FACIAL EMOTION RECOGNITION USING YOLO AND XCEPTION

NACHAMMAI DEVI POOJA S PRATHYUSH S RAMA PRABHA R
SUPERVISOR: Ms. ANGEL DEBORAH S, ASSISTANT PROFESSOR

Department of CSE, Sri Sivasubramaniya Nadar college of Engineering
Final Year Project, May 2022

Highlights of Proposed Model

- Detects the facial objects in the image.
- Hybrid model of using image with feature is proposed.
- Detects distinct emotions for each faces in the image.
- The proposed model has been deployed in Web App.

Challenges in recognizing emotions:

- Dataset consists of people with different facial characteristics, multiple faces in a frame.
- Occlusion, illumination conditions, Pose variations.

Functional Modules

- Face Detection
  - Train YOLO model with WIDERFACE
  - Annotate affectnet data set
  - Bounding Box Prediction
- Feature Extraction and Data Preparation
  - HOG sliding Window to extract gradient
  - SMOTE Oversampling
  - Image Augmentation using data generator
- Emotion Recognition using Xception

Dataset Description

- AffectNet is an in-the-wild dataset of facial expressions that contains over 1,000,000 facial images and 1250 emotion-related keywords.
- WIDER FACE dataset is a bench-mark dataset comprises of 32,203 images and labelled 393,703 faces with a high degree of variance.
- Since AffectNet does not have face annotations, the YoloV3 model is trained with WiderFace Dataset
- From the new augmented dataset 80% of total is taken for training 20% for validation and separate test-set for testing.

Bounding box and emotion prediction

- Based on the extracted features, different sizes of anchor boxes are defined per grid on the image.
- Non maximum suppression will remove the other bounding boxes with confidence levels below the threshold.
- A Deep learning framework captures the contextual information and appearance to discriminate among various emotions in the given frame.

Proposed Model for Facial Emotion Recognition

Face Detection

- YOLOv3 real-time object detection algorithm
- Since AffectNet does not have face annotations, the YoloV3 is trained with WiderFace Dataset.
- WIDER FACE dataset has shown that YOLOv3 algorithm has significantly increased accuracy while maintaining fast detection speed.

Oversampling and Image Augmentation

- Actual data set has imbalanced classes, which leads to a decrease in accuracy.To rectify this problem we balance the dataset
- The dataset is balanced using SMOTE (Synthetic Minority Oversampling Technique) Oversampler
- Each minority class sample is over-sampled by introducing synthetic examples along the line segments connecting any/all of the k minority class nearest neighbours.

Feature extraction

- We using HOG Sliding Window for feature extraction.
- HOG Sliding Window is the technique that counts occurrences of gradient orientation in localized portions of an image.
- It works to extract local features from a given image patch, which is a part of the image overlapped with the sliding window.

Emotion Recognition

- Xception Architecture is used for predicting emotion.
- Basic hypothesis: the mapping of cross-chain correlations and the separation of spatial correlations.
- Consists of 36 convergence layers that provide the foundation for network extraction.
- It consists of 14 modules with linear residual connections, excluding the first and last modules.

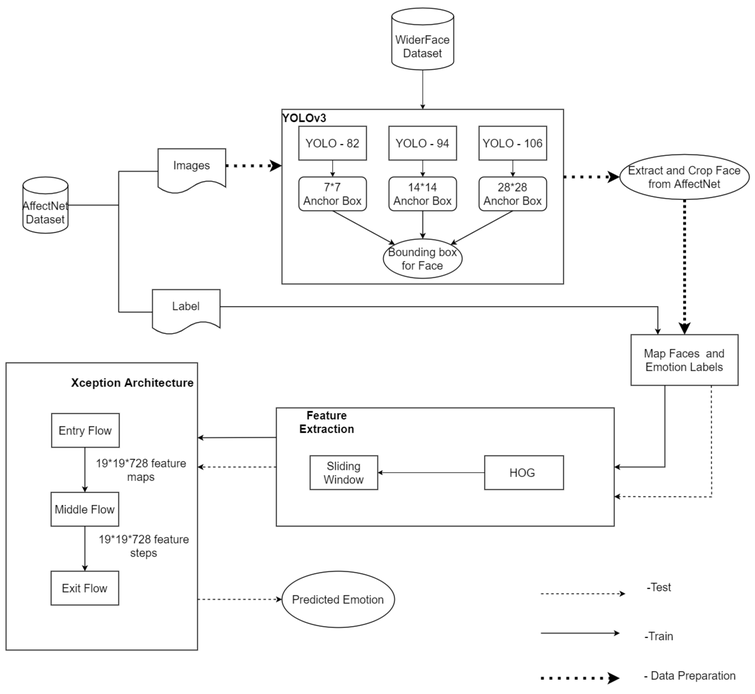

Figure 1. Architecture of the proposed system

Emotions in the Dataset

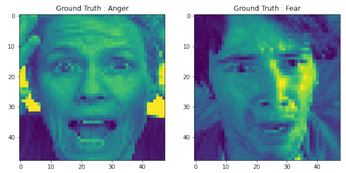

Figure 2. Representation of emotions in the dataset

Emotions recognized by the model

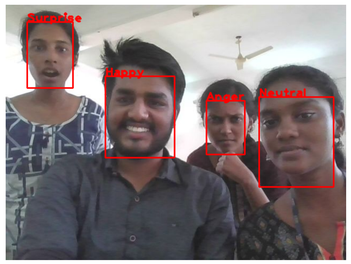

Figure 3. Different emotions predicted in real-time

Performance metrics of the Model

Model	Epochs	Precision	Recall	F1 score
Xception Image	100	42	46	45
Xception Hybrid	100	56	58	58
CNN Image	100	50	51	51
CNN Hybrid	100	53	53	53

Table 1. Performance Metrics for Facial Emotion Recognition

References

- Lin Zheng Chun. "YOLOv3: Face Detection in Complex Environments". In: *International Journal of Computational Intelligence System)* (Accessed 3 Dec. 2020). DOI: 10.2991/ijcis.d.200805.002
- M. Dahmane and J. Meunier. "Emotion recognition using dynamic grid-based HoG features". In: *Face and Gesture 2011 Santa Barbara, CA*, (2011)
- K. M. Rajesh and M. Naveenkumar. "A robust method for face recognition and face emotion detection system using support vector machines". In: 2016
- Syed Fathima, Ashwani Kumar, and Syed Raoof. "Real Time Emotion Detection of Humans Using Mini-Xception Algorithm". In: *IOP Conference Series: Materials Science and Engineering* 1042 (Jan. 2021), p. 012027. DOI: 10.1088/1757-899X/1042/1/012027