Video-Based Autism Spectrum Disorder Detection with Integrated Chatbot Assistance

Mahima S
Information Technology
College Of Engineering Guindy
Anna University ,Chennai
mahimaselvam5868@gmail.com

Kavitha E
Information Technology
College Of Engineering Guindy
Anna University ,Chennai
kavithaelumalai2005@gmail.com

Dhanalakshmi G
Information Technology
College Of Engineering Guindy
Anna University ,Chennai
dhanalakshmig03092004@gmail.com

***Abstract*—Autism Spectrum Disorder (ASD) is a developmental condition that affects communication, behavior, and social interaction. Early diagnosis is critical for effective intervention, but traditional methods are often time-consuming, costly, and not readily accessible. This project introduces a video-based analysis system that uses deep learning to detect signs of autism by analyzing behavioral patterns such as facial expressions, eye contact, and gestures. The system allows users to upload videos, which are processed by a trained model to provide an initial assessment of ASD presence.**

In addition to detection, the project features an AI-powered chatbot designed to assist parents and guardians by offering real-time guidance, therapy suggestions, and answers to common questions related to autism. This combination of automated detection and chatbot support creates a user-friendly, accessible platform that encourages early screening and offers emotional and informational assistance to families. The system aims to support timely intervention and improve access to resources for those affected by ASD.

***Index Terms*—Autism Spectrum Disorder (ASD), Video-Based Analysis, Deep Learning, Early Detection, Behavioral Analysis, Autism Detection, AI in Healthcare, Chatbot Assistance, Parent Support System, Facial Expression Analysis, Neurodevelopmental Disorder, Real-time Support**

I. Introduction

A. Motivation

Autism Spectrum Disorder (ASD) affects millions of children worldwide, with symptoms often going undiagnosed until later stages due to the complexity and variability in behavior. Early detection plays a critical role in improving long-term outcomes, yet access to professional diagnosis is limited in many parts of the world due to cost, availability, or social stigma. This gap in accessibility motivated the development of a solution that could provide an initial screening method that is both non-invasive and easily accessible. By leveraging video-based analysis, we aim to create a system that can observe and interpret key behavioral indicators of ASD without the need for in-person clinical assessments.

In addition to early detection, the emotional and informational support needed by parents and caregivers during this journey is often overlooked. Many are left searching for reliable guidance and resources after suspecting signs of ASD in their children. To address this, we integrated a chatbot into the system to assist users with answers to common concerns, therapy suggestions, and emotional support. The motivation behind combining deep learning with chatbot technology lies in our goal to not only detect autism efficiently but also provide a digital companion that empowers families with knowledge and confidence in taking early action.

B. Problem Statement

Autism Spectrum Disorder (ASD) is often misunderstood and difficult to identify, especially in its early stages. Many parents notice unusual behaviors in their child—like avoiding eye contact, delayed speech, or repetitive actions—but aren't sure what they mean or how serious they are. Getting a professional diagnosis can be challenging, especially in areas where access to specialists is limited or where families are hesitant due to social stigma or financial constraints. Even when help is available, the process can be long, stressful, and emotionally draining, leaving families without the clarity or support they need during a critical time.

In today's digital world, there's an opportunity to make this process easier, faster, and more accessible. By using video-based technology, we can observe a child's behavior naturally, without the pressure of clinical settings. But detection alone isn't enough—families also need guidance, reassurance, and emotional support. That's why this project aims to not only detect possible signs of ASD through AI-powered video analysis but also offer real-time assistance through a chatbot. This digital companion helps bridge the gap between uncertainty and action, giving families the comfort of immediate support and helping them take informed next steps.

C. Objective

The main objective of this project is to develop an intelligent, accessible, and user-friendly system that helps in the early detection of Autism Spectrum Disorder (ASD) through video-based behavioral analysis. By leveraging computer vision and deep learning techniques, the system aims to identify patterns such as lack of eye contact, repetitive movements, and atypical facial expressions—common indicators associated with ASD. This approach allows for non-invasive and remote screening, making the detection process more accessible, especially in underserved areas.

In addition to detection, the project also focuses on supporting parents and guardians by integrating a chatbot that can provide instant responses, guidance, therapy suggestions, and emotional support. The chatbot acts as a virtual assistant, helping users understand the results, reduce anxiety, and make informed decisions about seeking professional help. Overall, the project aims to create a dual-function platform that not only facilitates early identification of autism but also empowers families with timely information and compassionate support.

II. Related Work

Over the years, various projects and research studies have attempted to address the early detection of Autism Spectrum Disorder (ASD) using modern technology.

A. Video Gesture Analysis for Autism Spectrum Disorder Detection

The research demonstrated that machine learning models could effectively distinguish between neurotypical and autistic behavior based on subtle variations in hand and body movements. Their work provided a scientific basis for non-invasive, video-based diagnostic support tools.

B. AI-based Classification for Autism Spectrum Disorder Detection using Video Analysis

The study employs AI-based classification techniques, including machine learning algorithms and deep neural networks, to analyze videos of children's interactions and detect behaviors commonly associated with ASD, such as restricted movements, limited eye contact, and abnormal posture. This approach contributes to automating the early detection process, reducing the reliance on subjective human judgment.

C. A Machine Learning Approach for Early Detection and Diagnosis of Autism and Normal Controls and Estimating Severity Levels Based on Face Recognition

The study applies facial feature extraction techniques to assess and classify children with Autism Spectrum Disorder (ASD) and neurotypical controls, focusing on the behavioral cues visible in facial expressions. Additionally, the research estimates the severity of ASD based on the observed features, aiming to improve diagnostic precision by utilizing non-invasive face recognition methods. This approach significantly contributes to the automation of the ASD diagnosis process and the estimation of severity levels, which can be crucial for early intervention.

D. Differentiating from Existing Work

- **Video-Based ASD Detection:**
  While existing works focus on video analysis to detect ASD behaviors, our project incorporates advanced deep learning models to enhance accuracy and efficiency in detecting ASD-related traits from video data.
- **AI-Powered Chatbot Assistance:** Unlike other studies, our system integrates an AI-powered chatbot that provides real-time support, offering personalized advice, emotional guidance, and valuable resources to parents and caregivers.
- **Interactive Support System:** Our project goes beyond detection and diagnosis by offering a holistic, interactive support system that allows parents to engage with the system for continuous updates, helping them navigate the complexities of ASD in their child.
- **Real-Time Recommendations:** The system also includes recommendations for therapy and interventions, offering personalized suggestions based on the analysis, which is an enhancement not seen in many existing ASD detection projects.
- **Comprehensive Solution:** Combining diagnostic accuracy with ongoing support, our project addresses both technical and emotional needs, unlike existing works that primarily focus on either detection or clinical assessment.

III. SYSTEM DESIGN AND METHODOLOGY

A. Data Collection and Preprocessing

Data collection: The dataset consists of video recordings of children engaged in different activities such as social interactions, facial expressions, and body movements. These videos are sourced from various sources.The dataset consists of both Autistic and non-Autistic videos.

Preprocessing: The videos are preprocessed by extracting key features such as facial expressions, gestures, and body movements. Frames are analyzed to capture subtle behavioral cues associated with ASD. This step involves image enhancement and normalization to prepare the data for the machine learning model.

B. Deep Learning-Based ASD Detection

The BI-LSTM (Bidirectional Long Short-Term Memory) model is employed in this system to process video frames in sequences, enabling the model to capture both temporal context and long-range dependencies of a child's behavior. By analyzing the sequence of facial expressions, gestures, and body movements over time, the model is able to recognize patterns associated with Autism Spectrum Disorder (ASD).

Feature Recognition: The BI-LSTM model is used to recognize key features from the video data by processing frames in both forward and backward directions. This bidirectional processing allows the system to capture not only past behavioral cues but also anticipate future responses, which is vital for detecting subtle signs of ASD that might emerge over time. This helps the model to recognize patterns like reduced eye contact, delayed social interactions, and repetitive movements — all behaviors commonly associated with ASD.

Sequential Context: The model can learn complex dependencies in the sequence of video frames, where behaviors such as delayed responses or social withdrawal become more apparent when analyzed over time. The sequential nature of BI-LSTM allows for a more accurate identification of ASD traits by understanding the flow of behavior and recognizing temporal patterns that could be missed by analyzing individual frames alone.

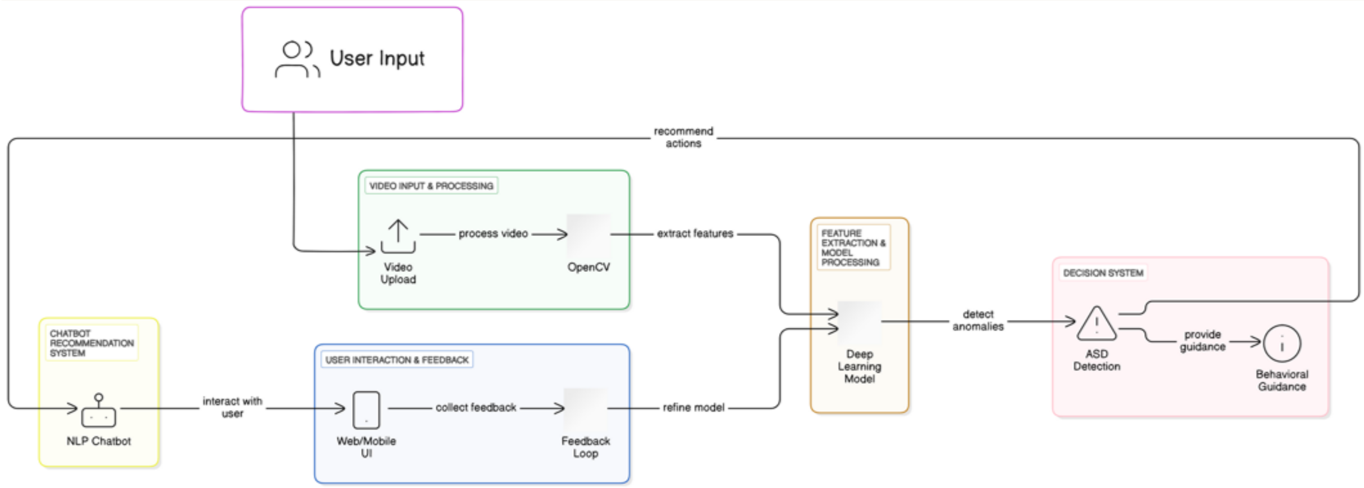

Fig. 1. System Architecture Diagram

C. Chatbot Integration and Activity Recommendations

As part of the AI-powered chatbot integrated into the ASD detection system, the severity level of Autism Spectrum Disorder (ASD) is assessed based on the answers provided to a series of diagnostic questions. These questions are designed to evaluate the social interactions, communication skills, and behavioral patterns of the child, which are key indicators of ASD. Based on the responses, the system computes a severity score, categorizing the level of ASD into Low, Moderate, or High severity.

Severity Scoring: The system uses a set of predefined questions focused on typical ASD traits such as eye contact, response to social cues, and repetitive behaviors. Each response is scored based on its alignment with known ASD indicators. The total score is used to classify the severity of the disorder, helping to determine the appropriate interventions and therapies.

Activity Recommendations: Once the severity level is determined, the chatbot provides personalized recommendations for activities that are tailored to the child's severity level. For example:

- Low Severity: The chatbot suggests basic social interaction exercises, such as role-playing games or interactive storytelling to encourage social engagement.
- Moderate Severity: The chatbot recommends communication-building activities, such as video-based therapy or guided interaction tasks, focusing on improving verbal and non-verbal communication.
- High Severity: The chatbot offers suggestions for intensive therapy sessions or structured behavioral interventions like Applied Behavior Analysis (ABA), as well as recommending local therapy centers or support groups.

D. System Integration and Testing

Component Integration: The system integrates the following modules:

1) Video Processing Module: Handles video upload and frame extraction using OpenCV. It ensures that video data is correctly captured, processed, and passed to the prediction model for classification.
2) Prediction Model: Uses BI-LSTM or any chosen deep learning architecture to classify behaviors as either Autistic or Non-Autistic based on the extracted features from the frames.
3) Chatbot Interface: Delivers personalized recommendations based on the severity of the detected ASD. The chatbot provides guidance on therapeutic activities and links to relevant resources based on the severity score of the child.

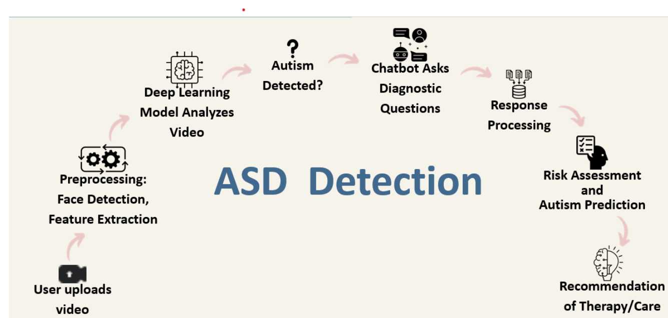

Fig. 2. Flow Diagram

Data Flow: A video is uploaded by the user. The video processing module extracts frames from the video. The frames are processed and passed through the model, which classifies the video data as Autistic or Non-Autistic. Based on the output, the severity level is assessed, and personalized recommendations are generated by the chatbot based on severity scores (Low, Moderate, or High).

User Interface (UI): A simple and intuitive interface is designed for ease of use by parents, caregivers, or clinicians. Users can easily upload videos, receive the results of ASD detection, and interact with the chatbot for further recommendations.

Integration with Backend Services: The backend services, such as the prediction model and chatbot logic, are hosted on a server that interfaces with the frontend (the user interface). This allows real-time processing of videos and immediate feedback for the users.

IV. Implementation Details

The implementation process begins with dataset collection, where video data is gathered from publicly available datasets containing both Autistic and Non-Autistic individuals. The dataset includes videos that capture various behavioral patterns and interactions, which are essential for training the model. The next step involves frame extraction. Using OpenCV, individual frames are extracted from each video at a fixed rate. This process converts the video into a sequence of frames, capturing the dynamic behavior of the individual. Due to the imbalance in the dataset, with fewer non-autistic samples compared to autistic ones, data augmentation techniques are applied to balance the data. Techniques such as horizontal flipping, slight rotations, and scaling are used to generate additional training samples, thus addressing the data imbalance and enhancing the model's ability to generalize to new, unseen data. Once the dataset is prepared and augmented after the frames are extracted, they are passed through MediaPipe for feature extraction. MediaPipe detects key points on the individual's face and body, such as facial landmarks and joint positions, which are crucial for analyzing gestures, facial expressions, and body movements. These key points are transformed into feature vectors, which represent the behavioral characteristics of the individual.

The extracted features are then processed through a BI-LSTM (Bidirectional Long Short-Term Memory) model for training. The BI-LSTM model is specifically designed to handle sequential data, making it well-suited for analyzing temporal dependencies between frames. The model learns to recognize patterns in facial expressions, gestures, and body movements that distinguish between autistic and non-autistic individuals. It is trained on the augmented dataset and is optimized to improve accuracy in classifying the behavior as either autistic or non-autistic.

After training and testing the model, the system is ready to make predictions. When a new video is uploaded, the model predicts whether the individual in the video exhibits signs of Autism Spectrum Disorder. Furthermore, the system integrates a chatbot to provide personalized support. The chatbot presents a series of questions to the user (usually a caregiver or clinician) to assess the severity of the autism spectrum disorder. Based on the answers, the system calculates a severity score and categorizes the condition into Low, Moderate, or High severity. Depending on the severity level, the chatbot suggests appropriate activities, such as social skills exercises, communication games, and behavioral therapy, to assist in the individual's development.

A. Dataset Collection and Preprocessing

The dataset for the Autism Spectrum Disorder (ASD) detection system is gathered from publicly available datasets containing video data of both Autistic and Non-Autistic individuals. These datasets are crucial for training the model to recognize behavioral patterns specific to ASD. The videos in the dataset represent various daily activities, gestures, facial expressions, and interactions, which are key to understanding behavioral differences.

B. Frame Extraction

After the dataset collection, the next step is to process the video data for the model. Frame extraction is performed using OpenCV, a powerful library for computer vision. The video is broken down into individual frames at a fixed rate, usually 30 frames per second. This frame-by-frame breakdown captures the dynamic behavior of the individual over time, which is critical for analyzing sequential patterns in facial expressions, gestures, and movements. The extracted frames are stored and prepared for further processing, allowing the system to analyze each frame individually while preserving the temporal dependencies between them.

C. Augmentation Techniques

Since the dataset is imbalanced, data augmentation techniques are employed to enhance the model's ability to generalize. Augmentation involves artificially increasing the size of the training set by applying random transformations to the original frames. Techniques such as horizontal flipping, rotation, and scaling are used to create new variations of the existing frames. These transformations ensure that the model learns robust features and is not overly sensitive to small variations in the input data, such as orientation or perspective. By augmenting the Autistic samples, the model becomes more capable of recognizing behavioral patterns from a diverse range of inputs, improving its accuracy in classification tasks.

D. Feature Extraction

This is a critical phase in the system as it transforms raw data (individual video frames) into meaningful features that can be used by the machine learning model. The feature extraction is performed using MediaPipe, a framework specifically designed for analyzing facial and body landmarks. MediaPipe provides tools for detecting key points on the face and body, such as facial landmarks (e.g., eyes, nose, mouth) and joint positions (e.g., elbows, shoulders, knees). These landmarks represent the critical physical markers that can be used to analyze behavior and detect subtle differences between autistic and non-autistic individuals.

The extracted key points are then converted into feature vectors. These vectors represent the unique behavioral characteristics of the individual in the video. For example, variations in facial expressions, gaze direction, and the way the individual uses gestures to communicate can be captured in the feature vectors. These features are essential for understanding how the individual interacts with their environment, providing valuable insights for detecting autism. Since these features are crucial for capturing behavioral patterns related to autism, accurate and robust feature extraction is vital to the success of the mode

E. BI-LSTM Model

. In the context of autism detection, the BI-LSTM model analyzes the temporal patterns in the extracted features (such as facial landmarks, body movements, and gestures) over time. These features are sequential, meaning the model needs to understand how an individual's behavior evolves throughout the video. By considering both the previous and future frames, the BI-LSTM can learn complex relationships between the behavioral patterns that are indicative of autism. For example, the way an individual's facial expression changes over time, or the sequence of hand gestures, may reveal significant behavioral cues that differentiate autistic and non-autistic individuals.

During training, the BI-LSTM model is optimized to minimize the classification error by learning from both the augmented Autistic and Non-Autistic data. This allows the model to improve its prediction accuracy, ensuring that it can classify a given video as either autistic or non-autistic with high confidence. Additionally, since the model is trained on sequential data, it is also capable of identifying subtle behavioral changes and inconsistencies that may not be immediately apparent from individual frames. This makes BI-LSTM a powerful tool for temporal pattern recognition in video-based autism detection systems.

F. Chatbot Integration

The AI-powered chatbot integrated into the ASD detection system plays a critical role in assessing the severity of Autism Spectrum Disorder (ASD) based on user responses to a series of diagnostic questions. These questions are designed to evaluate key ASD traits such as social interactions, communication skills, and behavioral patterns. By analyzing the answers provided by the user, typically a parent or clinician, the system computes a severity score and categorizes the level of ASD into Low, Moderate, or High severity. This process helps identify the degree of the disorder and determine the most appropriate interventions.

The severity scoring is based on a predefined set of questions that focus on common ASD traits, such as the child's eye contact, response to social cues, and repetitive behaviors. Each answer is scored according to its relevance to these traits, and the total score is used to assess the severity level of ASD. This classification enables caregivers and clinicians to better understand the individual's needs and plan for suitable therapies and interventions.

Once the severity level is determined, the chatbot provides personalized activity recommendations tailored to the child's specific needs. For Low severity, the chatbot suggests simple social interaction exercises, such as role-playing games or interactive storytelling, to help encourage social engagement. For Moderate severity, it recommends communication-building activities like video-based therapy or guided interaction tasks, focusing on enhancing both verbal and non-verbal communication. For High severity, the chatbot may suggest more intensive therapy sessions, such as Applied Behavior Analysis (ABA), or guide the user to local therapy centers or support groups that offer specialized interventions. By combining severity scoring with real-time activity suggestions, the chatbot ensures that users receive targeted support to help manage and address ASD effectively.

V. EXPERIMENTAL RESULTS ANALYSIS

The model's performance was evaluated using the validation dataset, and the following key metrics were obtained:

- **Accuracy**: 96%, indicating that the model correctly predicted the classification of 96% of the test samples.
- **Precision**: 97%, meaning that 97% of the positive predictions made by the model were correct, indicating a relatively low number of false positives.
- **Recall**: 95%, meaning that the model successfully identified 95% of the actual positive cases, with a few false negatives.
- **F1-Score**: 96%, balancing precision and recall. A high F1-score of 96% indicates that the model is performing well in terms of both identifying ASD cases and minimizing false positives.

These results show a strong performance, with the model achieving high precision, recall, and F1-score, although there is a slight reduction from perfect scores, reflecting the slight imperfections in real-world data.

A. Chatbot Integration and Severity Assessment

Alongside the model's classification ability, the system also includes a chatbot that assesses the severity of Autism Spectrum Disorder based on a set of diagnostic questions. The chatbot evaluates the responses provided by users (typically parents or clinicians) and computes a severity score that categorizes the level of ASD as Low, Moderate, or High.

- Low Severity: The chatbot suggests light interaction activities to promote basic social skills and communication.
- Moderate Severity: The chatbot recommends video-based therapy or structured interaction tasks to improve communication and social engagement.
- High Severity: Intensive therapy sessions and behavioral interventions like Applied Behavior Analysis (ABA) are suggested.

The integration of the chatbot not only adds a layer of interactivity but also allows the system to offer personalized support to individuals based on their specific needs and severity levels

VI. DISCUSSION

The user input—a video file containing the behavior of the individual to be analyzed. Upon uploading the video, the system processes the video frames to extract relevant features. This is done using OpenCV, which reads and extracts each frame from the video at a specified rate. These frames are then passed through MediaPipe, a library used for pose extraction. MediaPipe detects key points on the individual's face and body, which are essential for understanding facial expressions, hand gestures, and body movements. These extracted key points are then used as the input for the deep learning model.

The extracted frames are then normalized to ensure uniformity across all input data. Frame normalization includes resizing each frame to a fixed size and standardizing the pixel intensity values, ensuring that all frames are processed uniformly regardless of their original resolution or quality. This step ensures that the model can focus on the essential features of the frames without being affected by external inconsistencies such as lighting or video quality.

After the frame extraction and normalization steps, the processed data is fed into a BI-LSTM model (Bidirectional Long Short-Term Memory model). This model, designed to capture both past and future dependencies in sequential data, analyzes the temporal features of the video frames. The BI-LSTM model classifies the behavior of the individual, determining whether the video depicts an Autistic or Non-Autistic individual based on behavioral patterns such as facial expressions, body posture, and interaction styles. The model provides a prediction that categorizes the individual as either autistic or non-autistic.

Once the classification is complete, the system proceeds to chatbot-based recommendations. A set of questions is presented to the user (typically a caregiver or clinician), which helps assess the severity level of the individual's condition. Based on the responses, a severity score is calculated, categorizing the ASD condition into Low, Moderate, or High severity. The chatbot then provides tailored activity recommendations based on the severity level. These recommendations include activities such as communication games, behavioral therapy exercises, and social skill training, ensuring that the interventions are personalized to the individual's needs.

The results obtained from the Autism Spectrum Disorder (ASD) detection system highlight the effectiveness of using video-based analysis in combination with deep learning techniques, specifically the BI-LSTM model, for early detection of ASD. This model, integrated with chatbot assistance, provides both diagnostic and therapeutic functionalities, making it a comprehensive tool for supporting individuals with ASD. In this section, we discuss the strengths, limitations, and potential future directions of the system.

A. Strengths of the System

High Accuracy and Precision: The model achieved an impressive 96% accuracy, demonstrating its ability to correctly classify both autistic and non-autistic individuals from video data. The high precision (0.97) indicates that the model rarely misidentifies non-autistic individuals as autistic, which is crucial for ensuring the reliability of the detection process in clinical environments.

Effective Use of Video Data: The use of frame extraction combined with MediaPipe for feature extraction enabled the model to capture important behavioral cues, such as facial expressions, gestures, and body movements. These features are essential for understanding the subtleties of ASD behavior, which may not be easily detectable through conventional diagnostic methods.

Personalized Support through Chatbot: The integration of the chatbot for severity assessment is a key strength of this system. It provides real-time recommendations tailored to the severity level of the individual, making it a dynamic tool for early intervention and continuous support. The chatbot also ensures that the suggestions are adapted to the individual's needs, promoting better engagement with the therapeutic process.

Data Augmentation to Handle Imbalanced Dataset: The application of data augmentation techniques, such as horizontal flipping and rotation, to balance the dataset significantly improved the model's robustness. This approach mitigates the effects of the imbalanced distribution between autistic and non-autistic samples, ensuring that the model performs well across both classes.

B. Limitations and Challenges

Dataset Size and Diversity: Although the model demonstrated strong performance on the test dataset, the dataset used for training was relatively small, consisting of video samples from publicly available sources. A larger and more diverse dataset that includes a broader range of individuals and behaviors would help improve the generalizability of the model. Variations in age, ethnicity, and environmental factors could potentially affect the model's ability to detect ASD across diverse populations.

Real-time Video Analysis: While the system performed well on pre-recorded video data, real-time video analysis may present challenges in terms of processing speed and accuracy. The system would need to be optimized for real-time applications, which could involve addressing issues related to frame rate, latency, and hardware capabilities.

Limited to Behavioral Features: The system relies primarily on behavioral cues (e.g., facial expressions, gestures) for detecting ASD. While these cues are important, they may not capture the full spectrum of ASD symptoms. Incorporating additional diagnostic modalities, such as speech analysis or genetic testing, could further enhance the system's diagnostic accuracy.

Severity Scoring: The severity scoring mechanism used by the chatbot is based on predefined questions related to key ASD traits. While this approach provides a helpful framework, it is important to note that severity levels can vary widely across individuals. Therefore, the system's scoring should be continuously refined based on clinical feedback to ensure it accurately reflects the nuanced nature of ASD.

VII. Conclusion

This project presents an innovative approach to Autism Spectrum Disorder (ASD) detection by combining video-based analysis, deep learning models, and chatbot assistance to create a comprehensive and effective solution. The system leverages advanced techniques such as frame extraction, feature extraction with MediaPipe, and a BI-LSTM model to accurately differentiate between autistic and non-autistic individuals based on behavioral patterns. This technology offers a promising avenue for early diagnosis, which is crucial for implementing timely interventions that can significantly improve the quality of life for individuals with ASD.

The high accuracy achieved in the system demonstrates its ability to reliably identify ASD traits from videos, making it a powerful tool for early detection. The integration of a chatbot enhances the system's capabilities by providing real-time severity assessment and personalized recommendations for activities that cater to the needs of individuals at different levels of severity. This ensures that interventions are not only timely but also tailored, offering support that adapts to each individual's unique requirements.

The augmentation techniques used to balance the dataset and the model's generalization capabilities further emphasize the robustness of the system. By utilizing AI-powered technologies, the project offers an accessible, scalable solution that can be deployed in both clinical and home settings, providing invaluable support to parents, caregivers, and clinicians alike.

In conclusion, the project demonstrates the power of combining deep learning, computer vision, and conversational AI to create a system that provides early detection, personalized interventions, and continuous support for individuals with Autism Spectrum Disorder.

VIII. Future Work

While the current ASD detection system demonstrates strong performance and potential, there are several areas for improvement and expansion that could enhance the system's effectiveness, usability, and impact. The following outlines key directions for future work:

Expansion of Dataset: The current system was trained on a limited dataset, which may impact its generalization to diverse real-world scenarios. Expanding the dataset to include more varied and representative samples of both autistic and non-autistic individuals, including different age groups, cultures, and behavioral characteristics, would help improve the model's accuracy and robustness.

Multimodal Data Integration: While the system relies on video analysis, integrating additional data modalities such as audio (e.g., speech patterns), sensor data, and textual inputs could provide a more comprehensive understanding of ASD symptoms. This multimodal approach would enable the model to capture a wider range of behavioral cues, further improving its diagnostic capabilities.

References

[1] Andrea Zunino, Pietro Morerio, Andrea Cavallo, Caterina Ansuini, Jessica Podda, Francesca Battaglia "Video Gesture Analysis for Autism Spectrum Disorder Detection"

[2] Shivani Pandya, Swati Jain, Jai Prakash Verma, "AI based Classification for Autism Spectrum Disorder Detection using Video Analysis."

[3] Shubhangi D.C, Baswaraj Gadgay, Shaista Farheen, M.A Waheed "A Machine Learning Approach for Early Detection and Diagnosis of Autism and Normal Controls and Estimating Severity Levels Based on Face Recognition."

[4] Mayar Elshentenawy, Merna Ahmed, Mariam Elalfy, Ahmed Bakr, Mahmoud Heidar, Eslam Amer, "Intellibot: A Personalized Behavioral Analysis Chatbot Framework Powered by GPT-3 ."