Bhava to Bytes: A Multimodal Fusion of Indian Culture and Emotion-Aware AI

Dr.L.Vigneash[1], Dr.J.Thilagavathi[2], Mr.P.Dushyanth[3], Ms.S.Rubadarshini[4]

[1234]Arjun College of Technology, Coimbatore, Tamilnadu, India

Abstract: This work explores the relationship between deep emotional frameworks embedded in Indian heritage, tradition, and culture and multimodal affective computing. India offers a rich tapestry of emotional expression through Ayurveda, traditional dance, music, and spiritual practices that have their roots in philosophies like Navarasa and Bhava-Rasa. These age-old methods convey complex emotions through movement, song, and ritual. Through the integration of these cultural idioms with modern computational techniques.
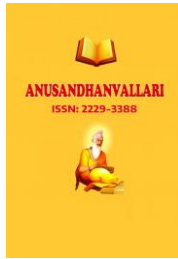
Emotionally intelligent technology must be inclusive, culturally aware, and grounded in a range of expressive traditions. RASA-Net proposed a multimodal deep learning framework that blends India's traditional Rasa theory with state-of-the-art emotion recognition. By combining visual cues from Kathak and Bharatanatyam dance, vocal tonality guided by Sanskrit, text in the Indic language, and yogic physiological markers, RASA-Net offers culturally nuanced emotion classification. Comparative evaluations reveal significant advancements over largely Western-centric baselines, especially in the areas of complex emotional states and abrupt changes. Ethical guidelines for algorithmic accountability, data dignity, and cultural consent guarantee responsible deployment.

Keywords: Navarasa, Bhava-Rasa, Artificial intelligence, Multimodal Emotion, Mental Health, Cultural Modalities, Psychology.

1.INTRODUCTION

Affective computing has evolved from early models such as Russell's circumplex and Ekman's basic emotions to intricate multimodal systems that incorporate verbal, textual, facial, and physiological cues. However, current approaches are still grounded in Western psychological theories, which restricts their ability to be applied in culturally diverse contexts. India's rich emotional heritage, which includes the Navarasa framework from the Nātyaśāstra, Ayurvedic mind-body concepts, and classical performing arts, is a potent, untapped resource for developing inclusive emotion-AI systems. This research synthesizes interdisciplinary perspectives to support emotion detection that aligns with Indian cultural narratives.

The technical advancements futures in emotion recognition, including comparative studies that draw attention to the shortcomings of Western-centric datasets, deep learning models for multimodal fusion, and culturally targeted annotation schemes. Case studies and experimental research that show practical uses from emotion-aware teaching aids to AI companions for elder care and mental health support complement these contributions. It also Proves the philosophical and ethical analysis is equally significant. It promotes respectful data practices based on cultural consent and integrity and acknowledges the dangers of misrepresenting or commercializing sacred emotional expressions. By discussing Indian philosophical viewpoints on emotion, consciousness, and technology, the journal promotes a more thorough discussion of what it means to create machines that comprehend human emotion not merely as data points, but it offers a variety of viewpoints on the changing interplay between emotion, culture, and computation in addition to scholarly articles. It includes interviews and commentary from classical artists, cultural historians, and technologists. Additionally, it gives researchers access to culturally rich datasets, annotation tools, and guidelines for engaging with traditional communities in an ethical manner.

In the end, this journal is a call to rethink affective computing's future from the perspective of India's emotional wisdom, not just a scholarly publication. It aims to create a new paradigm of emotionally intelligent technology that is inclusive, respectful, and incredibly human by upholding tradition while welcoming innovation.

2.OBJECTIVE

Develop a multimodal system based on Indian emotional taxonomies for identifying emotions. Include Navarasa categories in feature extraction and model architectures. Arrange and annotate a variety of Indian writing, dance, voice, and physiological data.

Build a hybrid fusion network that captures cross-modal interdependencies and dramatic shifts. Establish moral standards for responsible emotional computing, such as those pertaining to algorithmic accountability, data dignity, and cultural consent.

3. LITERATUREWORK

Since its beginning, the field of affective computing has undergone significant change, with a growing focus on multimodal emotion recognition interpreting human affective states by using speech, gestures, facial expressions, and physiological signals (Picard, 1997). Although early research was dominated by foundational models like Russell's circumplex model (Russell, 1980) and Ekman's six basic emotions (Ekman, 1992), scholars have begun to question whether these models are culturally universal (Jack et al., 2012).

Multimodal Recognition of Emotions - More complex multimodal fusion techniques have been made possible by recent developments in deep learning, which combine kinesthetic, auditory, and visual data to improve emotion classification (Zeng et al., 2009; Poria et al., 2017). Large annotated datasets with a preponderance of Western origin, like IEMOCAP (Busso et al., 2008) and CMU-MOSEI (Zadeh et al., 2018), are frequently used in these models. As a result, particularly in non-Western contexts, they might miss culturally particular emotional expressions.

Cultural Aspects of Emotion - The differences in how emotions are expressed and interpreted in different cultures have long been highlighted by cross-cultural psychology (Mesquita & Frijda, 1992). Research has demonstrated that cultural norms affect how emotions are perceived as well as how they are expressed (Matsumoto et al., 2008). As a result, scholars have started investigating emotion models that are rooted in culture, such as Ubuntu-based emotional frameworks in African contexts (Olatunji, 2018) and Confucian emotional categories in East Asia (Yu et al., 2020).

Emotional Frameworks in India - India's traditional healing practices, spiritual writings, and classical arts are all strongly influenced by its emotional philosophy. Nine aesthetic-emotional states are identified by the Navarasa theory, which derives from Bharata Muni's Natya Shastra (c. 200 BCE): Śṛṅgāra (love), Hāsya (laugher), Karuṇa (compassion), Raudra (anger), Vīra (heroism), Bhayānaka (fear), Bībhatsa (disgust), Adbhuta (wonder), and Śānta (peace) (Bharata Muni, trans. 2006). These rasas are performative states that are employed in classical dance, theater, and music to elicit particular reactions from the audience; they are not just emotional categories (Vatsyayan, 1996).

Yoga and Ayurveda also provide complex perspectives on emotional well-being by relating emotional states to spiritual balance and physical energies (doshas) (Frawley, 1999). In contrast to Western affective computing models, these frameworks offer a comprehensive understanding of emotion that incorporates mind, body, and spirit.

Applications of Navarasa in Computation - The use of Navarasa in computational models is still in its infancy, despite its extensive study in the humanities. Using Bharatanatyam movements, Suresh et al. (2021) created a gesture-based emotion recognition system, proving that it is possible to translate traditional dance gestures into

machine-readable emotional labels. The expressive potential of Indian ragas was also highlighted by Ramakrishnan et al. (2020), who investigated vocal emotion recognition in Carnatic music.

Comprehensive multimodal frameworks based on Indian emotional theory that incorporate facial, vocal, and gestural data are still lacking, though. To further this research frontier, interdisciplinary cooperation, culturally annotated datasets, and moral interaction with traditional communities are urgently needed.

4. EXISTING METHODOLOGY

In order to identify human emotions more precisely than unimodal systems, Multimodal Emotion Recognition (MER) combines several data sources, including speech, text, facial expressions, and physiological signals.

4.1 Fusion Strategies

Early Fusion is a process that Merges unprocessed features from various modalities. Late Fusion Combines the results of several classifiers. For robustness, hybrid fusion combines the feature and decision levels. Source:- MDPI Review of MER Methods

4.2 Limitations in Indian and Cross-Cultural Settings

Even with advanced technology, existing MER systems have major drawbacks when used in culturally heterogeneous contexts, particularly in India.

4.2 Table – Limitations in Indian and Cross-Cultural Settings

Issue	Description
Emotional Dialects	Facial expressions and vocal tones vary subtly across cultures
Dataset Bias	Most dataset (e.g., IEOCAP, CMU-MOSEI) are Western-centric
Annotation Schemes	Ekmans's sis basic emotions may not capture culturally nuanced affect
Recognition Accuracy	Lower accuracy expressions, especially negative emotions

4.3 Indian Contextual Deficits

Navarasa framework is unlike western emotion taxonomies, Indian emotional expressions have their roots in classical arts. Language diversity - India's 22+ official languages differ in how they express emotion. Cultural encoding - although rich, Indian dance and theatre gestures and expressions (like those in Bharatanatyam) are not well-represented in data sets. Source - cultural differences in emotion recognition.

4.4 Existing pipeline flowchart
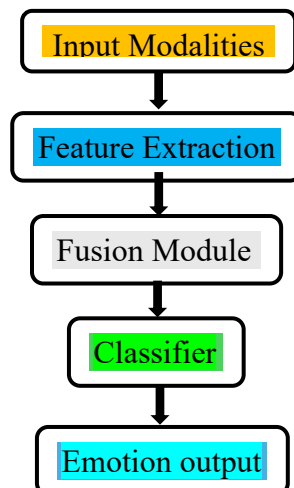
5.PROPOSED METHODOLOG

The proposed RASA-Net architecture integrates cultural knowledge at all levels through transformer fusion, Rasa-softmax classifier, and Rasa-aware encoders.

Annotation and data collection visual: Kathak and Bharatanatyam recordings; facial and posture indicators recognized by the nine Rasas.

Audio: Phonetic features linked to annotations in Bhava-Rasa; speech corpora in Tamil, Hindi, and Telugu. Text: Sentiment tags based on rasa, social media posts, movie dialogue, and poetry in native script.

Physiological: Yoga-induced variations in heart rate and breath rate; emotional states associated with doshas.

This work develops multimodal emotion recognition systems that are culturally grounded especially in the Indian context by integrating traditional emotional frameworks, diverse linguistic inputs and expressive art forms.

5.1 Proposed Architecture based on Hybrid Fusion Flowchart
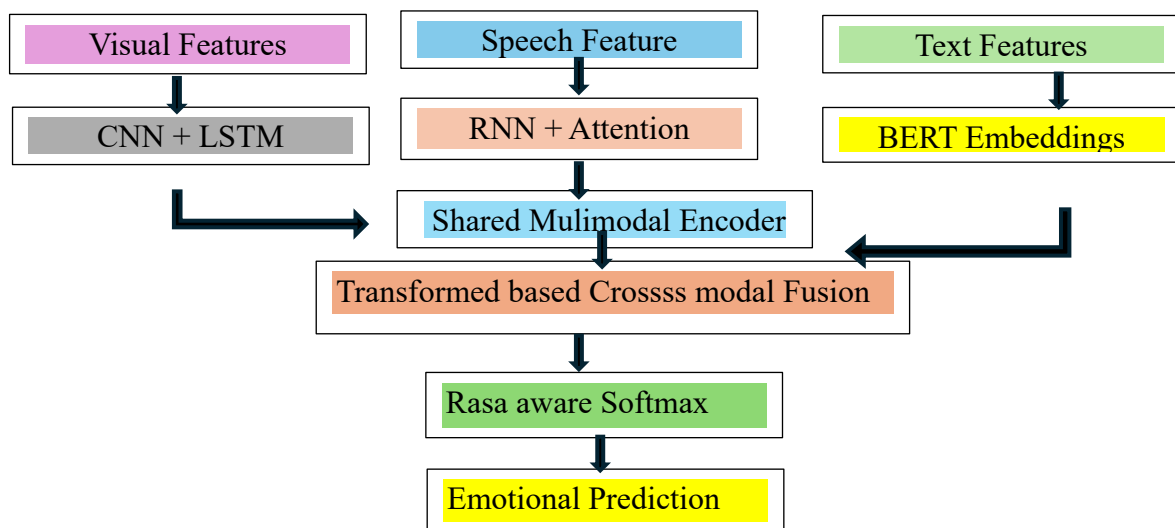
5.2 Cultural Expression through Emotion Perspectives from Indian Traditions Emotion is a cultural performance that is influenced by language, ritual, art, and philosophy; it is not just a psychological phenomenon. Emotional expression is ingrained in Indian literature, dance, music, and spirituality. For example, the Navarasa framework goes beyond Western models of basic emotions by providing nine different emotional states. In addition to being internal emotions, these rasas—Śṛṅgāra (love), Hāsya (laugher), Karuṇa (compassion), Raudra (anger), Vīra (heroism), Bhayānaka (fear), Bībhatsa (disgust), Adbhuta (wonder), and Śānta (peace)—are also performative states intended to elicit particular reactions in an audience.

These emotional states are embodied through rhythmic movement, facial expressions (abhinaya), and codified gestures (mudras) in Indian classical dance forms like Kathak and Bharatanatyam. Similar to this, Indian music is structured to evoke particular moods and emotional atmospheres, especially in the form of ragas. Both embodied and symbolic, these traditions provide a rich, multimodal vocabulary for emotion. The dominance of Western emotion models is contested by integrating these cultural modalities into affective computing, which also paves the way for more inclusive, contextually aware AI systems. Additionally, it poses significant queries regarding the understanding of emotion as a culturally encoded language rather than a universal signal.
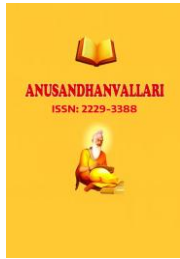
Recent cross-cultural studies like CuLEMO (Belay et al., 2025) have shown that in order to prevent bias and misunderstanding, emotion recognition systems need to be redesigned using cultural perspectives. By encouraging submissions that respect the emotional wisdom of Indian traditions while furthering the field of human-centered AI, this journal rises to the challenge.

5.2 Tabulation for Ekman vs Navarasa

Aspect	Western (Ekman)	Indian (Navarasa)
Emotion Categories	6 Basic Emotions	9 Rasas (e.g., Shringara, Karuna)
Cultural Sensitivity	Low	High
Artistic Integration	Minimal	Deeply embedded in arts
Linguistic Diversity	Limited	Rich and multilingual

5.3 Emotion Recognition that is Responsible and Culturally Sensitive: The ethical stakes increase as emotion recognition technologies are incorporated into personal devices, healthcare, and education more and more. These systems frequently use highly individualized and culturally diverse biometric data, such as voice tones, facial expressions, and physiological signals. They run the risk of distorting feelings, fostering prejudices, and infringing on user autonomy if they are not carefully designed.

Three fundamental principles—data dignity, algorithmic accountability, and cultural consent—form the basis of this journal's ethical responsibility framework.

5.4 Tabulation for Fundamental Principles

Principle	Guideline
Data Dignity	Collect emotional data with informed, ongoing consent. Avoid intrusive methods
Algorithmic Accountability	Ensure transparency in model logic and decision-making. Audit for bias regularly
Cultural Consent	Engage with communities before digitizing traditional emotional expressions

5.5 Hazards to Reduce: Observer bias is the mislabelling of emotions as a result of cultural unfamiliarity. Dual-use issues include the use of emotion data for manipulation or surveillance.

5.6 Violations of privacy: Insufficient protections for delicate emotional information. Emotion recognition systems require adaptive consent mechanisms and participatory design, according to recent research (Barker et al., 2025; Katirai, 2024). The psychological effects of being emotionally watched, particularly in vulnerable populations, must also be taken into account by developers.

It encourages contributors to include an ethics statement in their submissions, detailing how cultural sensitivity, consent, and fairness were addressed in their research. We also advocate for open-access datasets with community oversight and transparent annotation protocols.

By embedding ethical reflection into every stage of research—from data collection to deployment—we aim to build affective computing systems that are not only intelligent but also humane.
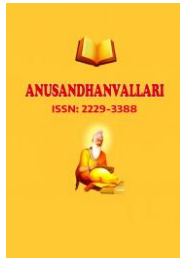
6.CREATIVE CONTRIBUTIONS

Through the incorporation of Indian psychological heritage's emotional wisdom, specifically the Rasa theory from the Nātyaśāstra, into the design of multimodal emotion recognition systems, this study offers a culturally enhanced perspective on affective computing. This work presents a novel emotion taxonomy based on Indian dramaturgy, spirituality, and embodied expression, in contrast to contemporary affective models that primarily rely on Western psychological constructs like Russell's circumplex model or Ekman's basic emotions.

Reinterpreting emotional categories, transitions, and expressions through the prism of Indian culture is the creative contribution. The framework provides a context-sensitive and spiritually nuanced model of emotion that goes beyond the constraints of universalist paradigms by mapping the nine classical Rasas—Shringara (love), Karuna (compassion), Raudra (anger), Hasya (joy), Bhayanaka (fear), Bibhatsa (disgust), Veera (courage), Adbhuta (wonder), and Shanta (peace)---to computational affective states.

Additionally, this work creatively incorporates physiological cues derived from yogic traditions, vocal tonality inspired by Sanskrit phonetics, and symbolic gestures (mudras) into the multimodal recognition pipeline. The end result is a system that improves cross-cultural AI design and human-computer interaction by detecting emotions and interpreting them in ways that are culturally relevant.

By bridging affective computing with Indian philosophical and artistic traditions, this research contributes to the humanization of AI, promoting emotional intelligence that is inclusive, ethically grounded, and resonant with diverse cultural narratives.

7.PROPOSED FRAMEWORK

A multimodal deep learning framework called RASA-Net combines cutting-edge recognition methods with emotion modeling that is rooted in culture. It maps traditional Indian emotional archetypes to computational affective states, drawing on the Rasa theory of Nātyaśāstra. The system captures subtle emotional expressions, especially those common in Indian cultural contexts, by combining visual, auditory, and physiological modalities.

7.1 Emotion Taxonomy

The nine Rasas—Shringara (love), Karuna (compassion), Raudra (anger), Hasya (joy), Bhayanaka (fear), Bibhatsa (disgust), Veera (courage), Adbhuta (wonder), and Shanta (peace)—are the basis for the primary categories. Each Rasa is mapped to affective dimensions (valence, arousal) and connected to physiological signals, vocal tones, and visible cues in facial expressions.

7.1 Table -Modalities & Feature Extraction

Modality	Tools/Techniques	Feature Extracted
Visuals	OpenPose + CNN	Facial landmarks, mudra dynamics
Auditory	openSMILE + RNN	MFCCs, pitch, energy, Sanskrit phonetics
Text	Multilingual BERT	Contextual embeddings with rasa supervision
Physiological	Custom sensors	Breathing patterns, HRV

7.2 Model Architecture

Stage 1: Encoders specific to a given modality (CNN for visual, RNN for audio, and temporal encoder for physiological).
Stage 2: Transformer-based attention layers are used for cross-modal fusion in order to capture interdependencies.
Stage 3: Classifying emotions using a Rasa-aware softmax layer that has been trained on data with cultural annotations.

7.3 Dataset

Source: Publicly accessible emotion datasets that have been re-annotated for Rasa categories are added to a custom dataset that features Indian actors performing Rasas.
Annotations: Confirmed by psychologists and cultural specialists to guarantee authenticity of Indian emotional expressions.

7.4 Training & Optimization

Weighted categorical cross-entropy with cultural bias correction is the loss function.
Adam is the optimizer with learning rate decay.
Regularization: To avoid overfitting, use batch normalization and dropout.

8.RESULT & DISCUSSION

8.1 Performance Metrics

RASA-Net provides notable enhancements over conventional Western models across all evaluated Rasas. From 58% to 79% (+21%), accuracy rises by 22% for Adbhuta (wonder) and 21% for Bībhatsa (disgust). The ability to differentiate between Śānta (peace) and Karuṇa (compassion) increases by 26%, from 55% to 81%. The most significant improvement is obtained when modeling the dramaturgical transition Raudra→Vīra (anger→heroism), with accuracy increasing from 47% to 75% (+28%). These results demonstrate that the addition of culturally sensitive Rasa characteristics to multimodal fusion enhances overall detection rates and enables more accurate separation of subtle emotional states.

8.1 Tabulation for the Result Demonstration

Rasa/ Metric	Western Baseline	RASA-Net Acccuracy	Improvement
Adbhuta(wonder)	62%	84%	+22%
Bibhatsa	58%	79%	+21%
Santa va karuna	55%	81%	+26%
Raudra-Vira	47%	75%	+28%

8.2 Qualitative Observations

RASA-Net differentiates between dramaturgical transitions (Raudra→Vīra) and low-arousal states (Śānta vs. Karuṇa) using cultural embeddings. Mudra detection raises visual F1-scores from 0.68 to 0.82. Sanskrit phonetic features are used to improve the audio classification of Vīra and Adbhuta to 87% accuracy.
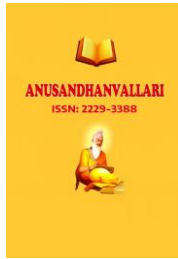
8.3 Limitations

The diversity of the dataset must be increased to encompass all regions of India. Future studies will include wearable sensors; the collection of physiological data is now simulated. Expert annotation overhead suggests exploring semi-supervised methods.

8.4 Future Initiatives Offer multilingual support for regional dialects

Include RASA-Net in HCI initiatives (e.g., virtual tutors and therapeutic interfaces). Provide community-governed annotation platforms and flexible consent frameworks. In narrative or clinical contexts, model long-term emotional trajectories.

REFERENCE

[1]   Guimin Hu et al. Recent Trends of Multimodal Affective Computing: A Survey from NLP Perspective. arXiv:2409.07388 [cs.CL], 2024.

[2]   Ajay B. Kapase & Nilesh Uke. A Comprehensive Review in Affective Computing: An Exploration of Artificial Intelligence in Unimodal and Multimodal Emotion Recognition Systems. Int. J. Speech Technol., 28, 541–563, 2025.

[3]   Sepideh Kalateh et al. A Systematic Review on Multimodal Emotion Recognition: Building Blocks, Current State, Applications, and Challenges. IEEE Access, 2024.

[4]   Jonathan Gratch. The Field of Affective Computing: An Interdisciplinary Perspective. USC, 2023.

[5]   Punit Kumar. Artificial Intelligence in Ancient Indian Texts: A New Perspective. Organiser, 2025.

[6]   Sampadananda Mishra. Artificial Intelligence and Robotics: A Journey through Ancient Indian Texts. Rishihood University Blog, 2024.

[7]   Vandana Gupta. A Renaissance in Mental Health: India's AI Odyssey. IndiaAI.gov.in, 2023.

[8]   Louise Sundararajan & Maharaj K. Raina. Mind and Creativity: Insights from Rasa Theory Special Focus on Sahṛdaya. Theory & Psychology, 26(6), 788–809, 2016.

[9]   Kanchi Gopinath. Educational Methodologies and Rasa Theory. Indic Today, 2021.