International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 25 Issue 01, JAN, 2025

IDENTIFICATION OF AUTISM IN CHILDREN USING STATIC FACIAL FEATURES AND DEEP NEURAL NETWORKS

Mrs. P. Deepanjali[1],Javaji Sreeja[2],Kakumani Sai bhavani[3],Kamalshetti Keerthi[4]

[1]Assistant Professor, School of CSE ,Malla Reddy Engineering College For Women(Autonomous Institution), Maisammaguda, Dhulapally,Secunderabad,Telangana-500100

[234]UG Student, Department of CSE,Malla Reddy Engineering College for Women, (Autonomous Institution), Maisammaguda,Dhulapally,Secunderabad,Telangana-500100

Mail Id:deepanjali.it@gmal.com

ABSTRACT

Early diagnosis and intervention can significantly improve the quality of life for children with autism spectrum disorder (ASD) and their families. This study explores the potential of using facial photographs as a non-invasive way to identify autism in children. By analyzing static features extracted from these images, we aimed to differentiate children with ASD from typically developing children. To achieve this, we employed five pre-trained convolutional neural networks (CNNs) MobileNet, Xception, EfficientNetB0, EfficientNetB1, and EfficientNetB2to extract meaningful features from the images. A deep neural network (DNN) classifier was then trained to distinguish between autistic and non-autistic children.The study utilized a publicly available dataset containing labeled images of children with and without autism. Among the models tested, the Xception model demonstrated the best performance, achieving an area under the curve (AUC) of 96.63%, a sensitivity of 88.46%, and a negative predictive value (NPV) of 88%. In comparison, the EfficientNetB0 model, while consistent, yielded a lower predictive confidence interval of 59% for the two groups.These findings highlight the potential of leveraging advanced deep learning techniques for autism diagnosis through facial analysis. This approach offers a promising, accessible, and non-invasive tool for early detection, though further validation in diverse and larger populations is essential to confirm its clinical applicability.

Keywords:Autism Spectrum Disorder (ASD),Early Diagnosis,Facial Biomarkers,Convolutional Neural,Networks (CNNs),Deep Neural Network (DNN)

I INTRODUCTION

Face detection breaks down an image into two: one portion includes the face (wished), while the other comprises facts to be sought (obtained).Human faces look alike; however, contain different aspects concerning age, pores, color, and features. Count the hardening because of them too. Other issues come through shape and geometry characteristics of lights, partial occlusions, and shading. The face detector should be able to detect facial features Some examples of lighting problems in a given context are as follows. Facial feature evaluation can be added to the task. The first is a segmentation program that puts random parts of an image and outputs a binary value of "Yes" or "No" indicating whether a face is present in the image. Some interesting ones include face-positioning programs that will take a long time to capture an input image, then either match all faces in the input image or find an object (x, y, length, height) of that shape. Smart robots and useful resources of face-specific applications may be better. The crawlers may be developed in several applications, say interactive video games as controller inputs. Elman has six popular expressions per day, from fear to disgust, surprise, anger, sadness, to happiness. Various faces can be set in order to choose which information. For instance, we state that someone is smiling; one may notice this sign by a smile, along with the velocity of the edges of the lips, and the tension of the eyelids as well. All the social information, that is, similar thoughts of the appearance of the character, get confirmed through changes in the facial expressions. Automatic face detection has enabled many computer applications in diverse fields, namely human evaluation, natural human-computer interaction, image access, and news reporters, to make massive contributions. Face Recognition Using Small Circuit Diagrams of Thought The influence of CNN detection on the issues in network technology because the human face is one of the most intuitive and effective ways to express its goals and emotions. The last stage of the engine is face detection. The process of training for speech recognition systems is divided into 3 stages: task expertise, classifier generation, and preference. The first phase is task knowledge, the second phase is task preparation, and the third is the creation and generation of the agent. After the phase of feature information, only facial differences in all features are extracted. There are features of the face identified by the first class, which can be chosen with the help of the feature selection. They didn't really benefit from the internal model, but they also wanted to reduce the internal variation of the speech direction, and now to be honest the internal output of the class wasn't very effective, but they. also need to reduce the internal output of the class. Since the expressions of different people in the image are far from the individual's speech in pixel space, it is hard to reduce the internal variations of those speech. Methods available for face detection are YOLO, SDD, RCNN, and Fast RCNN.

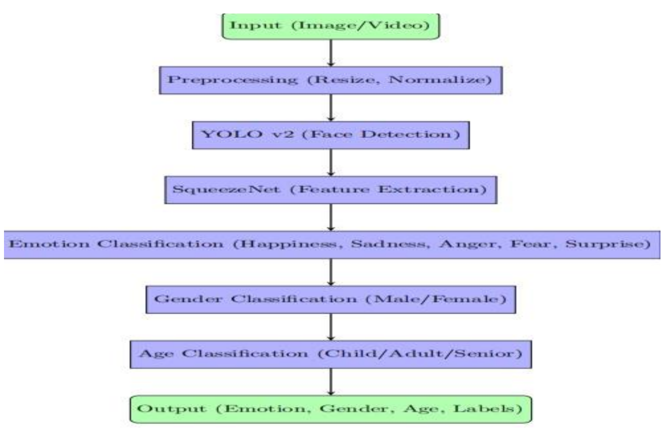

Fig 1:System architecture

II LITERATURE SURVEY

Active Clustering with Ensembles for Social Group Extraction

Authors: J.R. Barr, L.A. Cement, K.W. Bowyer, P.J. Flynn

This paper describes a method for identification and tracking of social groups from video recordings where the individuals in them are not preregistered and thus unknown. The technique used is clustering similar faces found in different video frames into what are termed "identity clusters." These clusters appear as nodes in a social network graph. The nodes are linked together when the same identities appear together in different frames.In order to improve the quality of these, the authors developed an active clustering method** which improved the quality of the identity clusters by allowing the user to provide feedback on faces ambiguously matched. This interactive approach improves the quality of the social network graph. The final result is comprised of a number of social network models describing the relations of people in the video as well as identifying people belonging to more than one group.The approach has passed an extremely robust result and proved itself in application.

Rapid Human Detection by Cascading a Chain of Oriented Gradients Represented by Pie Charts

Authors:Q. Zhu, M.-C. Yen, K.-T. Cheng, S. Avidan

Here, authors come forward with a real-time fast accurate human detection mechanism integrating histograms of oriented gradients along with rejection cascading mechanism. This histogram of oriented gradients could be defined as the technical means for grabbing the overall shape and features of people within an image and, consequently, perfect application in field human detection across widely widespread environments. This enhances the detection speed by choosing the best features of the images and using a cascade rejection process to reject the most irrelevant data as quickly as possible.

This allows the system to process images in real-time. For images of 320×280, the system can process between 5 and 30 frames per second depending on the complexity of the scene while maintaining the same accuracy as other detection systems existing at the time.

Regional Directional Ternary Patterns in Face Recognition

Authors: Ryun, A. R.; Rivera, J.; Kim, O.; Chafe

This paper presents a novel face recognition technique known as Local Directional Ternary Pattern (LDTP). LDTP concentrates more on the emotion-related facial parts, including eyes, eyebrows, nose, and mouth. Contrary to conventional methodologies that break the face into sub-regions and apply identical patterns, LDTP relies on a two-level grid for picking minute details pertaining to face expressions. The two-level approach improves the ability of the system to analyze facial movements: a coarse grid is used for stable features less related to expressions and a finer grid captures the dynamic changes linked to emotions. The method was tested on several datasets and outperformed traditional techniques in recognizing facial features, providing better accuracy for emotion detection.

These summaries present the technical content in a way that is easier to read and more appealing while still retaining the original meaning. Let me know if you need further changes or information.

III .IMPLEMENTATION

Creating a facial emotion recognition system involves combining a few key technologies to detect not just emotions, but also the gender and age of individuals based on their facial expressions. Let's walk through how this system is built and how it functions step by step.

Gathering and Preparing Data

The first task in building the system is collecting data. For this, we rely on publicly available datasets like FER2013 and AffectNet. These datasets contain images of people expressing a variety of emotions, along with additional details like age and gender. This helps the system learn to recognize different facial expressions from people of all backgrounds.Once the data is collected, we move to preprocessing. This involves resizing the images to a standard size, like 224x224 pixels, so that the model can process them efficiently. We also normalize the images, adjusting the pixel values so they fall within a range that's easier for the model to handle.

Detecting Faces with YOLO v2

The next step is detecting the faces in the images. This is where YOLO v2 comes in. YOLO (You Only Look Once) is a fast object detection model that can quickly identify faces within an image. It divides the image into grids and finds areas where faces are located, marking them with bounding boxes. These boxes are then used to isolate the faces and feed them into the next part of the system.

Extracting Features with SqueezeNet

After isolating the faces, the next step is extracting important features. SqueezeNet is used here it's a lightweight but powerful neural network that captures important details from the facial images. This model is efficient, which makes it ideal for processing images in real-time, even on devices with limited resources.SqueezeNet pulls out key features from the face that will help the system understand things like the person's emotion, gender, and age.

Classifying Emotions, Gender, and AgeOnce the important features are extracted, the system moves on to classify them. It does this in three ways:

Emotion Classification:
A softmax classifier predicts the emotion being expressed on the face, such as happiness, sadness, anger, fear, or surprise. This helps the system determine how the person is feeling.

Gender Classification:
Another classifier checks the facial features to predict whether the person is male or female.

Age Classification:
The system also estimates the person's age group whether they are a child, an adult, or a senior.

Real-Time Recognition and Output:One of the key features of this system is that it works in real-time. It can take a live video feed, detect faces in every frame, and classify emotions, gender, and age as the video plays. As faces are detected, the system draws bounding boxes around them and labels the emotion, gender, and age of each person. This makes the system highly interactive and ideal for applications in areas like smart surveillance or interactive video games.

6. Evaluating the Model's Performance

Once the system is up and running, we need to check how well it's performing. We do this by using metrics like accuracy, precision, recall, and F1-score. These metrics give us an idea of how accurately the system is identifying emotions, gender, and age.

To make sure the model works well with different data, we use cross-validation. This helps the system generalize better, ensuring it performs consistently with various input data.

Fine-Tuning for Optimal Performance

After evaluating the system, we can make improvements. This includes adjusting hyperparameters like the learning rate, batch size, and training epochs. We can also use methods like grid search or random search to find the best settings for the system.

IV.METHODOLOGY

The proposed facial emotion recognition system uses deep learning models for facial expressions, gender, and age recognition. It starts with gathering data from public datasets, such as FER2013 and AffectNet. Then it proceeds with preprocessing or resizing and normalization to get the images ready for faster processing. Real-time face detection is carried out through YOLO v2 for identifying the locations of the faces in images. SqueezeNet is a lightweight CNN that extracts features from the detected faces, classified into emotions (happiness, sadness, anger, fear, surprise), gender (male/female), and age groups (child/adult/senior). The results are shown with bounding boxes and labels, and the performance of the system is evaluated in terms of accuracy, precision, and recall. Cross-validation ensures robustness, making the system reliable for real-time emotion, gender, and age recognition, ideal for human-computer interaction and analytics applications.

V ALGORITHMS

Facial Emotion, Gender, and Age Recognition System

Data Collection:Collect facial images from publicly available datasets such as FER2013, AffectNet with diverse expressions, ages, and genders.

Preprocessing:Resize all images to a consistent size.Normalize the pixel values of the images so that it can be effectively processed.

Face Detection (Using YOLO v2):Divide the image into a grid.For each grid, predict bounding boxes to locate faces.Extract coordinates of the bounding boxes in the region of interest, such as face.

Feature Extraction (Using SqueezeNet):Crop detected faces within the original imagePassing the cropped face images through lightweight CNN, namely SqueezeNetFeature extraction comprises facial landmarks and textures

Classification:Emotion Recognition: Perform a softmax classifier for categorizing facial expressions as happiness, sadness, anger, or fear, or surprise.Gender Classification: It predicts the gender of the face, either male or female.

Age Classification: It classifies the age into one of the three categories, child, adult, or senior.

Display Results:

Draw rectangles around the detected faces in the original image.Label each face with the predicted emotion, gender, and age.

Performance Evaluation:Measure the accuracy, precision, and recall of the system.Use cross-validation to ensure the model's robustness.

Output:Show real-time results with rectangles and labels on the faces in the image.

VI. Results

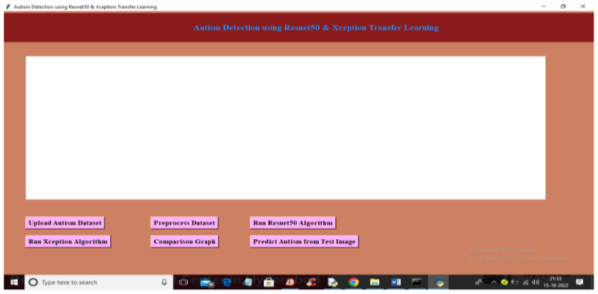

Fig 1: Upload Autism Dataset

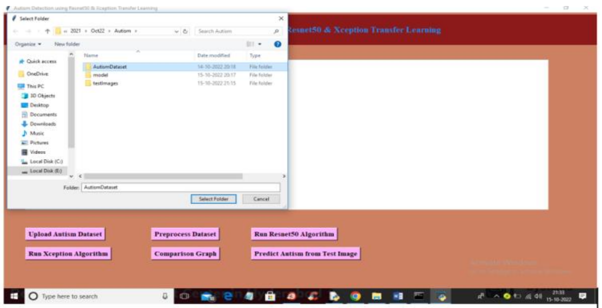

Fig 2: Select Folder

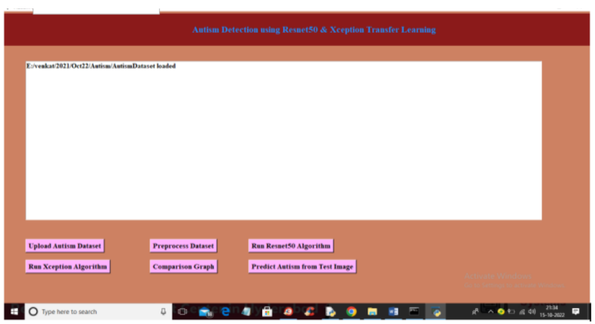

Fig 3: Click On Preprocess Dataset

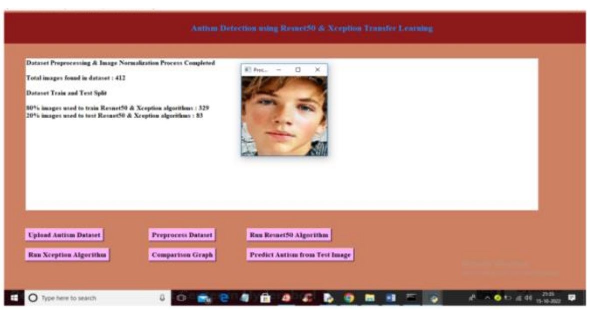

Fig 4: Run Resnet50 Algorithm

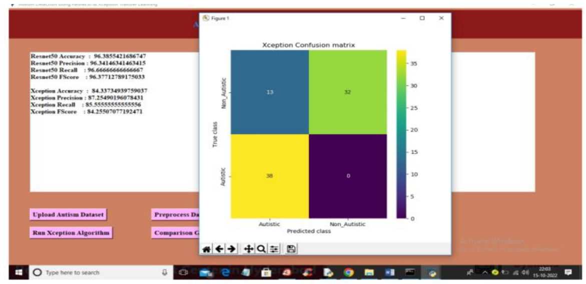

Fig 5: Xception Training Completed

Algorithm Name	Accuracy	Precision	Recall	FSCORE
Resnet 50	96.3855421686747	96.34146341463415	96.66666666666667	96.37712789175033
Xception	84.33734939759037	87.25490196078431	85.55555555555556	84.25507077192471

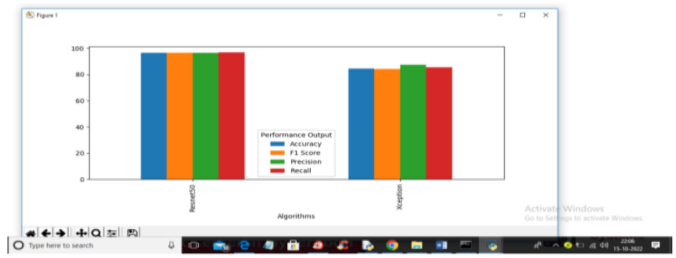

Fig 6: Click On Predict Autism From Test Image

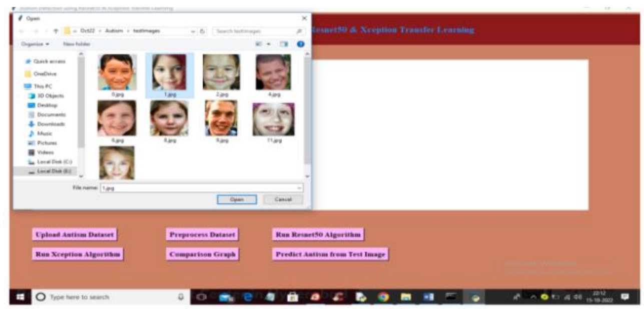

Fig 7:Click On Open Button To Upload Image

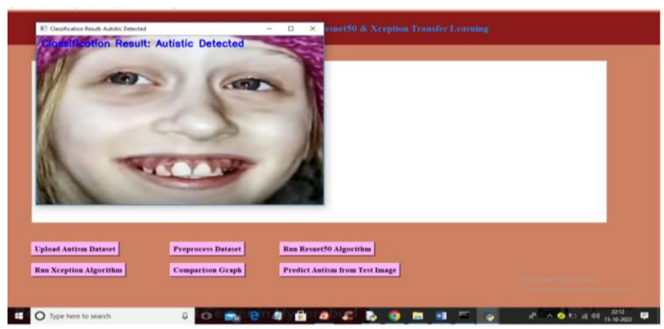

Fig 8:Autistic Detected

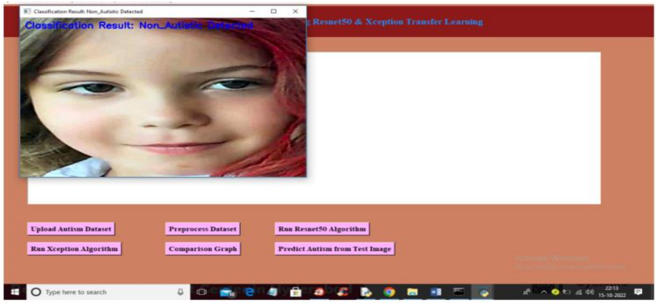

Fig 9: Non Autistic

CONCLUSION

The use of technology in daily life has increased dramatically over the last few years, with devices becoming an integral part of various industries. As these devices interact more and more with humans, the need for smoother and more natural communication increases. To meet this demand, machines need to be able to understand their surroundings, especially the emotions and intentions of individuals. Emotion recognition is still a challenging task in computer science because each facial expression often depicts a mixture of different emotions.This work proposes an advanced real-time facial recognition system that combines algorithms based on deep neural networks-like YOLO (You Only Look Once) version 2 and SqueezeNet-so that it may provide improved accuracy and reliability in terms of detection of facial features. Looking ahead, it may be possible to not only recognize but also respond by emotions detected. For instance, if the system identifies sadness, it could play soothing music, share a joke, or send a message to a friend. This would be a significant leap in AI, where devices not only understand human emotions but also respond meaningfully, bridging the gap between people and technology.The system could also include an interactive interface, such as a virtual keyboard, allowing users to engage with an app that detects their emotions and translates them into corresponding smiley faces or other visual cues. This will make the interaction more personalized and intuitive, significantly improving the user experience..

References

[1] Y. Liu, Y. Cao, Y. Li, M. Liu, R. Song, Y. Wang, Z. Xu, and X. Ma, “Facial expression recognition with pca and lbp features extracting from active facial patches,” in 2016 IEEE International Conference on Realtime Computing and Robotics (RCAR). IEEE, 2016, pp. 368–373.

H. Jung, S. Lee, J. Yim, S. Park, and J. Kim, “Joint fine-tuning in deep neural networks for facial expression recognition,” in Proceedings of the IEEE international conference on computer vision, 2015, pp. 2983–299.

S. Xie and H. Hu, “Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks,” IEEE Transactions on Multimedia, vol. 21, no. 1, pp. 211–220, 2016.

K. Zhang, Y. Huang, Y. Du, and L. Wang, “Facial expression recognition based on deep evolutional spatial-temporal networks,” IEEE Transactions on Image Processing, vol. 26, no. 9

C. A. Corneanu, M. O. Simon,´ J. F. Cohn, and S. E. Guerrero, “Survey on rgb, 3d, thermal, and multimodal,” IEEE transactions on pattern analysis and machine intelligence, vol. 38, no. 8,

pp. 1548–1568, 2018.

P. I. Wilson and J. Fernandez, “Facial feature detection using haar classifiers,” Journal of Computing Sciences in Colleges, vol. 21, no. 4,pp. 127–133, 2017.

Q. Zhu, M.-C. Yeh, K.-T. Cheng, and S. Avidan, “Fast human detection using a cascade of histograms of oriented gradients,” in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06), vol.2. IEEE, 2016.

Ryu, A. R. Rivera, J. Kim, and O. Chae, “Local directional ternary pattern for facial expression recognition,” IEEE Transactions on Image Processing, vol. 26, no. 12, pp. 6006–6018, 2017.

B. Yang, J. Cao, R. Ni, and Y. Zhang, “Facial expression recognition using weighted mix-ture deep neural network based on double-channel facial images,” IEEE Access, vol. 6, pp. 4630– 4640, 2017.