Estd : 1984

KONGU ENGINEERING COLLEGE

Team - 8

Facial Expression Recognition for Children with Autism Spectrum Disorder Using Swin Transformer

Mentor Name: Dr. M. N. Kavitha

Team Members :

Kaviya Priya N - 21CDR019

Ranjith S R - 21CDR037

Rahul Vanchivel N - 21CDR035

Project Domain

- Artificial Intelligence
- Deep Learning
- Computer Vision
- Affective Computing
- Machine Learning

Problem Statement

The problem this project addresses is the challenge of accurately detecting and classifying facial emotions in children with Autism Spectrum Disorder (ASD) to aid in emotional recognition and improve social interactions.

Scope and Objectives

SCOPE:

- Targeted Emotion Recognition for Neurodiverse Individuals
- Advanced Deep Learning Techniques

OBJECTIVES:

- Precise Face Detection
- Emotion Classification with Swin Transformer

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
ViTFER: Facial Emotion Recognition with Vision Transformers. (MDPI)	Aayushi Chaudhari et al.	ResNet-18 for feature extraction from facial images. Vision Transformers for capturing spatial features and global relationships.	Computational Complexity. Performance relies heavily on the quality and diversity of datasets like IEMOCAP and MELD.	2022

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
Hybrid Vision Transformer with Temporal Convolution for Multi-Emotion Recognition. (Springer)	Kamal Zakieldin et al.	Vision Transformer (ViT) is used for analyzing individual frames and temporal Convolution Network (TCN) is employed for accurate emotion recognition.	Computational resource limitations. Dependent on a single GPU for certain applications.	2024

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
Speech Emotion Recognition using CNN-LSTM. (Springer)	Yuvjeet Arora et al.	Mel Spectrograms with CNN-LSTM.	Limited to speech-based recognition. Slightly lower accuracy of 88.50%.	2024

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
An adaptive tuning privacy-preservin g approach for multimodal emotion detection. (Springer)	Lihua Yin et al.	Employs a generalized random response-based differential privacy method for multimodal data, combining pre-aggregator and iterative mechanisms.	Limited to vehicle emotion detection. Computational overhead.	2024

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
Exploring AI Facial Recognition for Real-time Emotion Detection: Assessing Student Engagement in Online Learning Environments. (IEEE)	J. Shan et al.	Utilizes a conditional Generative Adversarial Network (cGAN) for training a generative model, enabling De-expression Residue Learning (DeRL).	Depends on the quality of facial expression databases, limiting generalizability. Computational complexity.	2024

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
Music Emotion Recognition Based on Deep Learning. (IEEE)	Xingguo Jiang et al.	Deep neural networks for music emotion recognition, including emotion models, feature extraction, and recognition algorithms.	Handling the complexity of music's emotional expression. Improving model generalization. Need for large, labeled datasets.	2024

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
Speech Emotion Recognition Based on Swin -Transformer (iopscience)	Zirou Liao et al.	Swin Transformer to classify speech emotions by transforming acoustic signals into spectrograms for emotion recognition.	Model may struggle with noisy or unclear audio signals, affecting emotion classification accuracy.	2023

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
Speech Emotion Recognition Using Capsule Networks. (IEEE)	Xixin Wu et al.	Capsule Networks were employed to analyze spatial relationships in speech spectrogram features.	Limited to spectrogram based speech features.	2022

Literature Survey

PAPER	AUTHOR	TECHNIQUES	DRAWBACKS	YEAR
Conversational emotion recognition using multimodal features. (IEEE)	Zheng Lian et al.	The transformer-based structure, multi-head attention, bi-directional GRU, and speaker embeddings are employed to capture temporal information, and speaker sensitive dependencies.	Limited in handling complex speaker dynamics and context-sensitive nuances in some situations.	2021

Existing Method

- Haar Cascade for Face Detection followed by Support Vector Machines (SVMs) for Emotion Classification.
- Viola-Jones Algorithm for Face Detection followed by Convolutional Neural Networks (CNNs) for Emotion Classification.
- Dlib Facial Landmark Detector followed by Recurrent Neural Networks (RNNs) with LSTMs for Emotion Classification.
- YOLOv4 for Real-Time Face Detection followed by Vision Transformer (ViT) for Analyzing Facial Emotions.
- RetinaFace for Accurate Face Detection followed by Multi-Scale Vision Transformer for Emotion Recognition.

Proposed Method

- The proposed emotion recognition system focuses on accurately detecting emotions in children with Autism Spectrum Disorder by utilizing a combination of MTCNN (Multi-task Cascaded Convolution Neural Network) for face detection and Swin Transformer for emotion classification.
- First, the system employs MTCNN to detect and extract faces from images ensuring precise identification of facial features.
- The extracted facial data is then processed using Swin Transformer, which captures spatial hierarchies and relationships between facial features.

LIST OF MODULES

The list of modules:

- Data Collection and Preprocessing
- Face Detection
- Emotion Classification
- Emotion Recognition Validation

MODULE IMPLEMENTATION

Data Collection and Preprocessing:

- Facial images corresponding to different emotional states are collected from a specified dataset directory.
- The collected images are then resized to a consistent size of 224x224 pixels and normalized to ensure consistency in the input data, making it suitable for training deep learning models.
- Additionally, data augmentation techniques are applied to increase the dataset's diversity by generating new images through transformations like rotation, shifting, and flipping, ensuring robust model training.

MODULE IMPLEMENTATION

Face Detection:

- The MTCNN model detects faces in the input image by identifying facial landmarks and locating the bounding box for each face.
- Once faces are detected, MTCNN crops the detected face regions, removing unnecessary background and isolating the facial features for further processing.
- The cropped face images are passed into the emotion recognition model (Swin Transformer) to classify the emotions based on facial expressions.

MODULE IMPLEMENTATION

Emotion Classification:

- Each cropped face image from the previous face detection step is resized and preprocessed to match the input specifications of the Swin Transformer.
- The Swin Transformer model is trained on the prepared dataset to classify emotions accurately.
- Cross-Entropy Loss is employed during training to improve the model's predictive accuracy for emotions such as happiness, sadness, and anger.

SCREENSHOTS

Data Collection and Preprocessing:

```
Processing images...
Processed 0 images in anger folder...
Processed 100 images in anger folder...
Processed 200 images in anger folder...
Processed 300 images in anger folder...
Processed 400 images in anger folder...
Processed 500 images in anger folder...
Processed 600 images in anger folder...
Processed 700 images in anger folder...
Processed 800 images in anger folder...
Processed 900 images in anger folder...
Processed 1000 images in anger folder...
Processed 1100 images in anger folder...
Processed 1200 images in anger folder...
Processed 1300 images in anger folder...
Processed 1400 images in anger folder...
Processed 1500 images in anger folder...
Processed 1600 images in anger folder...
Processed 1700 images in anger folder...
Processed 1800 images in anger folder...
Processed 0 images in contempt folder...
Processed 100 images in contempt folder...
Processed 200 images in contempt folder...
Processed 300 images in contempt folder...
```

SCREENSHOTS

Data Augmentation:

```
Processing folder 1/3: neutral
Starting augmentation for folder: neutral...
Processed 100 images so far in neutral.
Processed 200 images so far in neutral.
Processed 300 images so far in neutral.
Processed 400 images so far in neutral.
Processed 500 images so far in neutral.
Processed 600 images so far in neutral.
Processed 700 images so far in neutral.
Processed 800 images so far in neutral.
Processed 900 images so far in neutral.
Processed 1000 images so far in neutral.
Processed 1100 images so far in neutral.
Processed 1200 images so far in neutral.
Processed 1300 images so far in neutral.
Processed 1400 images so far in neutral.
Processed 1500 images so far in neutral.
Processed 1600 images so far in neutral.
Processed 1700 images so far in neutral.
Processed 1800 images so far in neutral.
Finished augmenting images for folder: neutral. Total images processed: 1880
Processing folder 2/3: sad
Starting augmentation for folder: sad...
Processed 100 images so far in sad.
Processed 200 images so far in sad.
```

SCREENSHOTS

Face Detection using MTCNN:

```
1 face(s) detected! Drawing bounding box...
```

SCREENSHOTS

Training Swin Transformer:

```
Loading datasets...
Loaded images (Indices 46500 to 50646).
Total images left after this batch: 0
Loading pre-trained Swin Transformer model...
Model weights loaded from swin_transformer_46500_images.pth.
Training the model on the remaining images...
Batch 1/41: Image 1/100 - Loss: 0.0000
Batch 1/41: Image 2/100 - Loss: 0.0000
Batch 1/41: Image 3/100 - Loss: 0.0000
Batch 1/41: Image 4/100 - Loss: 0.0000
Batch 1/41: Image 5/100 - Loss: 0.0000
Batch 1/41: Image 6/100 - Loss: 0.0000
Batch 1/41: Image 7/100 - Loss: 0.0000
Batch 1/41: Image 8/100 - Loss: 0.0000
Batch 1/41: Image 9/100 - Loss: 0.0000
Batch 1/41: Image 10/100 - Loss: 0.0000
Batch 1/41: Image 11/100 - Loss: 0.0000
Batch 1/41: Image 12/100 - Loss: 0.0000
Batch 1/41: Image 13/100 - Loss: 0.0000
Batch 1/41: Image 14/100 - Loss: 0.0000
Batch 1/41: Image 15/100 - Loss: 0.0000
Batch 1/41: Image 16/100 - Loss: 0.0000
Batch 1/41: Image 17/100 - Loss: 0.0000
```

SCREENSHOTS

Resizing images suitable for Swin Transformer:

```
Starting resizing in folder: G:\Sem_8\Resized_Faces\train
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\contempt
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\disgust
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\fear
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\happy
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\neutral
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\sad
Created subfolder: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\surprise
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000213_aug_1.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000213_aug_1.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000213_aug_2.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000213_aug_2.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000213_aug_4.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000213_aug_4.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000343_aug_0.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000343_aug_0.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000343_aug_1.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000343_aug_1.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000343_aug_2.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000343_aug_2.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000343_aug_4.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000343_aug_4.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000400_aug_0.jpg
Resized and saved: G:\Sem_8\Resized_Images_for_Swin_Transformer\train\anger\10_image0000400_aug_0.jpg
Processing image: G:\Sem_8\Resized_Faces\train\anger\10_image0000400_aug_1.jpg
```

SCREENSHOTS

Evaluating with Test dataset:

```
Preparing the test dataset...
Loaded 6389 test images.
Loading the trained model...
Model weights loaded from: G:\Sem_8\Models\swin_transformer_50646_images.pth
Evaluating the model on the test dataset...
Image 1/6389 processed in Batch 1.
Image 2/6389 processed in Batch 1.
Image 3/6389 processed in Batch 1.
Image 4/6389 processed in Batch 1.
Image 5/6389 processed in Batch 1.
Image 6/6389 processed in Batch 1.
Image 7/6389 processed in Batch 1.
Image 8/6389 processed in Batch 1.
Image 9/6389 processed in Batch 1.
```

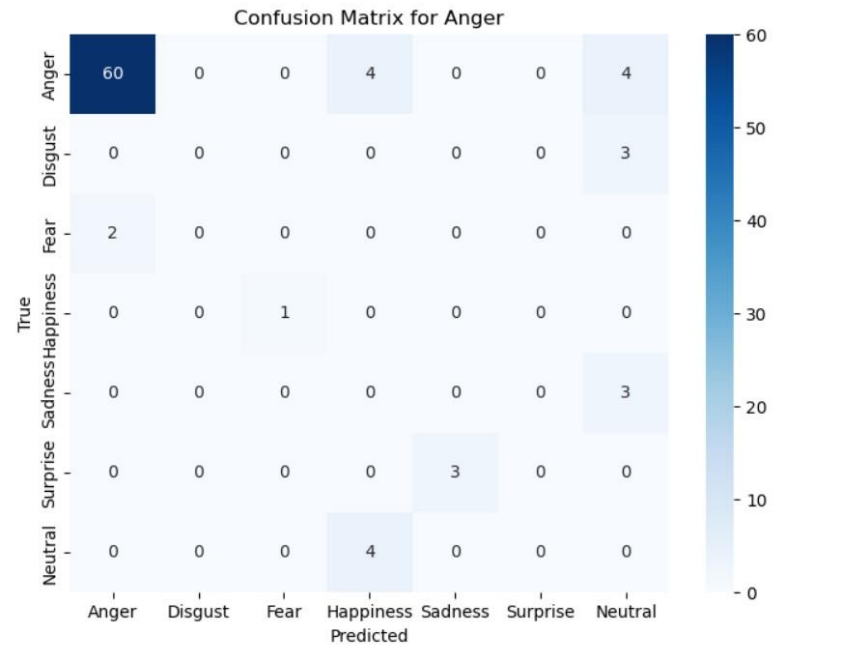
SCREENSHOTS

Emotion Classification using Swin Transformer:

```
Emotion detected : Anger
```

Confusion Matrix:

```
Confusion Matrix for detected emotion 'Anger':
Accuracy: 71.43%
```



SCREENSHOTS

Final Workflow for Emotion Classification:

```
1 face(s) detected! Drawing bounding box...
```


```
Emotion detected : Anger
```

KONGU ENGINEERING COLLEGE

ACCEPTANCE LETTER

ICEARS

3rd International Conference on Electronics and Renewable Systems (ICEARS 2025)

11-13, February 2025 | icears.com/conf2024/ | icears.con@gmail.com

LETTER OF ACCEPTANCE

To:

Dr. M. N. Kavitha, Saranya S S, Pragatheeswari E, N. Kaviya Priya, SR. Ranjith, N. Rahul Vanchivel
Kongu Engineering College

Decision: Acceptance with Major Revision

Herewith, the conference committee of the International Conference on Electronics and Renewable Systems ICEARS 2025 is pleased to inform you that the peer reviewed research paper **"Acceptance ID: ICEARS151"** entitled ***"Real-Time Facial Recognition System for Secure College Bus Transport using Deep Learning Techniques"*** has been accepted for oral presentation as well as it will be recommended in ICEARS Conference Proceedings. ICEARS will be held on 11-13, February 2025, in St.Mother Therasa Engineering College, Tuticorin, Tamil Nadu, India. ICEARS encourages only the active participation of highly qualified delegates to bring you various innovative research ideas.

We congratulate you on being successfully selected for the presentation of your research work in our esteemed conference.

ICEARS

--
Regards,
ICEARS 2025
Conference chair.

Estd : 1984

KONGU ENGINEERING COLLEGE

THANK YOU