

Narrative-Based Emotional Intelligence: Inferring Psychological States from Image-Elicited User Stories Using Generative AI

MADHURA MUKUND ROHINKAR¹, ASHWINI GARKHEDKAR²
^{1,2}MCA Department, P.E.S Modern College of Engineering Pune, India

Abstract- Conventional sentiment analysis methods primarily depend on direct textual input, which often limits the accurate understanding of deeper psychological states due to controlled or incomplete emotional expression. In response to this limitation, the present study introduces a narrative-based emotional intelligence framework that utilizes image-driven storytelling combined with generative artificial intelligence techniques. Within the suggested methodology, participants are exposed to image-based stimuli and asked to describe or narrate the scene in their own words. The collected narratives are processed through computational linguistic techniques and generative AI to identify emotional cues, sentiment patterns, and contextual meanings. The derived findings are subsequently applied to infer psychological indicators such as emotional state, anxiety tendencies, and optimism levels. The system also incorporates temporal analysis to track variations in user responses over time. The key contribution of this work lies in integrating visual prompts with narrative analysis to capture more natural and expressive emotional responses.

Keywords - Emotional Intelligence, Affective Computing, Generative AI, Natural Language Processing, Image Stimuli, Narrative Analysis, Psychological State Detection

I. INTRODUCTION

Understanding human emotions is essential for building intelligent systems. Traditional sentiment analysis mainly relies on direct text, which often fails to capture deeper psychological states as users may not fully express their emotion.

To overcome this, more natural methods of expression are needed. Visual stimuli, such as images, can encourage individuals to respond more freely through storytelling. This narrative expression helps reveal hidden emotional and cognitive patterns. This research proposes a narrative-based emotional intelligence system that combines image-elicited

stories with generative AI techniques. The system analyzes user narratives to identify emotional cues, sentiment, and context, and then infers psychological states such as mood, anxiety, and optimism. It also uses temporal analysis to track changes across successive sessions.

This approach contributes to affective computing by providing a more natural way to understand emotions, with applications in mental health, education, and human-computer interaction.

Additionally, the approach supports adaptive and personalized analysis, allowing more accurate emotional interpretation based on individual responses.

II. LITERATURE SURVEY

Picard, R. W. (1997). This foundational work introduced the concept of affective computing, enabling machines to recognize and interpret human emotions. It emphasized the role of emotional intelligence in human-computer interaction. However, the study primarily focused on physiological signals and did not explore narrative-based or image-driven emotional expression. This highlights the early focus on observable signals rather than internally expressed emotions.

Pang, B., & Lee, L. (2008). This study provided a comprehensive overview of sentiment analysis techniques for extracting opinions from textual data. It established the foundation for text-based emotion detection. However, it mainly captures surface-level sentiment and does not address deeper psychological states or implicit emotional patterns. Thus, it lacks the ability to interpret nuanced and context-rich emotional expressions.

Mikolov, T., et al. (2013). This research introduced word embeddings, enabling machines to understand contextual relationships between words. It significantly improved natural language processing tasks. However, it focuses on word-level representation and lacks the ability to analyze narrative structure or psychological depth. Thus, it does not capture higher-level narrative meaning.

Cambria, E. (2016). The author proposed sentic computing, which enhances sentiment analysis by incorporating semantic and concept-level understanding. This approach improves emotional interpretation compared to traditional methods. However, it remains dependent on structured text and does not consider spontaneous narrative expression. As a result, it overlooks unstructured and naturally expressed emotional content.

Radford, A., et al. (2018). This work introduced generative pre-trained transformer models capable of understanding and generating human-like text. These models enable deeper narrative interpretation. However, they are typically applied to existing datasets rather than elicited user narratives. Therefore, their potential for guided emotional expression remains underutilized.

Devlin, J., et al. (2019). This study introduced the BERT model, which improves contextual understanding of text by considering bidirectional relationships between words. It significantly enhanced performance in sentiment analysis and emotion detection tasks. However, it focuses only on textual input and fails to integrate visually prompted story generation. Consequently, it remains incapable of recognizing emotions that arise in response to visual cues.

Hazarika, D., et al. (2020). This work focused on conversational emotion recognition using contextual deep learning models. It captures emotional flow in dialogues and improves classification accuracy. However, it is restricted to conversational data and does not include image-based emotional elicitation. Hence, it does not explore emotion generation through external visual prompts.

Guntuku, S. C., et al. (2020). This study analyzed social media text for identifying psychological disorders including depressive symptoms and anxiety using linguistic features. It demonstrated how well NLP-based approaches perform in psychological analysis. However, it relies on passive data collection and lacks controlled narrative elicitation. This reduces the depth and consistency of emotional insights.

Li, J., et al. (2021). This research explored emotion detection using deep learning models trained on large-scale textual datasets. It improved classification of fine-grained emotions. However, it does not consider multimodal inputs or user-generated storytelling. As a result, it overlooks contextual and narrative-based emotional cues.

Kumar, A., et al. (2023). This research focused on AI-based mental health prediction using NLP and deep learning techniques, demonstrating strong potential in identifying psychological patterns from textual data. However, it relies on passive, pre-existing data and lacks interactive or elicitation-based approaches for capturing user responses.

This limits personalization and prevents the system from capturing spontaneous and context-driven emotional expression.

Zhang, Y., et al. (2024). This study explored advanced multimodal AI systems for emotion recognition, integrating text and visual data. It improved overall performance in emotion detection tasks. However, it does not specifically focus on narrative-based emotional analysis or temporal tracking. Thus, it does not fully capture evolving psychological states over time.

1.2 RESEARCH GAP

The majority of current emotion recognition approaches depend on either explicit written content or measurable behavioral indicators, both of which fall short in revealing deeper underlying psychological conditions. These approaches often depend on passive data and fail to encourage natural emotional expression. Furthermore, very few studies have investigated how visually presented prompts combined with participant-authored stories can serve as a basis for emotional evaluation. The integration of

generative AI with narrative understanding and temporal tracking is also underdeveloped.

This gap underscores the necessity of developing a system that enables natural expression and deeper psychological insights through image-elicited storytelling

III. NARRATIVE-BASED EMOTIONAL INTELLIGENCE FRAMEWORK

3.1 Image-Elicited Narrative Approach

This approach captures human emotions indirectly through storytelling using visual stimuli. Instead of asking users to express emotions directly, images are used as prompts to encourage natural and spontaneous responses. Users narrate their observations, frequently projecting their own thoughts and feelings into the narrative. This helps in capturing deeper psychological states with reduced bias.

3.2 Narrative Analysis and Processing

The collected narratives are analyzed using NLP and generative AI techniques for sentiment and contextual understanding. The process includes text preprocessing, sentiment detection, and contextual understanding. Generative AI helps identify subtle emotional patterns and hidden meanings within the narrative, providing a richer interpretation compared to traditional methods.



Fig 1: Proposed Narrative-Based Emotional Intelligence Framework

IV. PSYCHOLOGICAL STATE INFERENCE FRAMEWORK

The Psychological State Inference Framework is designed to translate narrative insights into meaningful psychological indicators. Instead of focusing solely on sentiment classification, the system aims to infer underlying emotional conditions. Key inferred dimensions include:

- Emotional State: Overall mood, such as positive, neutral, or negative.
- Anxiety Indicators: Detection of stress, uncertainty, or concern.
- Optimism Levels: Identification of hopeful or future-oriented expressions.
- Cognitive Patterns: Analysis of thought processes reflected in storytelling.

The framework uses contextual and semantic cues derived from narratives to map user responses to psychological interpretations. This approach offers a broader and deeper insight into a person's emotional condition than conventional sentiment-based approaches.

V. TEMPORAL EMOTIONAL ANALYSIS

Temporal analysis introduces a dynamic dimension to emotional understanding by tracking changes in user responses over time. Rather than treating each narrative as an isolated input, the system observes patterns across multiple interactions.

Key aspects include:

- Trend Analysis: Monitoring shifts in emotional tone across sessions.
- Behavioral Patterns: Identifying consistency or variation in responses.
- Emotional Stability: Assessing fluctuations in psychological states.
- Progress Tracking: Evaluating improvement or decline in emotional conditions.

This longitudinal perspective enables continuous and adaptive emotional assessment, making the system better suited for use cases involving continuous psychological monitoring and individually tailored feedback.

VI. DIMENSIONS OF INTERPRETABILITY AND PERSONALIZATION

This section highlights the importance of making emotional analysis systems transparent, understandable, and user-centric.

Traditional AI-based emotion detection models often function as black-box systems, providing outputs without clear reasoning. In contrast, the proposed framework emphasizes interpretability by linking emotional inferences directly to narrative content and linguistic patterns.

Key dimensions include:

- Interpretability: Clear explanation of how emotional conclusions are derived.
- Context Awareness: Understanding user-specific language and expression styles.
- Personalization: Adapting analysis based on individual response behavior.
- Transparency: Providing meaningful insights rather than just classification results.

By focusing on these dimensions, the system enhances trust and usability, making it suitable for sensitive applications such as mental health and education.

6.1 Interpretability in Emotional Analysis

Interpretability is essential for understanding how the system derives emotional and psychological insights from user narratives. Instead of functioning as a black-box model, the proposed framework ensures that outputs are explainable and traceable.

The system links emotional predictions directly to linguistic features such as keywords, tone, and sentence structure. This allows users and researchers to understand why a particular emotional state is inferred.

Key aspects include:

- Mapping emotional labels to specific words or phrases
- Providing reasoning behind predictions
- Enhancing trust in AI-generated insights
- Supporting validation in sensitive domains like mental health

6.2 Context Awareness and Personalization

Human emotions are highly context-dependent and vary across individuals. The proposed system incorporates context awareness to better understand user-specific language patterns and narrative styles.

Personalization enables the system to adapt to individual users over time. Instead of applying a generic model, the system learns from previous responses to improve accuracy.

Key aspects include:

- Understanding user-specific vocabulary and expressions
- Adapting to cultural and contextual differences
- Learning from historical user narratives
- Improving emotional interpretation accuracy over time

This makes the system more reliable and suitable for real-world applications where individual differences play a crucial role.

6.3 Transparency and Ethical Considerations

Transparency is critical when dealing with emotional and psychological data; keeps users informed about the manner in which their input is handled and the basis on which outcomes are drawn.

Responsible and ethical handling of data holds particular significance when the system is deployed in sensitive domains such as psychological well-being assessment.

Key aspects include:

- Clear communication of system functionality
- Data privacy and confidentiality
- Avoidance of biased or misleading interpretations
- Responsible use of emotional data

This ensures that the system is not only technically effective but also ethically responsible.

VII. SYSTEM ARCHITECTURE AND WORK FLOW

The proposed system integrates image-based input, narrative processing, emotional analysis, and psychological inference into a unified framework for understanding human emotions.

As illustrated in Fig 2, the workflow begins with the image stimuli module, in which participants are shown carefully selected images intended to stimulate both emotional reactions and cognitive engagement. Drawing from these visuals, individuals construct stories that mirror their subjective viewpoints and underlying emotional conditions.

The collected narratives are forwarded to the language processing pipeline, where foundational text preparation steps including tokenization, cleaning, and normalization are applied.

The system then performs emotion detection, extracting sentiment and affective features from the processed text. The extracted features are subsequently fed into the psychological reasoning component to evaluate core markers such as overall mood, anxiety tendencies, and degree of optimism

To provide a dynamic understanding, the temporal analysis module tracks changes in user responses over time, identifying trends and variations in emotional states.

Additionally, the system incorporates personalization and interpretability layers, which enhance user-specific analysis and provide clear explanations for inferred results. Results are displayed via an organized visual interface that delivers actionable understanding of the individual's affective and psychological profile.



Fig 2: System Architecture for Psychological State Inference

VIII. RESULTS AND DISCUSSION

The proposed framework provides deeper emotional insights compared to traditional sentiment analysis by encouraging natural storytelling through image stimuli. The use of NLP and generative AI helps in identifying subtle emotional patterns and contextual meanings.

Additionally, temporal analysis enables tracking of emotional changes over time, making the system more dynamic and effective. Overall, the approach improves emotional understanding, personalization, and interpretability.

CONCLUSION

This research presents a narrative-based emotional intelligence framework that uses image-elicited storytelling and generative AI to infer psychological states. The approach enables more natural and expressive emotional analysis compared to traditional methods.

By integrating NLP, generative AI, and temporal analysis, the system provides a more comprehensive and user-centric understanding of emotions, with applications in mental health, education, and human-computer interaction.

REFERENCES

- [1] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 1997.
- [2] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.
- [3] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *Proc. International Conference on Learning Representations (ICLR)*, 2013.
- [4] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.
- [5] A. Radford et al., "Improving language understanding by generative pre-training," *OpenAI*, 2018.

- [6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proc. NAACL-HLT, 2019, pp. 4171–4186.
- [7] D. Hazarika, S. Poria, R. Zimmermann, and E. Cambria, “Conversational memory network for emotion recognition in conversation,” in Proc. NAACL, 2020.
- [8] S. C. Guntuku, S. Yaden, M. L. Kern, L. Ungar, and J. Eichstaedt, “Detecting depression and mental illness on social media,” *Current Opinion in Behavioral Sciences*, vol. 18, pp. 43–49, 2020.
- [9] J. Li et al., “Deep learning for emotion classification: A survey,” *IEEE Transactions on Affective Computing*, 2021.
- [10] Y. Zhang et al., “Multimodal emotion recognition using deep learning: A survey,” *IEEE Access*, vol. 12, pp. 23456–23478, 2024.