Hybrid Deep Learning Framework for Personality Prediction in E-Recruitment

Om Sonawale, Prof. Pradnya V. Kulkarni, Ayaan Sayyad, Omkar Dadpe, Vivek Kumar

Om Sonawale, Computer Engineering, SAE
Prof. Pradnya V. Kulkarni, Computer Engineering, SAE
Ayaan Sayyad, Computer Engineering, SAE
Omkar Dadpe, Computer Engineering, SAE
Vivek Kumar, Computer Engineering, SAE

---***---

Abstract - The Personality prediction is a vital task in the domain of psychology, human-computer interaction, and user behavior analysis, with applications ranging from tailored advertisements to mental health assessments. Traditional methods rely heavily on self-report questionnaires or psychological assessments, which can be time-consuming and subjective. To overcome these limitations, researchers are exploring automated and objective methods using deep learning techniques, particularly Convolutional Neural Networks (CNNs) and Natural Language Processing (NLP) algorithms, which excel at capturing complex patterns in multimodal data. In this work, we propose a hybrid framework that combines CNN-based feature extraction and NLP algorithms for predicting personality traits using data such as facial expressions, speech patterns, and text analysis. The CNN model is leveraged due to its robust feature extraction capabilities, enabling it to learn intricate patterns directly from raw data inputs, which correlate with the Big Five personality traits. Additionally, we integrate a Support Vector Classifier (SVC) to classify personality traits based on the extracted features, offering improved prediction accuracy across diverse data sources.

Keywords—Personality Prediction, Big Five Personality Traits, CNN, NLP Algorithm, SVC Classifier

1. INTRODUCTION

Personality prediction is an emerging field in artificial intelligence and behavioral analysis with applications ranging from human resource management to personalized recommendations. Traditional personality assessments rely on self-report questionnaires, which can be subjective, time-consuming, and prone to bias. In contrast, modern AI-driven solutions offer objective and scalable methods to infer personality traits from observable data such as images, video, speech, and text.

This paper presents the technical implementation of a multimodal system that predicts personality traits using a hybrid framework based on deep learning. We combine Convolutional Neural Networks (CNNs) for visual data analysis and Natural Language Processing (NLP) techniques for text-based analysis. In addition to static image classification, we incorporate video analysis by extracting frames at regular intervals and processing them through the CNN model. Textual data is preprocessed and vectorized to extract meaningful linguistic patterns. Each modality contributes features that are fused and passed to a Support Vector Machine (SVM) classifier, which performs the final personality classification based on the Big Five personality model.

This system is designed to assist in e-recruitment platforms, where personality analysis can play a crucial role in identifying suitable candidates for specific roles. The implementation demonstrates how deep learning and machine learning techniques can be used in harmony to produce a robust and automated prediction pipeline.

2. LITERATURE SURVEY

SR	AUTHOR	TITLE	DESCRIPTION
1.	Marco A. Moreno-Armendariz, Carlos Alberto Duchanoy Martine	Estimation of Personality Traits From Portrait	Explores predicting the Big Five traits from text using linguistic features and NLP. Correlations were found, especially for Extraversion and Openness. Accuracy was limited by manual features.

2.	Sonali Pakhmode, Shaikh Ayan	Implementing Personality Prediction Using Machine Learning	Uses vocal features and SVM to predict personality traits. Moderate accuracy was achieved, with Extraversion and Neuroticism being most predictable.
3.	**I** Maliki, M A Sidik	Personality Prediction System Based on Signatures Using Machine Learning	Combines facial and vocal features for prediction using Decision Trees and Random Forests. Integration improved accuracy, but manual extraction limited scalability.
4.	Jia Xu, Weijian Tian	Prediction of the Big Five Personality Traits Using Static Facial Images of College Students with Different Academic Background	Uses CNNs to predict traits from facial expressions in videos. Outperforms traditional models, especially for Extraversion and Openness, by automatically learning features.

2.1. METHODOLOGIES

Gather a diverse set of images categorized as savory or unsavory. The dataset should include a wide range of food types and presentation styles to ensure robust model training. Additionally, collect corresponding text data, such as food reviews or descriptions, to complement the image data.

Clean and preprocess images to standardize formats, sizes, and quality. Augment the dataset to enhance model performance and generalization. Similarly, preprocess text data using (NLP) algorithm, performing tokenization, stop word removal, and lemmatization.

Develop a Convolutional Neural Network (CNN) model to predict personality traits based on the analysis of savory and unsavory food images. Additionally, integrate an NLP algorithm, to analyze the text data, extracting features that can correlate with personality traits based on linguistic patterns.

• Accuracy: To evaluate overall correctness of predictions.

• Precision, Recall, and F1 Score: To evaluate the model's performance on each personality trait, incorporating both visual and text-based predictions.

2.2. SYSTEM ARCHITECTURE

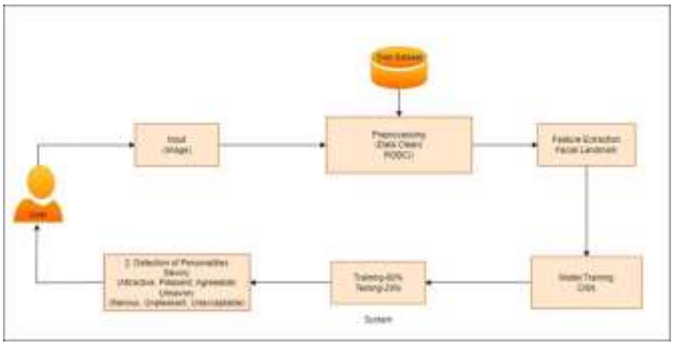

Fig-1 : figure

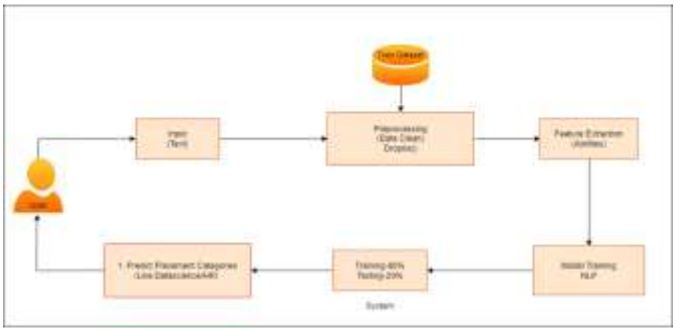

Fig-2 : figure

- CNNs are exceptionally well-suited for image and text data analysis tasks due to their ability to automatically learn hierarchical features from raw pixel data.
- CNNs can learn to extract meaningful features from images such as edges, textures, and shapes, which are crucial for recognizing patterns. Additionally, integrating NLP algorithms enables the analysis of accompanying textual data, such as incident reports or social media posts, enhancing the understanding of the context surrounding the visual content.

- CNNs are capable of achieving translation invariance, meaning they can recognize patterns regardless of their position within an image. Coupled with NLP techniques that analyze language structure and sentiment, this enables a comprehensive approach to understanding the relationships between visual elements and associated textual information, providing deeper insights into accident-related events

2.3. SYSTEM OVERVIEW

The proposed system is a hybrid framework that combines deep learning and traditional machine learning components to predict personality traits from three different modalities: images, video, and text. The system is divided into the following modules:

- **Image Module**: Processes savory and unsavory food-related images using a CNN model. This module extracts features correlated with users' preferences and associated personality labels.

- **Video Module**: Captures dynamic facial expressions by extracting frames from user videos. These frames are then passed through the same CNN pipeline used in the image module.

- **Text Module**: Analyzes user-provided textual data, such as reviews or comments, by vectorizing text using techniques like TF-IDF and feeding the features into a machine learning model.

- **Classifier Fusion**: The outputs from the CNN and text-based models are combined using a Support Vector Classifier (SVC) for final personality classification

2.4. IMPLEMENTATION

2.4.1. DATASET DESCRIPTION

1. Image Dataset The image dataset consists of **savory and unsavory food images**, categorized based on user preferences and associated with labeled personality traits. The images are divided into training and testing sets across multiple classes, each representing a dominant Big Five personality trait. The dataset was manually curated and organized into folders corresponding to each class.

2. Video Dataset To incorporate dynamic facial analysis, a custom video dataset was collected. Short video clips of subjects were processed using **OpenCV** to extract frames at regular intervals (e.g., every 0.5 seconds). These frames were treated similarly to static images and passed through the CNN pipeline.

3. Text Dataset

The text data was obtained from personality-labeled user reviews, written responses, and personality test descriptions. Each sample was associated with one of the Big Five personality categories. This dataset was preprocessed and transformed into feature vectors using **TF-IDF** techniques.

2.4.2. PREPROCESSING TECHNIQUES

Image and Video Preprocessing

- Images and extracted video frames were resized to **64x64 pixels**.
- Pixel values were normalized to the range [0, 1].
- Data augmentation techniques such as **rotation**, **zoom**, and **horizontal flipping** were used to improve generalization.
- Frame extraction from video clips was performed using OpenCV.

Text Preprocessing

- All text samples were converted to lowercase.
- **Tokenization** was applied to split text into words.
- **Stop words** and punctuation were removed.
- **Lemmatization** was used to reduce words to their base forms.
- The cleaned text was then transformed into numerical feature vectors using **TF-IDF vectorization**.

Label Encoding

For both image and text data, personality traits were encoded using **one-hot encoding** or **categorical labels** depending on the classifier.

2.4.3. MODEL ARCHITECTURE

1.CNN Model for Image and Video Frame Classification

The core model used for visual personality prediction is a Convolutional Neural Network (CNN) built using **Keras**. The architecture includes:

- **Input Layer**: 64x64x3 input shape
- **3 Convolutional Layers** with 32 and 64 filters, kernel size 1x1, and ReLU activation
- **Max Pooling Layers** to reduce spatial dimensions
- **Flattening Layer**
- **Fully Connected Dense Layer** with 256 units and dropout
- **Output Layer** with 6 softmax neurons (one for each personality trait)

The CNN was trained for **600 epochs** using **Stochastic Gradient Descent (SGD)** optimizer and categorical cross-entropy loss.

2.NLP Model for Text Classification

Textual data is analyzed using a classical NLP pipeline that includes preprocessing and vectorization, followed by classification using machine learning algorithms.

Pipeline Stages:

- **Text Cleaning**: Lowercasing, stop-word removal, lemmatization.
- **Vectorization**: TF-IDF transformation to represent text numerically.
- **Classification**: Support Vector Machine (SVM) classifier trained on TF-IDF vectors.

SVM Configuration:

- **Kernel**: Radial Basis Function (RBF)
- **C-parameter**: 1.0
- **Gamma**: ‘scale’

2.4.4. IMPLEMENTATION IMAGES

Fig 2 : Main Screen

Fig 3: Login/ Registr

Fig 4 : Functionality page

Fig 5: Image Analysis

Fig 6 : Text Analysis

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

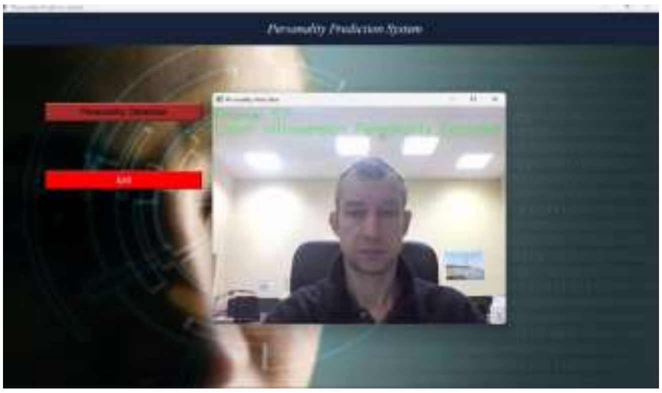

Fig 7 : Video Analysis

2.4.5. EVALUATION AND RESULTS

To measure the performance of the models, the following metrics were used:

- **Accuracy**: Overall correctness of predictions
- The CNN model shows an accuracy of 92%
- The NLP Model shows an accuracy of
- Where as the **Model's overall Training Accuracy** is **88.64 %**
- **Model's Overall Testing Accuracy is 86.89%.**

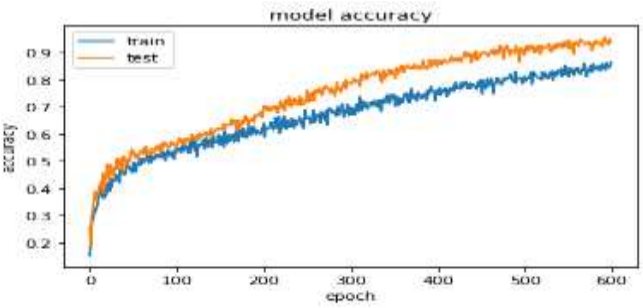

Fig-4 : CNN MODEL ACCURACY

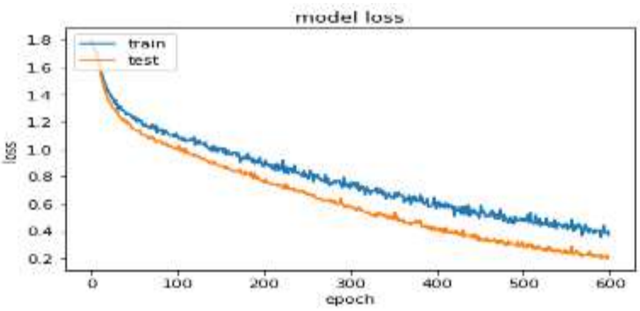

Fig-3 : CNN MODEL LOSS

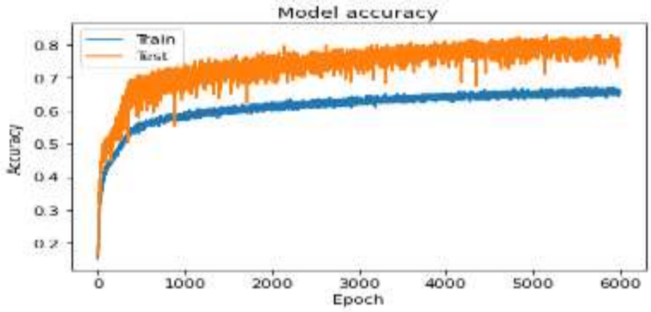

Fig 5 : OVERALL MODEL ACCURACY

2.5. LIMITATIONS

- The proposed system for personality prediction using NLP & Convolutional Neural Networks (CNNs) based on savory and unsavory images and text data has several limitations and challenges. Understanding these limitations is crucial for improving the system and managing expectations regarding its performance.

- The accuracy and generalization of the model are highly dependent on the size and diversity of the dataset. A small or non-representative dataset may lead to overfitting and poor performance on unseen data.

- The dataset might be biased towards certain cuisines or presentation styles, which can skew the predictions. Similarly, if the accompanying textual data such as reviews or comment lacks diversity or is biased, it can negatively affect the insights gained from the NLP analysis.

- Variations in image quality, such as resolution, lighting conditions, and background clutter, can impact the model's ability to accurately analyze and classify images. In the context of NLP, inconsistencies in the quality of textual data, such as spelling errors or informal language, may further complicate the extraction of meaningful features and context.

- Furthermore, the integration of image and text data may introduce additional complexity, as aligning insights from both modalities can be challenging. A lack of synergy between visual and textual features could lead to misinterpretations in the personality prediction process.

3. CONCLUSIONS

This paper presented the detailed implementation of a hybrid deep learning system for personality prediction using image, text, and video modalities. By leveraging CNNs for visual processing and NLP techniques for textual analysis, we created a modular and scalable architecture suitable for e-recruitment and psychological profiling. The inclusion of video frame analysis enhanced prediction by capturing dynamic visual cues. Future work includes real-time deployment, audio-based integration, and model explainability improvements.

ACKNOWLEDGEMENT

We would like to thank our project guide, Prof. P.V. Kulkarni, for her continuous guidance and support throughout the development of the software. We also acknowledge the contributions of the Computer Engineering Department at Sinhgad Academy of Engineering.

REFERENCES

[1]. M. Cao and Z. Wan, ''Psychological counseling and character analysis algorithm based on image emotion,'' IEEE Access, early access, Aug. 28, 2023, doi: 10.1109/ACCESS.2020.3020236.

[2]. M. Jayaratne and B. Jayatilleke, ''Predicting personality using answers to open-ended interview questions,'' IEEE Access, vol. 8, pp. 115345–115355, 2023.

[3]. H. A. Al Abdullatif and J. A. Velazquez-Iturbide, ''Who will continue using MOOCs in the future? Personality traits perspective,'' IEEE Access, vol. 8, pp. 52841–52851, 2023.

[4]. R. B. Tareaf, S. A. Alhosseini, P. Berger, P. Hennig, and C. Meinel, ''Towards automatic personality prediction using facebook likes metadata,'' in Proc. IEEE 14th Int. Conf. Intell. Syst. Knowl. Eng. (ISKE), Nov. 2023, pp. 714–719.

[5]. K. Kircaburun, S. Alhabash, Ş. B. Tosuntaş, and M. D. Griffiths, ''Uses and gratifications of problematic social media use among University students: A simultaneous examination of the big five of personality traits, social media platforms, and social media use motives,'' Int. J. Mental Health Addiction, vol. 18, no. 3, pp. 525–547, Jun. 2022.

[6]. R. M. Warner and D. B. Sugarman, ''Attributions of personality based on physical appearance, speech, and handwriting.,'' J. Personality Social Psychol., vol. 50, no. 4, pp. 792–799, 2023.

[7]. G. Farnadi, G. Sitaraman, S. Sushmita, F. Celli, M. Kosinski, D. Stillwell, S. Davalos, M.-F. Moens, and M. De Cock, ''Computational personality recognition in social media,'' User Model. User-Adapted Interact., vol. 26, nos. 2–3, pp. 109–142, 2022.

[8]. M. D. Back, J. M. Stopfer, S. Vazire, S. Gaddis, S. C.
Schmukle, B. Egloff, and S. D. Gosling, ''Facebook profiles reflect actual personality, not self idealization,'' Psychol. Sci., vol. 21, no. 3, pp. 372–374, Mar. 2023.

[9]. M. Skowron, M. Tkalčič, B. Ferwerda, and M. Schedl, ''Fusing social media cues: Personality prediction from Twitter and instagram,'' in Proc. 25th Int. Conf. Companion World Wide Web WWW Companion, 2023, pp. 107–108.

[10]. A. Souri, S. Hosseinpour, and A. M. Rahmani, ''Personality classification based on profiles of social networks' users and the five-factor model of personality,'' Hum.-Centric Comput. Inf. Sci., vol. 8, no. 1, p. 24, Dec. 2023.