**Gesture Recognition – Deep learning**

**Problem Statement**

You want to develop a cool feature in the smart-TV that can **recognize five different gestures** performed by the user which will help users control the TV without using a remote.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Experiment** | **Model** | **Key Changes** | **Training Accuracy** | **Validation Accuracy** | **Model Stability** | **Decision + Explanation** |
| Experiment 1 | Conv3D (4 Conv3D + 2 Dense Layer) | Baseline (4 CNN layers, 2 Dense layers, Dropout 0.5) | Very low | Low | Fluctuating | Underfitting, significant difference between training and validation accuracy |
| Experiment 2 | Conv3D (4 Conv3D + 2 Dense Layer) | Image size: 60x60 | Improved | Higher than training | Unstable | Unusual higher validation accuracy |
| Experiment 3 | Conv3D (4 Conv3D + 2 Dense Layer) | Batch size: 8 | Improved | Higher than training | - | High training time |
| Experiment 4 | Conv3D (4 Conv3D + 2 Dense Layer) | Image size: 30x30 | Increased | Increased | Fluctuating | - |
| Experiment 5 | Conv3D (4 Conv3D + 2 Dense Layer) | Frames: 10 | Dropped | Dropped | Fluctuating | - |
| Experiment 6 | Conv3D (3 Conv3D + 2 Dense Layer) | 3 CNN layers | Reduced | - | - | No significant impact on training time |
| Experiment 7 | Conv3D (3 Conv3D + 2 Dense Layer) | Dropout 0.4, 128 neurons in Dense layers | Improved | Improved | - | Significant improvement in accuracy |
| Experiment 8 | Conv3D + LSTM (3 Conv3D + 1 LSTM Layer +2 Dense Layer) | Add LSTM layer | Slightly reduced | Reduced gap | Gradual | - |
| Experiment 9 | Conv3D + LSTM (3 Conv3D + 1 LSTM Layer +2 Dense Layer) | Dropout 0.3, L2 regularization | Significantly improved | Same | - | Accuracy discrepancy |
| Experiment 10 | Conv3D + LSTM (3 Conv3D + 1 LSTM Layer +2 Dense Layer) | Epochs: 40 | 99% | 90% | Stable | Gradual learning |