Mathematical Foundation of the Ensemble Model Construction

This section describes the ensemble methodology for Facial Emotion Recognition (FER) using three deep learning models, each providing confidence scores. The ensemble leverages a rank-based fusion strategy with non-linear functions (exponential and hyperbolic tangent) to generate fuzzy ranks, enhancing the robustness and accuracy of emotion classification. Here's a breakdown of each step:

1. Confidence Scores of Base Models

Each base model i (i.e., DenseNet169, EfficientNetB7, InceptionV3) provides confidence scores for C emotion classes. Let the confidence scores from each base model i be:

$$Pr_{i1}, Pr_{i2}, Pr_{i3}, \ldots, Pr_{iC}$$

where $i \in \{1, 2, 3\}$.

Each model's confidence scores sum up to 1, ensuring a valid probability distribution:

$$\sum_{k=1}^{C} Pr_{ik} = 1, \forall i \in \{1, 2, 3\}$$

2. Non-linear Fuzzy Rank Transformation

To transform the confidence scores into ranks, two non-linear functions, exponential and hyperbolic tangent, are applied:

$$y = 1 - \exp\left(-\frac{(x-1)^2}{2}\right)$$

$$y = 1 - \tanh\left(\frac{(x-1)^2}{2}\right)$$

Here, x represents the confidence score for a class, and these functions transform scores into fuzzy ranks, adding non-linearity that enhances discrimination among classes.

3. Applying Non-linear Functions to Confidence Scores

The confidence scores Pr_{ik} are passed through each function to obtain fuzzy ranks:

$$\text{Rank}_{i1_k} = 1 - \tanh\left(\frac{(Pr_{ik} - 1)^2}{2}\right)$$

$$\text{Rank}_{i2_k} = 1 - \exp\left(-\frac{(Pr_{ik} - 1)^2}{2}\right)$$

Significance of Each Function:

- **Exponential function:** Has a downward concavity, creating a rank output that approaches 1 as probability approaches 1.
- **Hyperbolic tangent function:** Has an upward concavity in the [0,1] domain, resulting in rank values that increase as probability moves closer to 0.

4. Fused Rank Scores Calculation

The fuzzy ranks from both functions are combined to calculate a fused rank score for each class:

$$RS_{ik} = \text{Rank}_{i1_k} \times \text{Rank}_{i2_k}$$

Here, $\textbf{Rank}_{i1_k}$ offers a reward for correct classifications (higher scores for probable classes), while $\textbf{Rank}_{i2_k}$ captures the divergence, ensuring balanced influence on the final rank.

5. Final Fused Score Calculation

The fused scores RS_{ik} are aggregated across all models to produce a final score tuple:

$$FS_k = \sum_{i=1}^{L} RS_{ik}, \forall k = 1, 2, 3, \ldots, C$$

This step aggregates the rank information from each base model, capturing the degree of confidence towards each class.

6. Emotion Class Selection

The class with the minimum fused score FS_k is selected as the predicted emotion class:

$$\text{Emotion class (I)} = \min_{\forall k} FS_k$$

This approach helps in selecting the most probable emotion class, based on the fused confidence across models.

Conclusion

This ensemble strategy, with its rank-based fusion approach, provides a unique method to combine the strengths of different models, achieving a robust FER system suited for real-world scenarios. Each step contributes to the system's accuracy by fine-tuning the handling of confidence scores and enhancing the model's interpretability.