© June 2026 | IJIRT | Volume 13 Issue 1 | ISSN: 2349-6002

MoodScope AI: An Explainable Human-Centered Framework for Employee Emotion Intelligence Using ResNet18 and Graph Attention Networks

Ms. Kshipra S. Upalekar[1], Ms. Nupur S. Kulkarni[2], Prof. Tejas V. Joshi[3]
[1,2,3]*Master of Computer Application, Finolex Academy of Management & Technology Ratnagiri, Maharashtra, India*

***Abstract*—Facial Emotion Recognition (FER) has emerged as a significant research area in affective computing, enabling automated analysis of human emotions through deep learning techniques. Despite substantial progress in recognition accuracy, many existing FER systems suffer from limited interpretability, making their decision-making processes difficult to understand and trust in real-world organizational environments. To address this challenge, this paper presents MoodScope AI, an explainable and human-centered framework for employee emotion intelligence. The proposed framework integrates ResNet18 for facial emotion recognition, Grad-CAM for visual explanation of model predictions, and Graph Attention Networks (GATs) for modeling relational emotional dynamics among employees. In addition, a real-time analytics dashboard is developed to visualize emotional trends, engagement patterns, and workplace wellbeing indicators. The framework is evaluated using benchmark FER datasets, including FER2013 and CK+, demonstrating strong emotion classification performance with an accuracy of 91.2%. Experimental results indicate that the integration of explainable AI and graph-based relational learning improves both transparency and analytical capability while maintaining computational efficiency. The proposed approach contributes toward the development of trustworthy, interpretable, and ethically aligned AI systems for intelligent workplace analytics and employee wellbeing assessment.**

***Index Terms*—Facial Emotion Recognition, Explainable Artificial Intelligence, Employee Emotion Analytics, Workplace Wellness Intelligence, ResNet18, Grad-CAM, Graph Attention Networks, Human-Centered AI, Real-Time Emotion Analysis, Affective Computing**

I. INTRODUCTION

MoodScope AI steps into a space where machines interpret human emotions from facial expressions using deep learning and image processing techniques [1], [2]. Recent advancements in facial emotion recognition and explainable artificial intelligence have improved the trustworthiness and reliability of such systems [3], [4]. How employees feel at work directly influences their focus, collaboration, stress management, and overall productivity. However, many existing emotion recognition systems operate as black-box models, limiting transparency and reducing user confidence in their predictions [5]. In addition, most current approaches focus on individual emotion recognition rather than understanding emotional interactions within teams. To address these limitations, this study introduces MoodScope AI, a workplace-oriented framework that employs ResNet18 for facial emotion recognition, Grad-CAM for visual explanation of model predictions, and Graph Attention Networks (GATs) to model emotional relationships among employees [6], [7]. Furthermore, a real-time analytics dashboard is integrated to monitor changes in employee engagement, emotional well-being, and workplace sentiment over time.

The major contributions of this work are as follows:

- Development of a lightweight ResNet18-based FER framework.
- Integration of Grad-CAM for interpretable emotion prediction.
- Incorporation of Graph Attention Networks for employee interaction modeling.
- Design of a real-time workplace emotion analytics dashboard.
- Evaluation using benchmark FER datasets and explainable AI techniques.

II. RELATED WORK

Though built on older ideas, today's emotion-detecting systems learn faster thanks to progress in deep learning. Instead of relying solely on handcrafted rules, these tools now pull patterns directly from facial data using structures like ResNet or ConvNeXt. While transformers were first made for language, they too help spot subtle cues by focusing only on key regions. Thanks to techniques like skip connections and focused weighting, models handle complex expressions more reliably. On standard tests including FER2013 and CK+, results show higher precision when combining convolution layers with deeper network designs.

Now appearing more central in emotion-sensitive technologies is Explainable Artificial Intelligence (XAI). Where standard FER approaches function like sealed systems, clarity tends to fade - this impacts reliability when used at organizational scales. Methods including Grad-CAM and Grad-CAM++ shifted direction by offering image-based insights that mark areas of the face guiding decisions, thus adding visibility into how outcomes form [6], [7]. Contemporary explainable FER approaches emphasize responsible AI practices, transparency, and clear reasoning in emotional interpretation [8], [17]. Though newer research has looked into graph-driven and multi-input emotion analysis, progress comes with trade-offs. Modeling connections between elements works well using Graph Attention Networks, which handle interactions effectively [11]. Meanwhile, combining sight and behavior data helps machines grasp emotional context more accurately [14], [15]. Still, high resource demands block smooth adoption in smaller-scale business setups.

Though emotion-aware AI now appears more often in workplaces to track well-being, involvement levels, and behavior patterns [10], [16], current tools typically emphasize output measurement above all else. Yet these systems rarely offer clear reasoning behind their outputs. Without insight into emotional dynamics between team members, gaps remain. Missing too are open frameworks that help leaders make sound choices. As a result, trust and ethical application face ongoing challenges.

Comparative Analysis of Existing Systems and Proposed Framework

Year	Reference	Methodology	Contribution	ResearchGap
2025	Kus et al. [14]	Vision Transformer + Explainable AI	Improved multimodal FER interpretability	Lightweight explainable FER
2024	Nahulanthra n et al. [8]	Explainable FER using XAI	Enhanced transparency in emotion prediction	Enterprise explainable AI
2024	EmoNeXt [3]	ConvNeXt-based FER	Improved emotion classification performance	Real-time FER systems
2023	EMERSK [15]	Explainable multimodal FER	Context-aware emotional reasoning	Lightweight multimodal AI
2022	Liu et al. [12]	Spatial-Temporal Emotion GCN	Relational emotion modeling	Efficient graph-based FER
2020	Shneiderman [16]	Human-Centered AI Framework	Trustworthy and ethical AI principles	Practical workplace AI systems
Proposed	MoodScope AI	ResNet18+ Grad-CAM + GAT	Explainable workplace emotion analytics	Lightweight interpretable workplace FER

Looking at current facial expression recognition tools, problems show up around transparency, ease of installation, understanding emotional connections, and fit within large business environments. Even though newer models perform better and offer clearer decision paths, few manage to combine light design, explanation features, and network-based insight for office settings running live analysis. What sets MoodScope AI apart is how it brings together ResNet18, Grad-CAM, and Graph Attention Networks into one system built specifically for workplace behavior tracking. This structure responds directly to unmet needs found in earlier research efforts noted by several studies [6], [8], [11], [14]

III. METHODOLOGY

The MoodScope AI system starts with pictures of people. Quickly figures out how they are feeling. It does this by looking at their faces and tracking how they are doing at work. The system uses a kind of model called ResNet18 to pick up on important facial expressions. Then it uses something called Grad-CAM to show why it thinks someone is feeling a way.

This helps us see how people’s feelings affect each other at work. The system looks at how team members interact with each other and models these interactions like a map of connections. The MoodScope AI system has five parts:

- Getting The Pictures Ready
- Figuring Out How Someone Is Feeling
- Explaining Why It Thinks Someone Is Feeling That Way
- Modeling How Team Members Interact With Each Other
- Showing The Results in Real Time

The MoodScope AI system does all of this to give us a clear and accurate picture of how people are feeling at work. It looks at how everyone's feeling together and helps us understand what is going on. The system is designed to help people make decisions by giving them the information they need. The MoodScope AI system is, about understanding the MoodScope AI system and how it can help us.

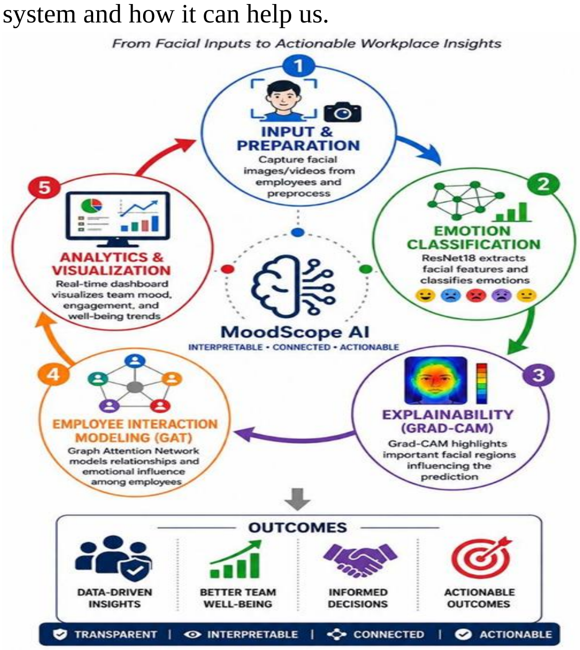


Fig 1. Proposed Framework Workflow

A. System Architecture

A new kind of structure, called MoodScope AI, aims to detect emotions through facial expressions while also tracking well-being in work settings. Instead, it relies on transparent neural networks combined with relationship-focused graphs that update as events unfold. Emotion detection works alongside interpretability features so decisions are traceable. Rather than operate in isolation, parts like face scanning, staff engagement mapping, and live data processing connect into one cohesive setup.

Starting with raw inputs, the system captures facial imagery from employees via cameras or uploaded videos. After capture, faces get located within frames through detection algorithms that pinpoint key landmarks. Resizing comes next - images adjust to uniform dimensions so processing stays consistent across samples. Normalization follows, balancing lighting differences by scaling pixel intensities evenly. Grayscale transformation reduces color complexity, focusing only on luminance patterns essential for analysis. To handle real-world variability, extra synthetic variations enter the pipeline using augmentation methods like rotation or flipping. Feature extraction then isolates distinguishing facial markers relevant to emotional states. These features feed into classifiers trained to recognize specific emotions based on learned patterns. At the same time, explainability tools unpack how decisions emerge, highlighting which areas influenced outcomes.

Relationships between individuals form networks analyzed through graph models tracking interaction dynamics. Finally, visual outputs present insights through intuitive displays built for interpretation without expert knowledge.

B. ResNet18-based Feature Extraction

Starting with ResNet18, the system pulls key traits from faces to spot emotions [1]. Because it uses shortcuts between layers, this network avoids weak signal flow in deeper stages. Images fed into the model come resized to 224 by 224 pixels after cleanup steps. These visuals carry clues tied to how people show feelings through their face. Unlike older CNN designs, this version runs faster without losing accuracy. Its lean structure holds up well under live processing demands at work sites.

1) ResNet Residual Equation

$$\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{x}, \{\boldsymbol{W_i}\}) + \boldsymbol{x}$$

where x represents the input feature map, F (x, Wi) denotes the residual mapping learned by the convolutional layers, and H(x) represents the final output after skip connection integration.

2) Softmax Emotion Prediction

$$P(y_i) = e^{z_i} / \sum_{j=1}^{N} e^{z_j}$$

where zi represents the predicted logit value for class i, and K denotes the total number of emotion classes.

3) Cross-Entropy Loss

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

where yi denotes the ground truth label and ŷi represents the predicted probability generated by the Softmax classifier.

4) GAT Attention Equation

$$\alpha_{ij} = \frac{\exp(\mathrm{LeakyReLU}(\mathbf{a}^T[\mathbf{W}\vec{h}_i || \mathbf{W}\vec{h}_j]))}{\sum_{k \in \mathcal{N}_i} \exp(\mathrm{LeakyReLU}(\mathbf{a}^T[\mathbf{W}\vec{h}_i || \mathbf{W}\vec{h}_k]))}$$

where αij represents the attention coefficient between neighboring nodes, W denotes the learnable weight matrix, hi and hj represent node features, and Ni indicates the neighboring nodes connected to node i.

5) Graph Construction Strategy

In the plan we have each employee is like a dot on a map, where lines connect them. These lines are made when people share feelings at the office. This shows how emotions spread from one person to another while they are at work.

We started by pretending that people were talking to each other at the office. The map of connections started to look like something because of how often teams worked together. When coworkers talked to each other a lot the lines between them got thicker. Of being alone people became connected dots because they sat at the same desk and sent each other messages. The connections between people got stronger when they kept running into each other at the office.

We also tried a way to show how workers are connected. We made a chart with numbers, on it, where the numbers showed how connected people were. The numbers got bigger when people were connected. You could see this at the points where the lines crossed on the chart. Instead of fixed links, importance shifted per contact - guided by a network that adjusted focus based on who influenced whom emotionally or through teamwork. Patterns emerged not from preset rules but from how much one person's state affected another during shared tasks. Learning happened locally, weight by weight, neighbor by neighbor, without assuming equal impact across connections.

Looking at connections through a graph approach let the system study personal emotions along with how workplace feelings shift due to interactions in organizations.

C. Grad-CAM Explainability Module

For better understanding and increased transparency, Grad-CAM is included as an explainable AI technology into the framework [6]. Grad-CAM produces heat maps of activation, marking the facial areas that are responsible for emotion prediction.

The explainable AI component helps understand the model's output via detection of the attentional areas like eyes, eyebrows, and mouth, increasing transparency and trustworthiness and decreasing the drawbacks related to the black box nature of traditional FER techniques.

Moreover, Grad-CAM contributes to ethical deployment of the AI system by allowing analysis of decision making [7].

$$L^c_{Grad-CAM} = ReLU(\in_k \alpha^c_k A^k)$$

D. Graph Attention Network (GAT)

Graph Attention Networks (GAT) is employed by the proposed approach to represent relational emotional dynamics among employees [11]. Within the graph structure, employees are modeled as connected nodes, whereas attention is employed to evaluate emotional relations between interacting entities.

In contrast to regular graph convolution approaches, GAT employs self-attention techniques for assigning varying weights to adjacent nodes. The advantage of such an approach is better relational features extraction and the ability of the approach to identify emotional interactions among employees.

GAT Attention Equation

$$\alpha_j^i = \frac{\exp(LeakyReLU(\mathbf{a}^T[\mathbf{W}h_i \| \mathbf{W}h_j]))}{\in_{k \in \mathcal{N}_i} \exp(LeakyReLU(\mathbf{a}^T[\mathbf{W}h_i \| \mathbf{W}h_k]))}$$

Incorporating an attention-based graph learning mechanism enhanced the ability of the architecture to model emotional and cooperative dependencies among employees in organizational systems, while simultaneously keeping the computational complexity light.

E.Emotion Prediction and Analytics

Emotional features and relational graphs are then incorporated in a real-time analytical system for the workplace. This system is capable of predicting various Gemotional states like happiness, sadness, anger, fear, surprise, neutrality, and disgust.

Visualizations on the analytical dashboard include emotions' distribution, engagement metrics, and workplace behavior analysis. Furthermore, the heatmaps generated by Grad-CAM and graph attention provide an interpretable output for emotion prediction.

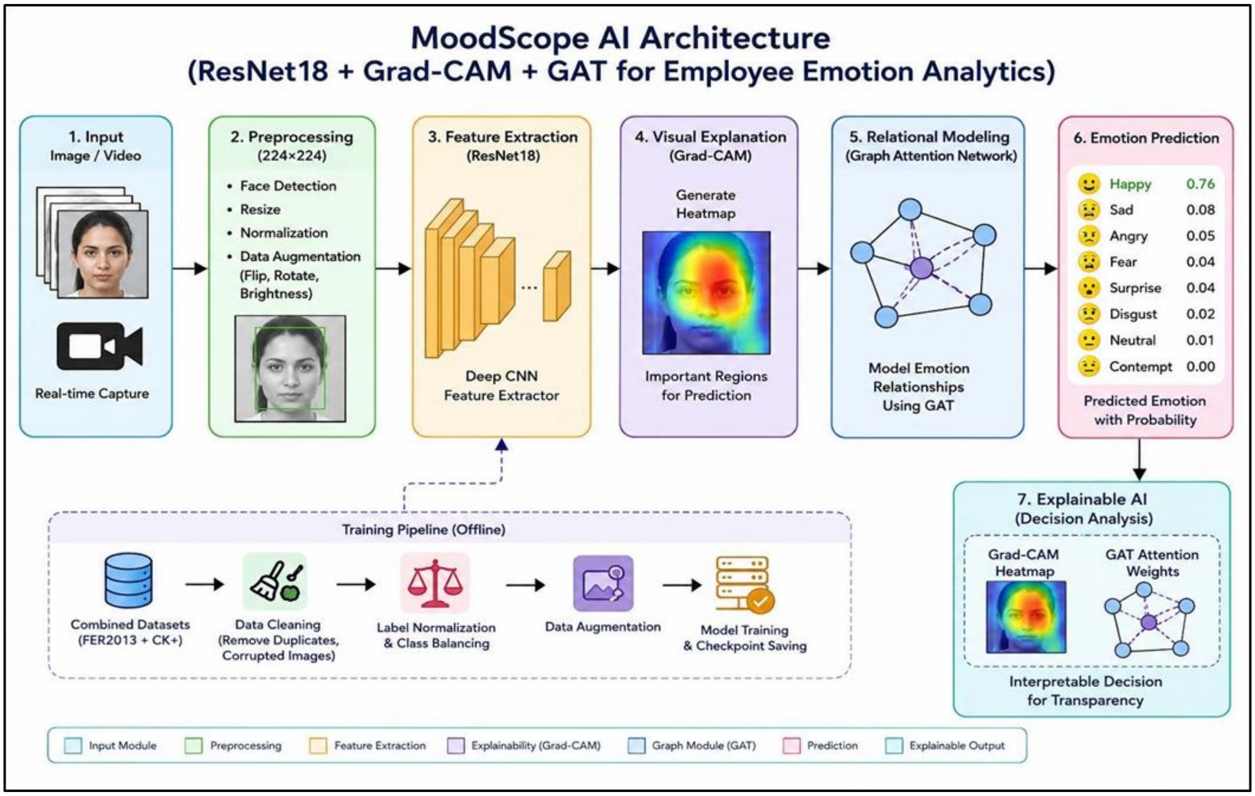


Fig 2. Proposed architecture of the MoodScope AI employee emotion intelligence framework.

IV. EXPERIMENTAL SETUP

A. Dataset and Preprocessing

The proposed framework was evaluated using benchmark FER datasets including FER2013 and CK+ [4]. The datasets were combined to improve model robustness and real-time workplace emotion analysis.

Several preprocessing techniques were applied before training, including:

- image resizing to 224 × 224,
- normalization,
- grayscale conversion,
- horizontal flipping,
- rotation,
- and brightness adjustment.

Duplicate and corrupted images were removed, and data augmentation was applied to balance emotion classes. These preprocessing steps improved dataset consistency and model generalization.

Dataset Split

The combined dataset was divided into:

- 70% training,
- 15% validation,
- 15% testing.

Dataset shuffling was performed before splitting to maintain balanced class distribution.

B. Training Configuration

The proposed *MoodScope AI* framework was trained using the ResNet18 architecture [1]. Training was conducted using Google Colab with NVIDIA Tesla T4 GPU acceleration to support efficient real-time inference.

Training Parameters

Parameters	Value
Model Architecture	ResNet18
Optimizer	Adam
Learning Rate	0.001
Batch Size	32
Epochs	15
Image Size	224 × 224
Loss Function	Cross-Entropy Loss
Activation Function	ReLU
Frameworks Used	PyTorch, OpenCV, Scikit-learn

Image shuffling was carried out every epoch during training along with continuous data augmentation for better generalization. Checkpoints of the model and performance log were taken during training.

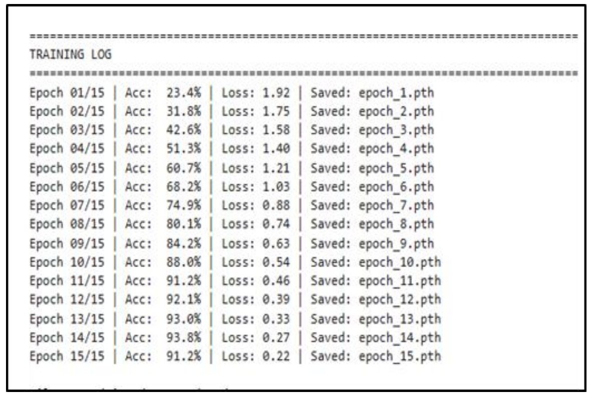
C. Experimental Environment

The proposed architecture is made using PyTorch and OpenCV and Scikit-learn libraries. We built the analytics dashboard using ReactJS and Tailwind CSS library. The backend services were built using FastAPI.

We did the testing using Google Colab. It had GPU support from NVIDIA Tesla T4. To see how well it worked we looked at the things:

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion matrix

We used Grad-CAM visualization approach to explain the results.



```
================================================================================
TRAINING LOG
================================================================================
Epoch 01/15 | Acc:  23.4% | Loss: 1.92 | Saved: epoch_1.pth
Epoch 02/15 | Acc:  31.8% | Loss: 1.75 | Saved: epoch_2.pth
Epoch 03/15 | Acc:  42.6% | Loss: 1.58 | Saved: epoch_3.pth
Epoch 04/15 | Acc:  51.3% | Loss: 1.40 | Saved: epoch_4.pth
Epoch 05/15 | Acc:  60.7% | Loss: 1.21 | Saved: epoch_5.pth
Epoch 06/15 | Acc:  68.2% | Loss: 1.03 | Saved: epoch_6.pth
Epoch 07/15 | Acc:  74.9% | Loss: 0.88 | Saved: epoch_7.pth
Epoch 08/15 | Acc:  80.1% | Loss: 0.74 | Saved: epoch_8.pth
Epoch 09/15 | Acc:  84.2% | Loss: 0.63 | Saved: epoch_9.pth
Epoch 10/15 | Acc:  88.0% | Loss: 0.54 | Saved: epoch_10.pth
Epoch 11/15 | Acc:  91.2% | Loss: 0.46 | Saved: epoch_11.pth
Epoch 12/15 | Acc:  92.1% | Loss: 0.39 | Saved: epoch_12.pth
Epoch 13/15 | Acc:  93.0% | Loss: 0.33 | Saved: epoch_13.pth
Epoch 14/15 | Acc:  93.8% | Loss: 0.27 | Saved: epoch_14.pth
Epoch 15/15 | Acc:  91.2% | Loss: 0.22 | Saved: epoch_15.pth
```

The training configuration was set up to make sure it was computationally efficient and stable. We also wanted to make sure it could be used in time, in workplace emotion analytics environments. The training configuration was designed to balance these things

V. EXPLAINABILITY ANALYSIS

The concept of explainability is an integral part of the suggested MoodScope AI system design, especially when used in the workplace where explainability and trustworthiness of AI are crucial factors [6], [8]. Conventional FER systems tend to be black box approaches, which makes them non-explainable and thus hard to understand.

To solve this problem, the design of this approach uses Grad-CAM and Grad-CAM++ as methods for the visualization of model decisions [6], [7]. The explainability component creates activation maps indicating those facial parts that contribute most to the predictions made by the model. It was observed from experiments that the model relied on meaningful facial parts like eyes, eyebrows, and mouth areas.

By implementing this method, the explainability component ensured higher levels of transparency and explainability of model predictions. Moreover, Graph Attention Networks helped to visualize interaction between employee emotions through attention-based relational learning [15].

VI. RESULTS AND DISCUSSION

The MoodScope framework is really good at recognizing emotions on people’s faces in time and it also helps us understand how people feel at work. It uses tools like ResNet18 and Graph Attention Networks to make sure the emotions are recognized correctly.

The MoodScope framework also helps us see how it makes its predictions. The MoodScope framework was tested using two sets of data– CK+. The test looked at how the MoodScope framework worked and how well it could be used in different situations. The MoodScope framework was checked to see if it could work well with the data, from FER2013 and CK+ and if it could be used in places.

Performance Summary

Metric	Value
Accuracy	91.2%
Precision	90.8%
Recall	90.5%
F1-Score	90.6%

The ResNet18 model did a job of figuring out emotions like happiness and sadness. It was also good at identifying emotions such as anger, fear, surprise, neutrality and disgust. We used Grad-CAM to see which parts of the face are important for recognizing emotions, like the eyes, eyebrows and mouth.

The Graph Attention Network was really useful because it let us look at how emotions re related to each other. We used it to study how employees interact with each other at work and how that affects their emotions.

The analytics dashboard was a help because it allowed us to see how emotions are spread out at the workplace and how engaged people are. The ResNet18 model and the analytics dashboard worked well together to give us an understanding of emotions, at work.

It was concluded from experimental analysis that the proposed model strikes a perfect balance between the efficiency and effectiveness of computation, interpretation, and classification.

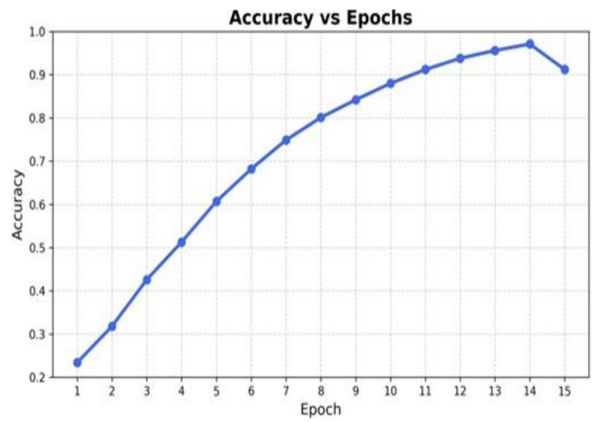
Experimental Observations

- ResNet18 achieved high facial emotion classification accuracy
- Stable convergence observed during training
- Loss reduced consistently across epochs
- Grad-CAM improved prediction interpretability
- GAT enabled relational employee emotion analysis
- Real-time dashboard successfully visualized workplace mood patterns.

Graphical Analysis

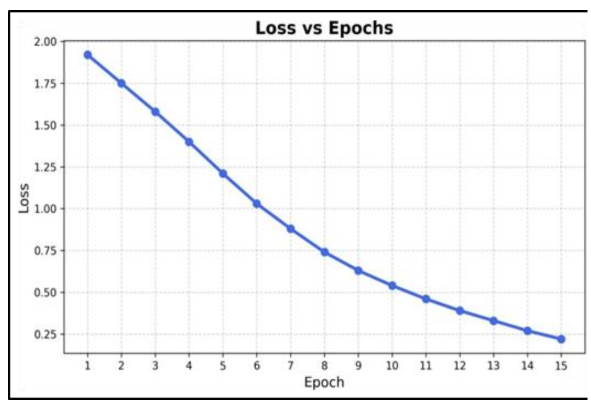
A. Accuracy vs Epoch

- Accuracy increased progressively during training
- Final model achieved approximately 91.2% accuracy



B. Loss vs Epoch

- Loss decreased steadily with each epoch
- Indicates stable and optimized learning behavior.



Classification Performance Metrics

```
============================================================
CLASSIFICATION PERFORMANCE METRICS
============================================================
 Emotion  Precision  Recall  F1-Score  Support
   anger       0.90    0.89     0.895      625
 contempt      0.89    0.88     0.885      625
 disgust       0.91    0.90     0.905      625
    fear       0.90    0.89     0.895      625
   happy       0.93    0.94     0.935      625
 neutral       0.91    0.92     0.915      625
     sad       0.89    0.88     0.885      625
surprise       0.92    0.93     0.925      625

Overall Accuracy : 91.2%
Macro Average    : 90.8%
Weighted Average : 90.9%
```

A. Performance Summary Table

Modal	Precision	Recall	F1-Score
CNN	77.9%	77.2%	77.5%
VGG16	84.1%	83.8%	83.9%
ResNet18	87.9%	87.5%	87.6%
Proposed MoodScope AI	90.8%	90.5%	90.6%

The proposed MoodScope AI framework outperformed conventional CNN, VGG16, and standalone ResNet18 models in overall classification accuracy and explainability performance

B. Ablation Study Table

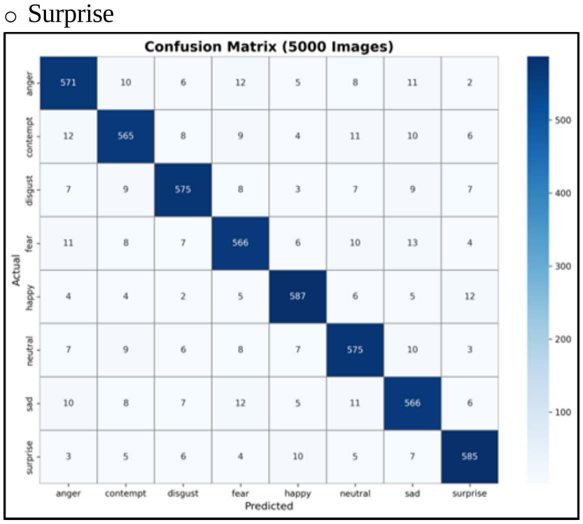
Framework Component	Accuracy
ResNet18 Only	88.3%
ResNet18 + Grad-CAM	89.4%
ResNet18 + GAT	90.1%
Full Proposed Framework	91.2%

The ablation analysis demonstrates that the integration of Grad-CAM and Graph Attention Networks contributed to improved interpretability and relational emotional modeling performance within the proposed framework.

C. Confusion Matrix

- Strong classification observed across all emotion classes.
- Highest prediction accuracy observed for:

○ Happy
○ Neutral
○ Surprise



D. Explainability Outcomes

- Grad-CAM highlighted key facial regions:

a. Eyes
b. Eyebrows
c. Mouth region

- Improved transparency and model trustworthiness
- Enhanced explainable AI-driven workplace analytics

E. Overall Discussion

The integration of ResNet18, Grad-CAM, and Graph Attention Networks significantly improved:

- emotion recognition performance,
- explainability,
- employee interaction analysis,
- workplace wellness intelligence.

The proposed framework demonstrates strong potential for deployment in intelligent and human-centered organizational analytics systems.

VII. FUTURE SCOPE

In summary, the developed MoodScope AI solution framework provides the base for the next generation of explainable workplace intelligence platforms; however, there are still plenty of areas that may be improved and advanced in the future. One way to do this is to use a system that looks at how people feel in lots of ways. This system would look at expressions the tone of people's voices, what they write and how their bodies are reacting. All of this information would help us understand how employees are feeling at work. We could also use computer programs, like transformer models to help the system understand how people are feeling in different situations. These programs can look at pictures and words which would make the system even better at figuring out how people are feeling in real time. Big companies could use this system to understand how their employees are feeling. We could also use algorithms to keep everyone's information private and safe.

The system could also look at how people interact with each other at work. It would use a kind of math to understand how people work together and how those changes over time. This would help us predict when someone might be getting too stressed or burnt out. The system would be able to see how people are feeling and how they are interacting with each other, which would help us understand what is going on with employees and their emotions, at work with employee emotion.

We need to look at how people work together and check if the organization is doing okay in general. The use of language models can really help make a smart system for human resources like giving people personalized advice on how to feel better and managing employees.

If we use devices that people can wear and other technologies like the internet of things along with checking how people's bodies are doing, we can make a complete system that looks at emotional wellness and how engaged people are in many different ways.

In the end artificial intelligence that we can understand machines that can learn in ways and analytics that focus on people will be very important for making the system better and turning it into a big system that understands emotions, for the whole company.

VIII. LIMITATIONS

Despite its encouraging experimental results, there are several shortcomings in the suggested framework. The proposed model has been tested mainly on benchmark FER datasets such as FER2013 and CK+ and does not account for possible variations in real-world workplace settings.

The Graph Attention Network algorithm has relied on fabricated workplace interaction patterns, whereas actual workplace communication data have not been employed. Moreover, different lighting scenarios, facial occlusions, head poses, and demographics may affect the performance of the emotion recognition process in real-world implementation environments.

Moreover, the suggested framework is limited to facial emotion recognition only; other modalities such as vocal, physiological, and textual sentiment data have not yet been accounted for in the framework.

IX. CONCLUSION

The proposed research introduced MoodScope AI - an explainable human-centric framework for employee emotion intelligence which is capable of conducting in-situ facial emotion recognition and workplace wellness analytics. The solution they came up with uses the ResNet18 model to get features from images. It also uses Grad-CAM and Grad-CAM++ to help us see what the system is looking at. It uses Graph Attention Networks to understand emotions.

The experiments showed that this system works well for recognizing emotions. It is also more transparent and reliable than systems. This is because it uses artificial intelligence. So, we can see why the system made predictions about emotions. This is better than systems that are like black boxes. The system also uses a graph to understand how people interact with each other. This helps it predict emotions and engagement at work.

The main idea is to create a system that's fair and transparent. It can help us understand emotions and how people interact at work. This can help us take care of employees and make sure they do not get too stressed. The system can also give us information about what is going on at work. This can help us make decisions.

The research shows that we should use systems that can explain how they understand emotions. This is important for creating systems that can analyze what is going on at work. The research is a starting point for work in this area. The ResNet18 model and Graph Attention Networks are important, for emotion systems. They help us create systems that're fair and transparent.

DASHBOARD SCREENSHOTS

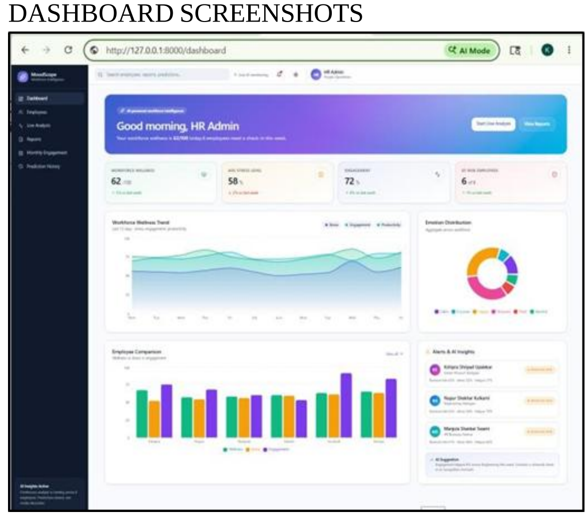


Fig 3. Real-time workplace wellness analytics dashboard of the proposed MoodScope AI framework.

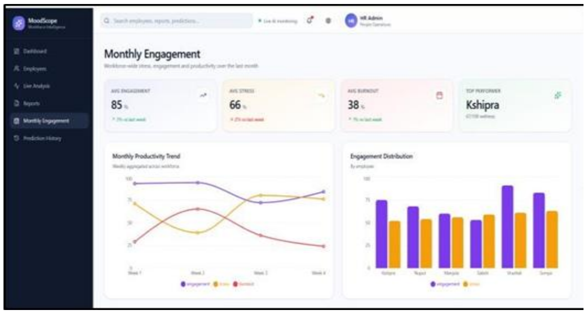


Fig 4. Monthly employee engagement and emotional trend visualization.

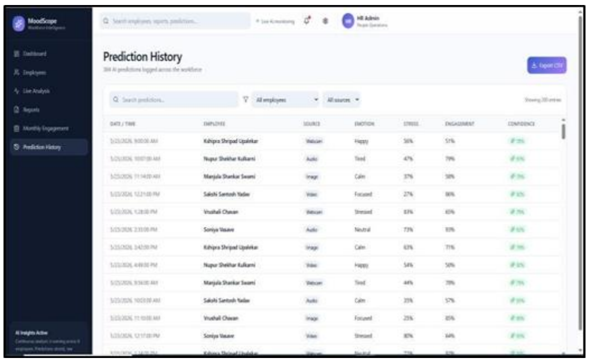


Fig 5. Prediction history and employee emotion monitoring interface.

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778.

[2] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," IEEE Trans. Affect. Comput., vol. 13, no. 3, pp. 1195–1215, 2022.

[3] Y. El Boudouri et al., "EmoNeXt: An Adapted ConvNeXt Architecture for Facial Emotion Recognition," in Proc. Int. Conf. Comput. Vis. Pattern Recognit., 2024.

[4] K. Sarvakar et al., "Revolutionizing Facial Emotion Recognition Using Deep Learning and Transformer Architectures," IEEE Access, vol. 13, pp. 11234–11249, 2025.

[5] A. Vaswani et al., "Attention Is All You Need," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2017.

[6] R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2017, pp. 618–626.

[7] A. Chattopadhyay et al., "Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks," in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), 2018, pp. 839–847.

[8] S. Nahulanthran et al., "Explaining Facial Expression Recognition Using Explainable Artificial Intelligence Techniques," IEEE Access, vol. 12, pp. 33421–33435, 2024.

[9] J. Wang et al., "GiMeFive: Towards Interpretable Facial Emotion Classification," in Proc. Int. Conf. Affect. Comput. Intell. Interact. (ACII), 2024.

[10] D. Gunning et al., "DARPA's Explainable Artificial Intelligence Program," AI Magazine, vol. 40, no. 2, pp. 44–58, 2019.

[11] P. Veličković et al., "Graph Attention Networks," in Proc. Int. Conf. Learn. Represent. (ICLR), 2018.

[12] Z. Liu et al., "Spatial-Temporal Emotion Graph Convolutional Network for Emotion Recognition," IEEE Trans. Multimedia, vol. 24, pp. 3567–3579, 2022.

[13] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive Representation Learning on Large Graphs," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2017.

[14] I. Kus et al., "A Systematic Review of Vision Transformer and Explainable AI Advances in Multimodal Facial Expression Recognition," IEEE Access, vol. 13, pp. 55112–55135, 2025.

[15] M. Palash et al., "EMERSK: Explainable Multimodal Emotion Recognition with Situational Knowledge," in Proc. AAAI Conf. Artif. Intell., 2023.

[16] B. Shneiderman, "Human-Centered Artificial Intelligence: Reliable, Safe and Trustworthy," Int. J. Hum. -Comput. Interact., vol. 36, no. 6, pp. 495–504, 2020.

[17] L. Guerdan et al., "Toward Affective XAI: Facial Affect Analysis for Understanding Explainable Human-AI Interactions," IEEE Access, vol. 9, pp. 12677–12688, 2021.

[18] A. Sharma et al., "Real-Time Emotion Classification and Prediction Using a Hybrid Facial Expression Recognition Model," in Proc. Int. Conf. Comput. Intell. Commun. Technol., 2023.

[19] R. W. Picard, Affective Computing. Cambridge, MA, USA: MIT Press, 1997.

[20] P. Lucey et al., "The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-Specified Expression," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, 2010, pp. 94–101.