Recognizing Cultural Differences in Fear Social Signals

By: Arda Cifci & Tushrima Kelshikar & Zihan Yu

Background

Fear is a primal human emotion, with its expression deeply rooted in both physiological responses and social signals.

The emotion encompasses several subcategories including nervousness, surprise, and feeling threatened.

It varies widely among individuals and is even more diverse across cultures.

While we all feel fear, how we show it—through facial expressions, body language, and vocalizations—can be heavily influenced by cultural norms and practices.

Project Goals

This complexity presents a challenge for current emotion-recognition technology, which often fails to account for these nuances, leading to oversimplified interpretations.

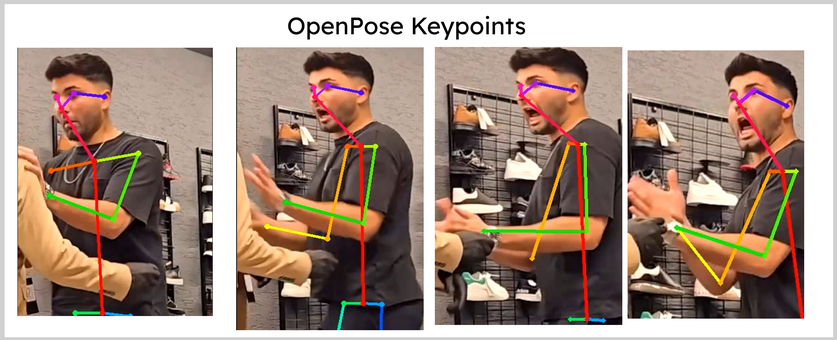
Our project aims to bridge this gap by analyzing and recognizing fear and its sub-categories within the rich tapestry of cultural context.

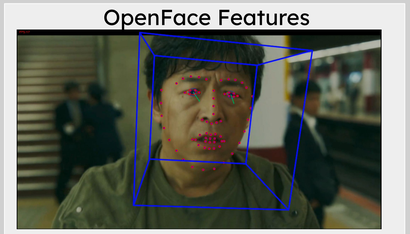
Dataset

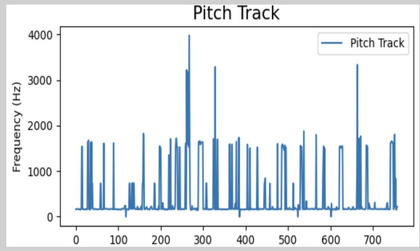
We collected around 200 dynamic short videos (1-10 seconds) from YouTube which included natural and staged expressions of fear.

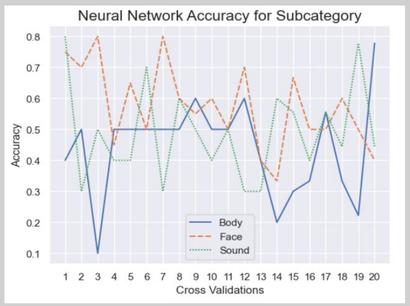
Our data focused on Canadian, Chinese, Indian, and Turkish cultures capturing the fear subcategories nervousness, surprise, and threatened.

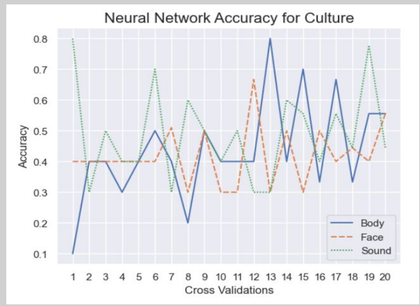
Annotations within this dataset meticulously categorize facial expressions, vocal attributes, and body language keypoints.

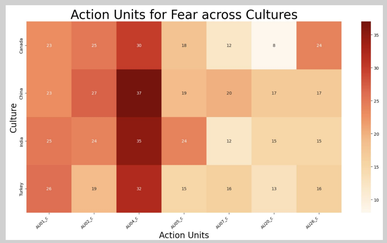


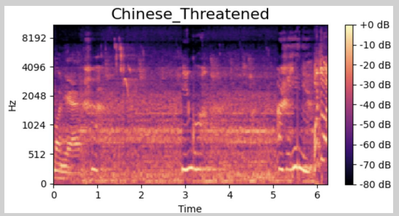


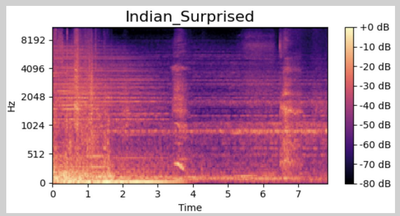












Methods

COLLECT : For data collection, we used OBS Studio to capture videos at 720P and 60FPS, saved as mp4 files.

SOUND : Audio was extracted using FFmpeg and processed to extract features such as pitch, frequency, and amplitude

FACE : OpenFace was used to extract facial feature data from our dataset which includes facial action units (AU's), gaze, and landmarks. We then processed and dimensionally reduced these extracted facial features to fit within our classification models.

BODY : OpenPose was used to extract body keypoint data from out dataset. This included 25 bodypart location landmarks. We then processed and dimensionally reduced these features using PCA to fit them with our classification models.

Using our extracted features, we trained several SVM and Neural network model's to recognize the cultural difference and fear subcategories of our dataset.

Results

The project reveals Chinese individuals express fear subcategories subtly, mainly through facial expressions. Neural networks achieved 60% accuracy in classifying Chinese individuals' fear sounds, compared to 55% for Turkey, 50% for Canada, and 25% for India.

Turkish express fear and its subcategories with more body movement and vocals, less facial expressions. Neural networks recognized Turkish facial fear with 52% accuracy and lowest in India at 46%.

Indians expressed fear subcategories moderately across all features. They were neither loud or quiet, had small movement or large movement, or showed extensive facial expressions. With neural networks, we saw India had the highest body movement classification accuracy of 60% and Canada with the lowest of 17%

Canadian individuals had a mix of everything. As Canada is a melting pot of culture and people, each Canadian video of the dataset had diverse features.