Fusion Methods Results Summary

Overview

This document summarizes the results of fusing text and image/video emotion scores to match Gemini's combined emotion scores.

Fusion Method

Method 4: Score-Proportional Dynamic Fusion

Dynamically assigns weights proportional to the strength of text and visual scores.

Formula:
w_text = text_score / (text_score + visual_score + epsilon)
w_visual = visual_score / (text_score + visual_score + epsilon)

where epsilon is a small constant to prevent division by zero.

The final fused score is:
final_score = w_text * text_score + w_visual * visual_score

Results Summary

Method	MSE (lower is better)	Cosine Similarity (higher is better)	Pearson Correlation (higher is better)
Method 1 (Max-based)	0.014301	0.942984	0.915619
Method 2 (Corrected Weighted)	0.011331	0.945652	0.919694
Method 3 (Entropy-based Weighted)	0.029760	0.846670	0.799234
Method 4 (Score-Proportional Dynamic)	0.010431	0.949523	0.925804

Conclusion

Score-Proportional Dynamic Fusion is the most effective method for combining text and visual emotion scores, providing the closest match to Gemini's combined scores.