Affect2Act: Graph Attention Networks for Emotion-Informed Decision Making

Jash Vora JASH.VORA22@SPIT.AC.IN
Sardar Patel Institute of Technology, Mumbai, India

Yash Shah YASH.SHAH22@SPIT.AC.IN
Sardar Patel Institute of Technology, Mumbai, India

Editors: List of editors' names

Abstract

Current affective computing systems focus on recognizing emotions but struggle to use them for reasoning, which limits human-AI interaction. We introduce Affect2Act, a graph attention model that represents emotions as interconnected nodes and learns context-dependent relationships. This structure enables flexible reasoning over emotional states. In a synthetic decision-making benchmark, Affect2Act achieves 87.67% accuracy with an emotional balance of 65.47%, slightly exceeding Graph Convolutional Networks (86.73%) and clearly surpassing the attention-only (70.67%) and MLP (45.67%) baselines. These findings suggest that graph-structured emotional representations can support more robust, emotion-informed AI systems and open new directions for affective reasoning.

Keywords: Graph Neural Networks, Affective Computing, Attention Mechanisms, Emotion Modeling

1. Introduction

Modern AI systems largely overlook emotional context, despite its central role in human cognition (Barrett, 2017). While significant progress has been made in emotion *recognition* from facial expressions (Ekman, 1971), speech (Cowie et al., 2001), and multimodal signals (Kollias et al., 2019; Zhang et al., 2018), the ability to use emotions for *reasoning* remains an open challenge. This limitation is critical as AI systems increasingly operate in human-centered environments.

Most existing methods treat emotions as independent features, in contrast to psychological evidence that emotions form structured, interdependent systems (Plutchik, 1980; Russell, 1980). For example, fear and surprise often co-occur in threat assessment, while joy and trust interact differently in social contexts. Capturing these relationships requires models that move beyond flat feature representations.

Graph Neural Networks (GNNs) provide a natural framework for modeling such relational structures (Kipf and Welling, 2016; Veličković et al., 2018). They have been applied successfully across domains such as molecules (Gilmer et al., 2017) and social networks (Hamilton et al., 2017), yet they remain underexplored in affective computing. Psychological models like Plutchik's emotion wheel and Russell's circumplex (Plutchik, 1980; Russell, 1980) suggest that graph-based approaches could capture context-dependent links between emotions more effectively than traditional methods.

In this work, we introduce **Affect2Act**, a Graph Attention Network that represents emotions as interconnected nodes and learns their context-sensitive relationships through multi-head attention. This enables more flexible emotion-informed decision making compared to approaches that assume fixed or independent emotion structures.

Our contributions: (1) a graph-based representation of emotions grounded in psychological theory; (2) a multi-head Graph Attention Network integrating emotional and environmental context; (3) regularization strategies that encourage balanced use of emotions and improve decision performance.

2. Method

We frame emotion-informed decision making as learning a mapping $f : (\mathcal{E}, \mathcal{C}) \to \mathcal{A}$, where $\mathcal{E}$ are emotional states, $\mathcal{C}$ is environmental context, and $\mathcal{A}$ are possible actions. Emotions are represented as a graph $G = (V, E)$ with six fundamental emotions (joy, fear, anger, sadness, surprise, trust) as nodes, grounded in psychological theory (Plutchik, 1980). Each node is associated with features encoding contextual activations, intensity, valence, and emotion-specific profiles. Edges model relationships between emotions and are refined during training.

To capture both intrinsic and context-dependent emotion relationships, we employ a multi-head Graph Attention Network (GAT) (Veličković et al., 2018). For each layer l, node representations are updated via attention-weighted message passing:

$$h_i^{(l+1)} = \sigma\left(\sum_{k=1}^{K} \sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(k)} W^{(k)} h_j^{(l)}\right),$$

where $\alpha_{ij}^{(k)}$ denotes attention coefficients computed from concatenated features of nodes i and j. Stacked GAT layers yield emotion-specific embeddings that preserve relational structure.

Environmental context is encoded separately and fused with emotional embeddings using cross-modal attention:

$$o = \text{MultiHeadAttention}(Q_{\text{env}}, K_{\text{emotion}}, V_{\text{emotion}}),$$

allowing decisions to depend on both situational cues and relational emotion dynamics. A feedforward decision network maps the fused representation to discrete actions.

A common failure mode in affective models is collapse to dominant emotions. To mitigate this, we apply lightweight regularization encouraging (i) higher entropy of predicted emotion distributions, (ii) balance across emotions, and (iii) variance across samples. These penalties prevent over-reliance on a single emotion and promote richer emotional reasoning. Full formulations are provided in our code[1].

We evaluate Affect2Act on a synthetic dataset of six scenario types (e.g., *emergency*, *social*, *analytical*), where each sample defines environment features, emotion activations, and a ground-truth decision label. The generator introduces structured variability, noise, and scenario-specific emotion patterns, providing a controlled testbed for studying emotion-informed reasoning.

1. Code available in supplement.

3. Experiments

We evaluate Affect2Act on a synthetic benchmark spanning six scenario types (*emergency*, *social*, *analytical*, *conflicted*, *creative*, and *supportive*). Each dataset contains 10,000 samples, split into 70% training, 15% validation, and 15% test. Each sample specifies environmental features, emotion activations, and a ground-truth decision label.

We compare against: (1) **GCN**, using the same emotion graph but without attention; (2) **Multi-Head Attention**, treating emotions as independent sequences; (3) **MLP**, using concatenated features without explicit structure.

The model is trained with a combined objective:

$$\mathcal{L}_{total} = \mathcal{L}_{classification} + \lambda_{reg}\,\mathcal{L}_{regularization},$$

where $\mathcal{L}_{classification}$ is standard cross-entropy and $\mathcal{L}_{regularization}$ encourages entropy, balance, variance, and low correlation across emotions. This prevents collapse to dominant emotions and promotes diversity. We set $\lambda_{reg} = 0.1$ with fixed weights for each component. A learnable temperature parameter τ is applied to adjust the sharpness of emotion distributions. Full definitions of all regularization terms are provided in the supplement and released code.

We report two complementary metrics: - **Decision Accuracy (↑)**: proportion of correct action predictions. - **Emotional Balance (↑)**: normalized entropy of predicted emotion distributions, measuring diversity in emotional activation.

4. Results

Table 1 shows that our approach achieves superior performance across both metrics.

Table 1: Performance Comparison

Model	Decision Accuracy	Emotional Balance
Affect2Act (GAT)	**0.8767**	0.6547
GCN	0.8673	0.4053
Multi-Head Attention	0.7067	**0.6619**
MLP	0.4567	0.4737

Affect2Act achieves the highest decision accuracy (87.67%) while maintaining a balanced use of emotions (65.47%). GCN reaches comparable accuracy but collapses to dominant emotions (40.53% balance), showing that fixed graph structures cannot sustain diversity. Multi-Head Attention preserves diversity (66.19%) but suffers in accuracy (70.67%), suggesting that sequential modeling is insufficient. MLP performs the weakest overall, underscoring the limitations of unstructured features. While the improvement in decision accuracy over GCN is modest (+0.94%), Affect2Act achieves a substantially higher emotional balance (+25 points). This joint optimization highlights that our model reaches a more stable equilibrium between decision accuracy and emotional diversity which crucial for robust and interpretable affective reasoning.

Discussion and Future Work

Our findings show that graph-structured representations enable richer, emotion-informed reasoning in AI systems. Although the numerical improvement over GCN is modest, the rise in emotional balance demonstrates that attention-based modeling enhances diversity and stability in emotional reasoning—something fixed-structure models fail to capture. The emotional balance metric thus emerges as a key indicator of reasoning robustness rather than a secondary measure. Affect2Act achieves strong accuracy while maintaining balance, demonstrating that both can coexist through regularization-driven equilibrium.

Attention analysis also reveals interpretable patterns: fear–anger links during emergencies reflect fight-or-flight responses, while joy–trust interactions align with theories of positive emotions. A visual and quantitative analysis of these emotional graphs would further support the interpretability claims.

Future Directions: While Affect2Act shows consistent gains, its scope is limited by the fixed six-emotion design and synthetic setting, which we acknowledge as current limitations. Future work will explore richer or data-driven emotion taxonomies, multimodal extensions, and real-world evaluations while considering computational and ethical aspects such as bias amplification.

5. Conclusion

We presented Affect2Act, a Graph Attention Network for emotion-informed decision making. By modeling emotions as interconnected nodes with context-dependent attention, Affect2Act achieves strong decision accuracy (87.67%) while preserving emotional balance (65.47%). Compared to baselines that either collapse to dominant emotions or neglect relational structure, our approach demonstrates that graph-structured emotional representations open new directions for interpretable and emotionally intelligent AI.

Acknowledgments

We thank anonymous reviewers for feedback and acknowledge ethical considerations in emotion-aware AI systems.

References

L. F. Barrett, *How Emotions Are Made: The Secret Life of the Brain*. Pan Macmillan, 2017.

A. S. Cowen and D. Keltner, “Self-report captures 27 distinct categories of emotion bridged by continuous gradients,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 38, pp. E7900–E7909, 2017.

R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, “Emotion recognition in human-computer interaction,” *IEEE Signal Processing Magazine*, vol. 18, pp. 32–80, Feb. 2001. doi:10.1109/79.911197.

P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, vol. 17, no. 2, p. 124, 1971.

N. H. Frijda, *The Emotions: Studies in Emotion and Social Interaction*. Cambridge University Press, Cambridge, UK; New York, USA, 1986.

J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 1263–1272, 2017.

W. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," arXiv preprint arXiv:1706.02216, 2017.

T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations (ICLR)*, 2017. Available at: https://openreview.net/forum?id=SJU4ayYgl.

D. Kollias, P. Tzirakis, M. Nicolaou, A. Papaioannou, G. Zhao, B. Schuller, I. Kotsia, and S. Zafeiriou, "Deep affect prediction in-the-wild: Aff-Wild database and challenge, deep architectures, and beyond," *International Journal of Computer Vision*, vol. 127, 2019. doi:10.1007/s11263-019-01158-4.

R. W. Picard, *Affective Computing*. MIT Press, Cambridge, MA, 1997. Available at: https://arl.human.cornell.edu/linked%20docs/Picard%20Affective%20Computing.pdf.

R. Plutchik, "A general psychoevolutionary theory of emotion," in *Theories of Emotion*, pp. 3–33, Elsevier, 1980.

J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, p. 1161, 1980.

J. A. Russell, "Core affect and the psychological construction of emotion," *Psychological Review*, vol. 110, no. 1, p. 145, 2003.

P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *International Conference on Learning Representations (ICLR)*, 2018. Available at: https://openreview.net/forum?id=rJXMpikCZ.

S. Zhang, S. Zhang, T. Huang, and W. Gao, "Multimodal deep convolutional neural network for audio-visual emotion recognition," in *Proceedings of the 2016 ACM International Conference on Multimedia*, pp. 281–284, 2016. doi:10.1145/2911996.2912051.

Appendix A. Appendix

A.1. Full Training Objective

The total training loss is the combination of classification error and regularization terms:

$$\mathcal{L}_{total} = \mathcal{L}_{classification} + \lambda_{reg}\,\mathcal{L}_{regularization}, \tag{1}$$

where the classification component is standard cross-entropy:

$$\mathcal{L}_{classification} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log(\hat{y}_{i,c}). \tag{2}$$

A.2. Regularization Terms

The regularization term is defined as a weighted sum of four components:

$$\mathcal{L}_{regularization} = \alpha\mathcal{L}_{entropy} + \beta\mathcal{L}_{balance} + \gamma\mathcal{L}_{variance} + \delta\mathcal{L}_{correlation}, \tag{3}$$

with coefficients $\lambda_{reg} = 0.1$, $\alpha = 1.0$, $\beta = 0.3$, $\gamma = 0.2$, $\delta = 0.1$ in all experiments.

Entropy Regularization. Encourages non-degenerate emotion distributions by penalizing low entropy:

$$\mathcal{L}_{entropy} = \left|\sum_{j=1}^{E} p_j \log p_j + 0.7\log E\right|^2. \tag{4}$$

Balance Regularization. Prevents dominance of specific emotions by aligning average activations with a uniform distribution:

$$\mathcal{L}_{balance} = \sum_{j=1}^{E} \bar{p}_j \log\left(\frac{\bar{p}_j}{1/E}\right). \tag{5}$$

Variance Regularization. Encourages sample-wise diversity across the dataset:

$$\mathcal{L}_{variance} = \sum_{j=1}^{E} \max(0, 0.01 - \mathrm{Var}(p_{\cdot,j})). \tag{6}$$

Correlation Regularization. Discourages highly correlated emotion activations:

$$\mathcal{L}_{correlation} = \frac{1}{E(E-1)}\sum_{i\neq j} |\mathrm{Corr}(p_{\cdot,i}, p_{\cdot,j})|. \tag{7}$$

Here, p_j represents predicted probabilities for emotion j, $\bar{p}_j$ is the dataset mean activation, and $E = 6$ is the number of emotion nodes.

A.3. Temperature Scaling

We apply a learnable temperature parameter τ to control the sharpness of emotion probability distributions:

$$p_j^{scaled} = \frac{\exp(z_j/\tau)}{\sum_{k=1}^{E}\exp(z_k/\tau)}. \tag{8}$$

Lower τ sharpens distributions, while higher τ smooths them.

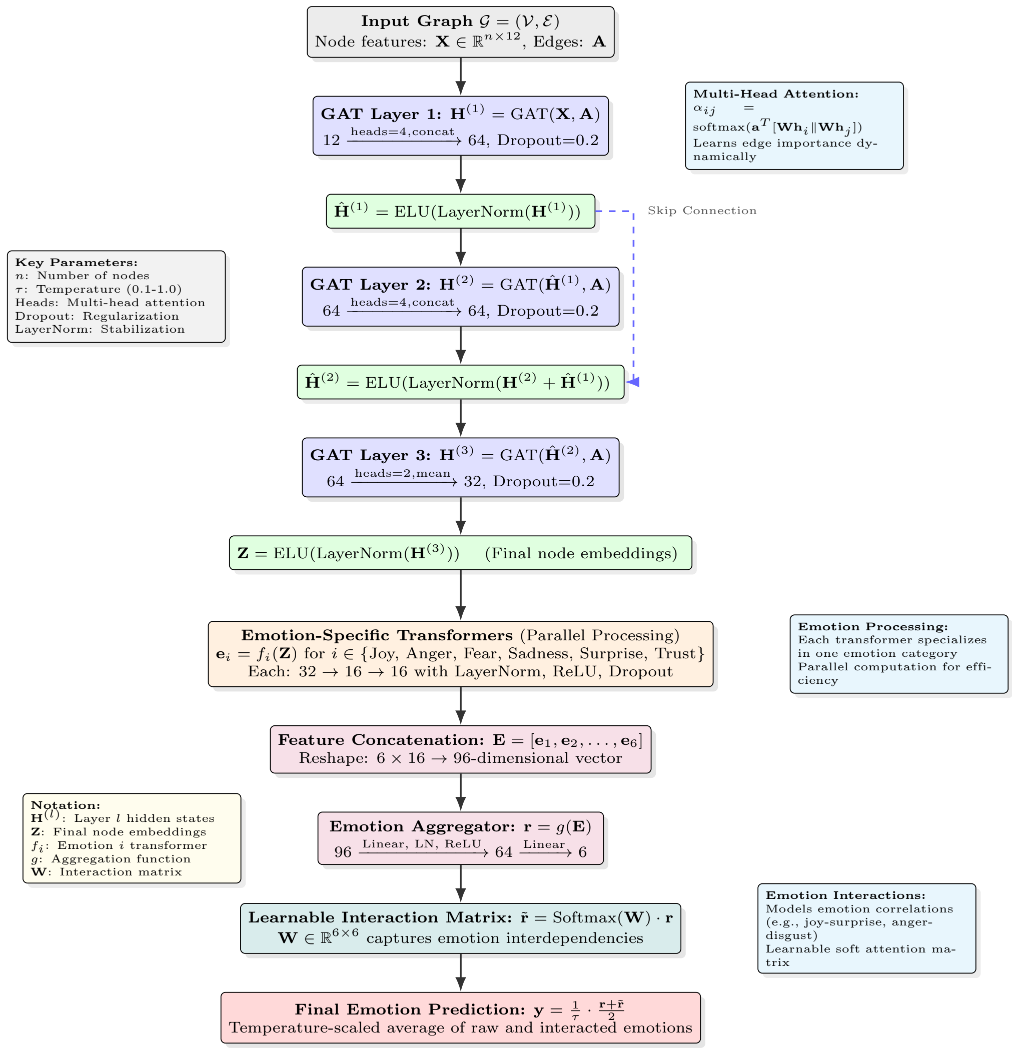

Figure 1: Affect2Act architecture with GAT layers, emotion-specific transformers, and an interaction matrix for decision prediction.

A.4. Model Architecture Details

Figure 1 outlines Affect2Act. The model applies three GAT layers with residual connections to encode context-dependent emotion nodes. Each node is then refined by emotion-specific transformers, and their outputs are aggregated into a joint representation. A learnable 6×6 interaction matrix further adjusts relationships between emotions, with the final decision derived from a temperature-scaled blend of the aggregated and interaction-enhanced vectors. This design combines relational modeling, emotion specialization, and interaction reasoning in a compact architecture.

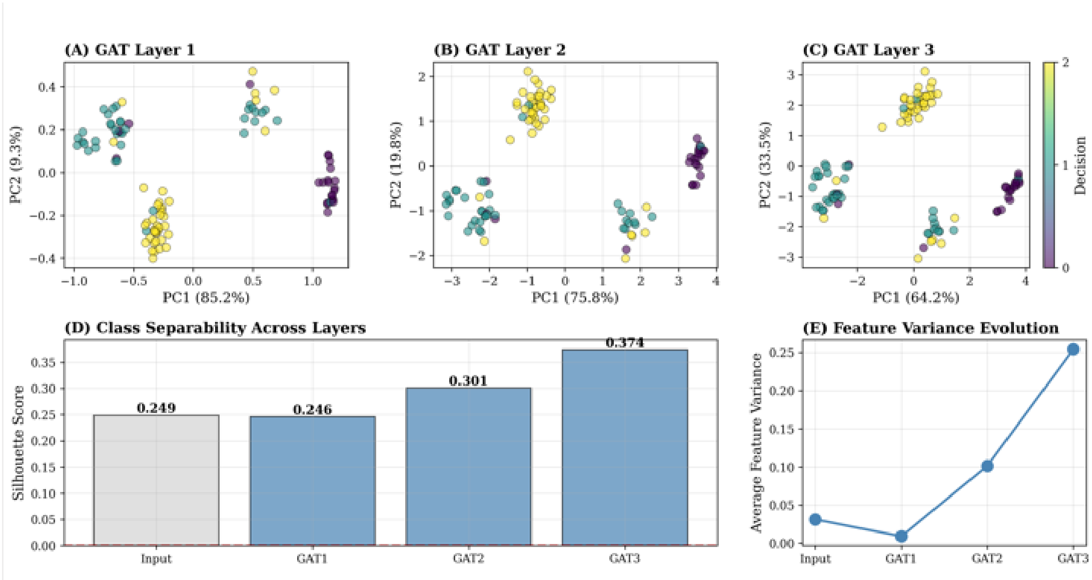

Figure 2: Layer-wise representation learning in GAT. (A-C) PCA projections of emotion embeddings across three GAT layers, colored by decision class. (D) Silhouette scores showing progressive class separability. (E) Feature variance evolution demonstrating representation refinement through successive layers.

A.5. Layer-wise Representation Learning

Figure 2 visualizes how emotion representations evolve through successive GAT layers. The PCA projections (A–C) show that emotion embeddings become increasingly separated by decision class as they pass through deeper layers, with principal component variance gradually concentrating (PC1: 85.2% $\rightarrow$ 64.2%). The silhouette score analysis (D) confirms progressive improvement in class separability, rising from 0.249 at the input to 0.374 after the third GAT layer. Feature variance (E) initially decreases as the model filters noise, then increases as discriminative patterns emerge. This progression demonstrates that the GAT architecture successfully refines emotion representations for decision-making, with each layer contributing to more structured and task-relevant embeddings.

A.6. Implementation Notes

All models are implemented in PyTorch with `torch-geometric`. Affect2Act uses three GAT layers followed by emotion-specific transformations, and a decision network integrating environmental context via multi-head attention. We optimize with AdamW, use cosine annealing learning-rate schedules, and apply gradient clipping for stability. Synthetic datasets are generated with scenario-specific emotion patterns, controlled variance, and reproducible seeds. Full implementation details are provided in the released code.