control study to understand the association between socio-economic and environmental risk factors," Acta Neurological Belgica, vol. 119, no. 3, pp. 393–401, 2019.

[7] A. Zunino, P. Morariu, A. Cavallo, C. Ansuini, J. Podda, F. Battaglia, E. Veneselli, C. Becchio, and V. Murino, "Video gesture analysis for autism spectrum disorder detection," in 2018 24th international conference on pattern recognition (ICPR), pp. 3421–3426, IEEE, 2018.

[8] G. Wan, X. Kong, B. Sun, S. Yu, Y. Tu, J. Park, C. Lang, M. Koh, Z. Wei,Z. Feng, et al., "Applying eye tracking to identify autism spectrum disorder in children," Journal of autism and developmental disorders, vol. 49, no. 1, pp. 209–215, 2019.

[9] Z. Sherkatghanad, M. Akhondzadeh, S. Salari, M. Zomorodi-Moghadam, M. Abdar, U. R. Acharya, R. Khosrowabadi, and V. Salari, "Automated detection of autism spectrum disorder using a convolutional neural

network," Frontiers in neuroscience, vol. 13, p. 1325, 2020.

[10] J. Kang, X. Han, J. Song, Z. Niu, and X. Li, "The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data," Computers in biology and medicine, vol. 120, p. 103722, 2020.

[11] M. Leo, P. Carcagn`ı, C. Distante, P. Spagnolo, P. L. Mazzeo, A. C. Rosato, S. Petrocchi, C. Pellegrino, A. Levante, F. De Lum`e, et al., "Computational assessment of facial expression production in asd children," Sensors, vol. 18, no. 11, p. 3993, 2018.

[12] S. Raj and S. Masood, "Analysis and detection of autism spectrum dis-order using machine learning techniques," Procedia Computer Science, vol. 167, pp. 994–1004, 2020.

[13] M. F. Misman, A. A. Samah, F. A. Ezudin, H. A. Majid, Z. A. Shah, H. Hashim, and M. F. Harun, "Classification of adults with autism spectrum disorder using deep neural network," in 2019 1st International Conference on Artificial Intelligence and Data Sciences (AiDAS), pp. 29–34, IEEE, 2019.

[14] K. S. Omar, P. Mondal, N. S. Khan, M. R. K. Rizvi, and M. N. Islam, "A machine learning approach to predict autism spectrum disorder," in 2019 International conference on electrical, computer and communication engineering (ECCE), pp. 1–6, IEEE, 2019.

[15] J. Kosmicki, V. Sochat, M. Duda, and D. Wall, "Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning," Translational psychiatry, vol. 5, no. 2, pp. e514–e514, 2015.

[16] A. Gulli and S. Pal, Deep learning with Kera's. Packt Publishing Ltd, 2017.

[17] A. S. Heins Feld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F Meneguzzi, "Identification of autism spectrum disorder using deep learning and the abide dataset," Neuroimage: Clinical, vol. 17, pp. 16–23, 2018.

[18] J. Liu, K. He, Z. Wang, and H. Liu, "A computer vision system to assist the early screening of autism spectrum disorder," in International Conference on Cognitive Systems and Signal Processing, pp. 27–38, Springer, 2018.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.

[20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions, "in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9, 2015.