SENTIMENTAL ANALYSIS FOR PRODUCT FEEDBACK REVIEW

Dovari Kiran
Department of Information Technology
Vasavi College of Engineering
Hyderabad, India
kiran.kunnu55@gmail.com

Karrenolla Sunil
Department of Information Technology
Vasavi College of Engineering
Hyderabad, India
sunilkarrenolla872@gmail.com

Maranganti Sathya Devi
Department of Information Technology
Vasavi College of Engineering
Hyderabad, India
satyamaraganti@gmail.com

***Abstract*— In the realm of e-commerce, product reviews wield significant influence over consumer purchasing behaviors. Online platforms brim with user-generated content, offering diverse perspectives on products and services. Sentiment analysis, a vital tool in this landscape, systematically deciphers the emotional undercurrents within text, empowering businesses with actionable insights. Videos and audio recordings introduce richer dimensions to these reviews, capturing nuanced expressions and vocal nuances. By integrating both modalities into sentiment analysis, a deeper comprehension of reviewers' emotional states emerges, providing more nuanced insights into product perception and satisfaction. Our innovative approach merges the Haar cascade algorithm for facial detection and sentiment analysis in videos with the BERT model for audio sentiment analysis, synergizing visual and auditory cues. This integration enhances accuracy and robustness, facilitating a more precise assessment of product sentiment. Through the fusion of outputs from these modalities, we endeavor to pinpoint the most salient emotional cues expressed by reviewers, offering businesses a comprehensive understanding of consumer sentiment and aiding in informed decision-making processes.**

***Index Terms*—, Video and audio integration, Sentiment Analysis, Multimodal Fusion, Product Reviews, Haar Cascade Algorithm, BERT Model, Emotional Analysis. Speech recognition.**

I. INTRODUCTION

This paper presents an innovative approach to sentiment analysis in product reviews by integrating video and audio modalities, thereby extracting nuanced emotional insights. Leveraging the Haar cascade algorithm for video analysis and the BERT model for audio processing, our method offers a holistic understanding of reviewer sentiment. By scrutinizing facial expressions, gestures, speech patterns, and vocal intonations, we capture the intricate emotional context of product reviews, providing deeper insights into consumer perceptions and satisfaction. Additionally, our adaptable solution caters to diverse platform capabilities and user preferences, making it suitable for both offline and online product review systems.

Through empirical evaluation and comparative analysis, we demonstrate the effectiveness and robustness of our approach across various product categories and review platforms. This research significantly contributes to the advancement of sentiment analysis techniques and multimodal fusion,

highlighting the potential for enhanced sentiment understanding in product review systems. Furthermore, our findings offer practical implications for businesses seeking to utilize consumer feedback effectively and make informed decisions. By integrating the video-based Haar cascade algorithm and the audio-based BERT model outputs, our method aims to identify prominent emotional cues expressed by reviewers, thereby providing a comprehensive assessment of product sentiment.

Moreover, our proposed approach addresses the dynamic needs of both offline and online product review systems. It can efficiently analyze pre-recorded video and audio reviews offline, offering valuable insights into product sentiment without real-time processing requirements. Conversely, for online platforms hosting live video or audio reviews, our method seamlessly adapts to real-time processing, enabling immediate feedback and analysis of consumer sentiment as reviews are generated. This versatility makes our approach applicable across various review platforms, enhancing the depth and accuracy of sentiment analysis in product reviews.

II. SURVEY

Gina Khayatun Nufus, Mustafid Mustafid, and Rahmat
"Sentiment Analysis for Video on Demand Application User Satisfaction with Long Short Term Memory Model"
[1] This study proposes a methodology utilizing Word2vec and Long Short-Term Memory (LSTM) models for aspect-level sentiment classification in video-on-demand application reviews. The integration of Word2vec embeddings with LSTM architecture enhances sequential data modeling, facilitating accurate sentiment classification.

Gaurav Meena, Krishna Kumar Mohbey, Sunil Kumar
"Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach"
,[2] This research focuses on image-based sentiment analysis. Through transfer learning and pre-trained deep CNN models, the study aims to detect and classify emotions in images, achieving superior results compared to traditional machine learning approaches.

M. A. H. Akhand, Shuvendu Roy, Nazmul Siddique and Tetsuya Shimamura
"Facial Emotion Recognition Using Transfer Learning in the Deep CNN"
[3] This study proposes a very deep Convolutional Neural Network (DCNN) model leveraging transfer learning. By fine-tuning pre-trained DCNNs and handling diverse facial angles, the proposed approach achieves remarkable accuracy in facial emotion recognition.

Nur Hasifah A Razak, Muhammad Firdaus Mustapha.
"Amazon Product Sentiment Analysis using RapidMiner"
[4] This paper presents a systematic methodology for sentiment analysis of Amazon product reviews across different categories. Utilizing RapidMiner, the study involves data acquisition, preprocessing, and model validation to analyze sentiment trends and their impact on product attitudes.

Dorca Manuel-Ilie, Pitic Antoniu Gabriel, Crețulescu Radu George
"Sentiment Analysis Using Bert Model"
[5] Employing the BERT model, this research proposes a comprehensive approach to sentiment analysis that integrates deep learning concepts. By annotating topics and emotions, customizing neural network architectures, and systematic model evaluation, the methodology aims to optimize sentiment analysis performance across diverse datasets.

Amira Samy Talaat.
"Sentiment analysis classification system using hybrid BERT models".
[6] Introducing a sentiment analysis framework with hybrid BERT models, this study combines BERT with BILSTM and BIGRU algorithms. Through preprocessing, model comparison, and evaluation on multiple datasets, the research demonstrates the superiority of hybrid models over classical machine learning methods.

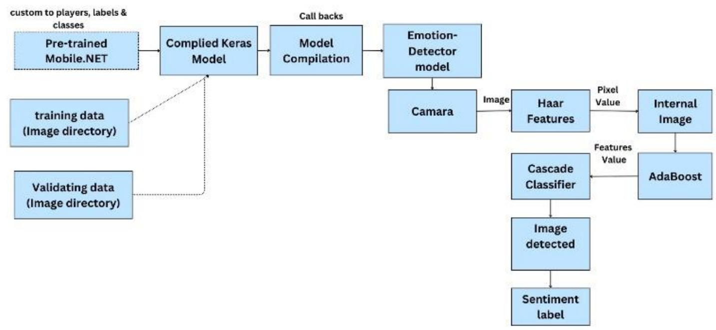

Fig. 1. Video Sentiment Analysis block diagram

III. PROPOSED METHODOLOGY

The proposed methodology for sentiment analysis of product reviews using audio and video data involves specific key steps. First, the audio data is preprocessed and fed into the BERT model, a state-of-the-art deep learning architecture for Natural Language Processing, to extract sentiment features from the audio content. Meanwhile, the video data undergoes preprocessing and facial expression recognition using the Haar Cascade algorithm to detect emotional cues from facial expressions. These extracted features from both modalities are then fused using a multimodal fusion technique, such as late fusion or early fusion, to combine the complementary information from audio and video sources.

A. Harr Cascade

This document outlines a methodology for preprocessing video data for face detection using the Haar cascade algorithm.

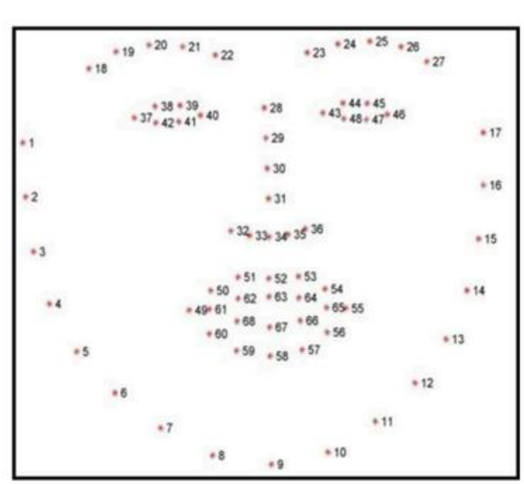

Fig. 2. Facial Landmarks Points

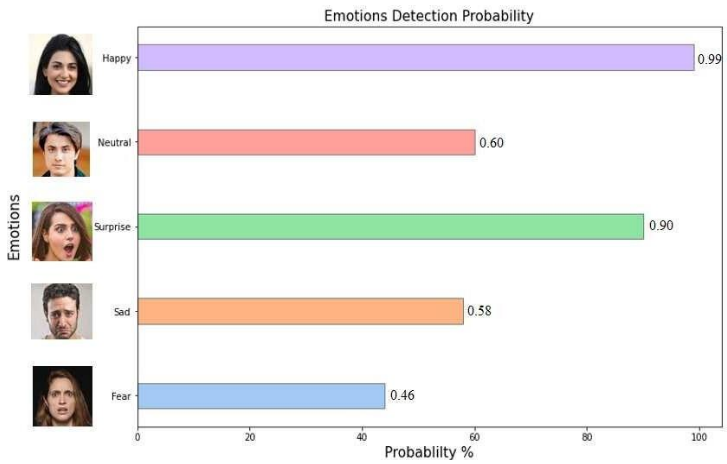

Fig. 3. Emotions Detection

The Haar cascade is a machine learning approach well-suited for real-time face detection due to its computational efficiency.

Methodology–

Video Acquisition:
Utilize a video capture library like OpenCV's VideoCapture class to read the target video file.

Frame Extraction:
Iterate through the video using a loop to extract individual frames. Each frame represents a single image captured at a specific point in time.

Grayscale Conversion:
The Haar cascade classifier typi- cally operates on grayscale images. Convert each extracted frame from RGB (color) to grayscale format for optimal performance. This can be achieved using OpenCV's cvtColor function.

Resizing:
Consider resizing frames to a smaller resolution to reduce processing time. This is a trade- off between speed and accuracy, where smaller sizes are fasterbut might miss smaller faces. Use OpenCV's resize function for resizing.

Face Detection:
Load a pre-trained Haar cascade classifier for frontal face detection. Popular choices include OpenCV's built-in-haarcascade-frontalface-default.xml or haarcascade-frontalface-alt2.xml. Employ the classifier's detectMultiScale function on each preprocessed frame. This function identifies regions within the frame that correspond to likely face locations. It returns a list of bounding boxes (rectangles) enclosing the detected faces.

Post-processing:
Refine the detected faces based on your application's needs. This could involve: Applying additional filtering to remove false positives. Extracting facial features (e.g., eyes, mouth) using specialized classifiers within the detected regions.

ResultGeneration:
Depending on our objective, we can choose the following outputs. Generate a new video with bounding boxes drawn around the detected faces in each frame. (OpenCV's rectangle function can be used for drawing boxes). Extract and store the bounding box coordinates and other relevant data from each detected face for further analysis. Haar Cascade uses the facial landmarks to predict the labeled emotions by the following modules:
1. Haar Features Calculation: Calculates features in adjacent regions of an image, mainly involving pixel intensity sums and differences.

2. Integral Image Creation: Reduces calculation by creating integral images, focusing only on relevant features. Adaboost helps in selecting useful features.

3. Adaboost Training: Combines weak classifiers to form a strong classifier. Weak learners are trained to distinguish objects from non-objects based on computed Haar features.

4. Cascading Classifiers Implementation: Uses a cascade of classifiers to efficiently reject negative samples and quickly identify positive ones. Boosting ensures accurate classification by combining predictions from multiple weak learners.

B. Audio Sentiment

In our approach to sentiment analysis of audio data using BERT, we incorporate a real-time audio-to-text conversion system powered by the Google API. This system is instrumental in swiftly and accurately converting spoken content into text, enabling subsequent sentiment analysis with the BERT model. By efficiently handling incoming audio streams, the Google API module promptly converts them into text while maintaining content fidelity. This conversion process is essential as it effectively bridges the gap between the unstructured nature of audio data and the text-based analysis required by the BERT model. Leveraging the Google API for real-time transcription ensures that the textual representation of the audio remains current and closely aligned with the original spoken content throughout the analysis process.

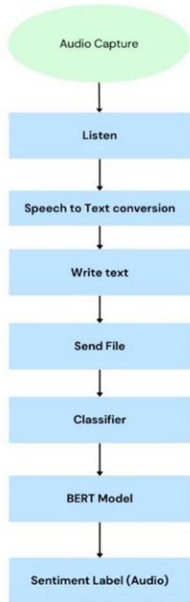

Fig. 4. Audio Sentiment Analysis Flow Datagram

Once the audio data is converted into text, we integrate a pre-trained BERT model into our system for sentiment analysis. BERT is a state-of-the-art language representation model that excels in understanding contextual nuances and relationships within text data. By incorporating BERT, we can capture the sentiment expressed in the transcribed audio text accurately. The BERT model's bidirectional nature allows it to consider the entire context of the input text, resulting in more nuanced and contextually informed sentiment analysis outcomes. This integration of advanced NLP techniques with real-time audio transcription enhances the capability of our system to provide accurate and timely sentiment insights from spoken content, enabling applications across various domains such as customer service, market research, and voice-driven user interfaces.

C. Integrated Sentimental analysis

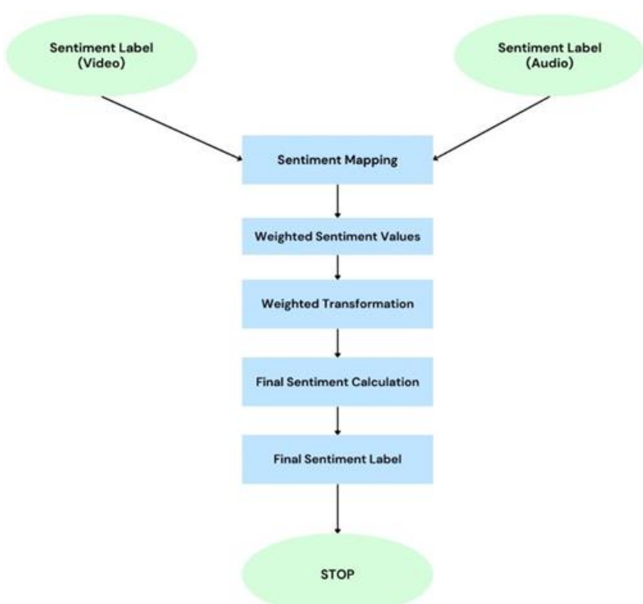

Fig. 5. Integrated Sentimental Analysis

The block diagram for this module involves a sentiment fusion approach where sentiment values from audio and video sources are combined to produce a final sentiment classifica- tion.

1) Sentiment Mapping: Convert sentiment labels ('Extremely Negative', 'Negative', 'Neutral', 'Positive',' Extremely Positive') to numerical values (0 to 4) using a dic- tionary (sentiment-mapping), facilitating calculation.

2) Weighted Sentiment Values: Calculate sentiment values (audio-value and video-value) based on sentiment labels (label-audio-sentiment and label-video) obtained from audio and video sentiment analyses, respectively. Multiply each sentiment value by predefined weights (audio-weight and video-weight) to determine relative importance.

3) Final Sentiment Calculation: Combine weighted sentiment values (audio-value and video-value) to generate a composite sentiment value (final-temp). Use the max() function to constrain the final sentiment value (final-temp) within the defined sentiment range (0 to 4). Round the composite sentiment value (final-temp) to the nearest integer using math.ceil() to obtain the final sentiment score (final).

4) Sentiment Label Conversion: Convert the final sentiment score (final) back into a sentiment label using the senti-final() function. This function maps the numerical sentiment score to its corresponding sentiment label ('Extremely Negative', 'Negative', 'Neutral', 'Positive', 'Extremely Positive').

This block diagram outlines a sentiment fusion process where sentiment values from audio and video sources are combined using weighted averages to produce a comprehensive final sentiment classification. Leveraging inputs from multiple modalities (audio and video) and considering their relative importance enhances the robustness of sentiment analysis.

IV. RESULTS AND DISCUSSIONS

Fig. 6. Video sentiment output of subject_ 1 emotion

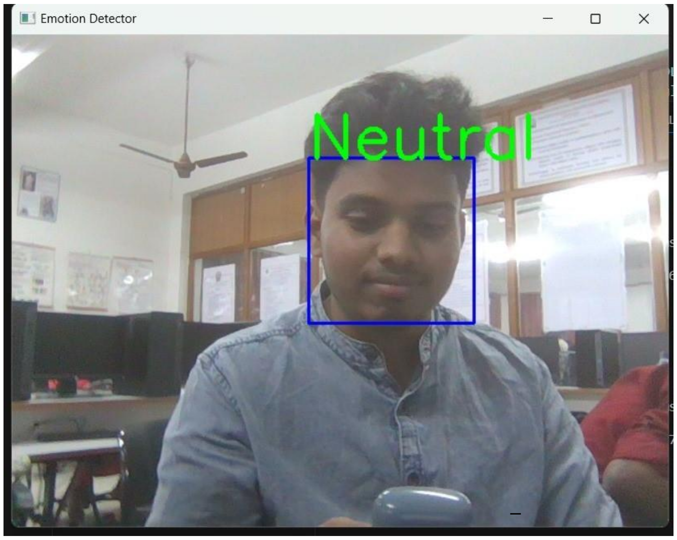

Fig. 7. Video sentiment output of subject_ 2 emotion

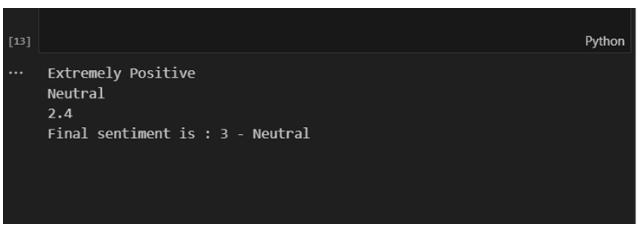

Fig. 8. Integrated Final sentiment output

In this section, we unveil the outcomes of our study, focusing on the criteria set for the Product Sentiment Feedback System. These results shed light on the performance and efficacy of the system, revealing insights into its usability, scalability, and accuracy. Through a comprehensive analysis, we aim to evaluate the system's effectiveness in delivering sentimental information corelating to the product and ensuring a seamless user experience. The findings presented here are pivotal in understanding the feedback system's suitability for addressing diverse user needs and guiding future enhancements to optimize its performance.

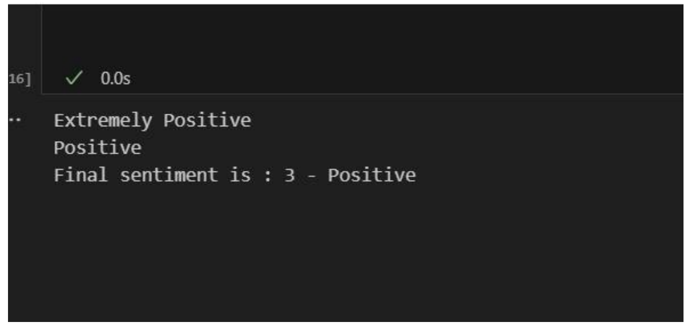

Fig. 9. Integrated Final sentiment output

Emotions	Precision		Recall		Accuracy (Individual)		Accuracy (Combined)
	Video (Haar)	Audio (Bert)	Video (Haar)	Audio (Bert)	Video (Haar)	Audio (Bert)	Video + Audio
Extremely negative	0.58	0.91	0.56	0.97	0.96	0.97	0.96
Negative	0.50	6.95	0.62	0.90	0.96	0.99	0.97
Neutral	0.60	0.88	0.66	0.85	0.97	0.99	0.98
Positive	0.87	0.98	0.81	0.96	0.97	0.99	0.98
Extremely positive	0.53	0.91	0.80	0.97	0.96	0.97	0.96

Table. 1. Analysis

Comparison of Accuracies:

Individual Modality Accuracies:

Haar Cascade (Video): Moderate to Good Accuracy.

BERT Model (Audio): High Accuracy.

Combined Modality Accuracy: The accuracy of combined sentimental analysis can potentially exceed that of individual modalities: Video provides visual cues (facial expressions) while audio provides verbal cues (spoken content). Combining these modalities can capture a more holistic representation of the user's emotional state.

Evaluation Metrics:

We evaluated the combined sentimental analysis using metrics like accuracy, precision, recall. Upon comparision of these metrics against the individual modality performances to assess the improvement achieved by combining modalities we learn that the Sentimental analysis is greatly improved and we get a precise sentiment relating to product.

Combining traditional video-based sentimental analysis (Haar cascade) with audio-based sentimental analysis using the BERT model can lead to enhanced accuracy and robustness in sentiment analysis tasks. While both individual modalities have their strengths and limitations, their fusion leverages synergies between visual and auditory cues, thereby providing a more comprehensive understanding of emotional expressions.

V. CONCLUSION & FUTURE SCOPE

The integration of traditional video-based sentiment analysisusing Haar cascade for facial emotion detection with audio- based sentiment analysis using the BERT model represents a promising approach to comprehensive emotion recognition from multimedia inputs.

1) Advantages of Integration:
   - Combining Visual and Auditory Cues: Integration of facial expressions (visual) and spoken content (audi- tory) provides a more comprehensive understanding of emotions as they often complement each other.
   - Leveraging Strengths: Visual cues can capture subtle facial expressions, while auditory cues can capture nuances in tone and speech, compensating for lim- itations in each modality.
   - Enhanced Accuracy: Integration of multiple modal- ities enhances the accuracy and robustness of emotion recognition systems by leveraging diverse sources of emotional information.
2) Potential Impact:
   - Improved Accuracy and Reliability: The integrated approach is expected to lead to improved accuracy and reliability in emotion detection tasks compared to uni-modal approaches.
   - Applicability in Various Domains: The approach has broad applicability in domains such as customer feedback analysis, mental health monitoring, and human-computer interaction systems, where under- standing emotions is crucial.
3) Future Scope: a. Enhanced Fusion Techniques: Research and development of advanced fusion techniques, such as deep learning-based multimodal fusion, can improve the integration of video and audio modalities. b. Model Optimization: Fine-tuning and optimizing the Haar cas- cade model for better detection of subtle facial expres- sions, and experimenting with different architectures and configurations of the BERT model for audio sentiment analysis. c. Multimodal Dataset Development: Creation and curation of multimodal datasets that include syn- chronized video and audio data to facilitate training and evaluation of integrated emotion recognition systems. d. Cross-Domain Applications: Extending the system's ca- pabilities to analyze emotions across different domains, such as analyzing emotions in music or gestures, to broaden its applicability. e. Language Considerations: Expanding the BERT model to handle multi-lingual datasets, enabling emotion analysis in diverse linguistic contexts.

REFERENCES

[1] Gina Khayatun Nufus, Mustafid Mustafid, and Rahmat "Sentiment Analysis for Video on Demand Application User Satisfaction with Long Short Term Memory Model"

[2] Gaurav Meena, Krishna Kumar Mohbey, Sunil Kumar "Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach"

[3] M.A.H. Akhand, Shuvendu Roy, Nazmul Siddique and Tetsuya Shimamura
"Facial Emotion Recognition Using Transfer Learning in the Deep CNN"
https://www.mdpi.com/2079-9292/10/9/1036

[4] Nur Hasifah A Razak, Muhammad Firdaus Mustapha.
"Amazon Product Sentiment Analysis using RapidMiner"

[5] Dorca Manuel-Ilie, Pitic Antoniu Gabriel, Crețulescu Radu George
"Sentiment Analysis Using Bert Model"
https://www.researchgate.net/publication/376670839_Sentiment_Analysis_Using_Bert_Model

[6] Amira Samy Talaat.
"Sentiment analysis classification system using hybrid BERT models".
https://journalofbigdata.springeropen.com/articles/10.1186/s40537-023-00781-w

[7] S. S. A. B. U. Rahul Ramachandrana,
" Exploring the relationship between emotionality and product star ratings in online reviews,"

[8] Mika V. Mantyla, Daniel Graziotin and Miikka Kuutila,
"The Evolu-tion of Sentiment Analysis-A Review of Research Topics".

[9] Sentiment Analysis, Available at:
https://insightsatlas.com/sentiment- analysis/

[10] Sentiment Analysis Explained, Available at: https://www.lexalytics.com/technology/sentiment-analysis

[11] https://github.com/ShivamGaurUQ/Sentiment-Analysis-of-Amazon- product-reviews