**ResEmoteNet Paper Summary**

1. **ResEmoteNet** – integrates Convolutional Neural Networks, Residual connections, and the Squeeze and Excitation network to effectively capture facial emotions.
   1. The model was evaluated using four open-source databases
      1. FER2013 - 79.79% accuracy
      2. RAF-DB – 94.76% accuracy
      3. AffectNet-7 – 72.39% accuracy
      4. ExpW (Expressions in the Wild) – 75.67% accuracy
2. Architecture:
   1. CNN composed of **3 convolutional layers** for hierarchical feature extraction. Each layer is followed by **batch normalization** and **max-pooling**.
   2. Squeeze-and-excitation block like the one used in EmoNeXt, with the squeezed output being processed by two FC layers; a dimensionality-reduction layer followed by a dimensionality-expansion layer, with a Rectified Linear Unit (ReLU) activation in between.
   3. Residual Network with 3 residual blocks used to prevent vanishing gradients and degradation in deep networks.
   4. Adaptive Average Pooling layer used as it enables the aggregation of input information into a constant output size, regardless of the original input dimensions. AAP adjusts kernel size and stride to reach a specific output size, instead of reducing spatial dimensions like traditional pooling methods. It ensures consistent output dimensions in various datasets and layers.
3. Training:
   1. Batch Size: 16
   2. Epochs: 80
   3. Learning Rate: 0.001 (1x10-3)
   4. Used Cross-Entropy Loss & Stochastic Gradient Descent optimiser
   5. Used class-weights
   6. Scheduler reduced learning rate by 0.1 when performance plateaued.