Name: Syed Rafay Hassan
Roll No: 22i-1955

1. Network Details

We tested two deep learning models for emotion recognition: **ResNet50** and **EfficientNetB0**.

- Both models were first trained on ImageNet and then used as the base.

- On top of them, we added two heads:

    - **Classification head**: Dropout → Dense layer (256 units, ReLU) → Dense (8 classes, Softmax).

    - **Regression head**: Dropout → Dense layer (128 units, ReLU) → Dense (2 values, Tanh).

- **Optimizer:** Adam (learning rate = 0.0001).

- **Losses:** Categorical cross-entropy (for classification) and Mean Squared Error (for valence/arousal regression). Both were given equal weight.

- **Training:** 12 epochs, batch size = 32, input size = 224×224.

2. Dataset Splits

- Dataset: **DL_Assignment1_Dataset**

- **Task 1 (Classification):** Predict 8 emotion categories.

- **Task 2 (Regression):** Predict valence and arousal values.

- **Splits:** 85% training and 15% validation. The split was stratified to keep class balance.

3. Training Curves

- **ResNet50**: Accuracy increased slowly and stabilized around **40% validation accuracy**.

- **EfficientNetB0**: Learned faster but plateaued at a lower **28% validation accuracy**.

- Both models showed steady loss reduction without major overfitting.

4. Results on Validation Set

Classification Results

Model	Accuracy	F1-Macro	Mean AP
ResNet50	**0.408**	**0.405**	0.396
EfficientNetB0	0.278	0.259	0.335

- ResNet50 performed much better in classification, especially in terms of F1-score.
- Confusion matrices showed that EfficientNetB0 often confused high-frequency classes.

Regression Results

Model	RMSE (Valence)	RMSE (Arousal)	Corr (Valence)	Corr (Arousal)	CCC (Valence)	CCC (Arousal)	SAGR (Valence)	SAGR (Arousal)
ResNet50	**0.425**	**0.361**	**0.484**	**0.388**	**0.434**	**0.320**	**0.743**	**0.778**
EfficientNetB0	0.429	0.384	0.441	0.244	0.323	0.187	0.712	0.765

- ResNet50 gave better results for both valence and arousal across all metrics.
- EfficientNetB0 performed weaker, especially in correlation and CCC scores.

5. Comparison and Observations

- **ResNet50 was the best overall model.** It achieved higher accuracy, F1-score, and better regression results compared to EfficientNetB0.
- **EfficientNetB0 underperformed** and could not match ResNet50's balance between classification and regression.
- The results suggest that **ResNet50 is more reliable for emotion recognition tasks** on this dataset.