**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. We experimented with various models, adjusting hyperparameters to evaluate accuracy. Below are the detailed findings from our tests

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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 (31.5%) | Low (46.0%) | Fluctuating | Underfitting, significant difference between training and validation accuracy |
| Experiment 2 | Conv3D (4 Conv3D + 2 Dense Layer) | Image size: 60x60 | Improved  (86.5%) | Higher than training (89.0%) | Unstable | Unusual higher validation accuracy |
| Experiment 3 | Conv3D (4 Conv3D + 2 Dense Layer) | Batch size: 8 | Dropped  (67.0%) | Higher than training (86.0%) | Unstable | Training Accuracy has reduced, higher validation accuracy |
| Experiment 4 | Conv3D (4 Conv3D + 2 Dense Layer) | Image size: 30x30 | Increased (87.9%) | Increased (82.0%) | Fluctuating | Accuracy improved |
| Experiment 5 | Conv3D (4 Conv3D + 2 Dense Layer) | Frames: 10 | Dropped  (81.7%) | Higher than training (88.7%) | Stable compared to previous | Training Accuracy has reduced, higher validation accuracy |
| Experiment 6 | Conv3D (3 Conv3D + 2 Dense Layer) | 3 CNN layers | Dropped  (54.4%) | Dropped  (59.0%) | No significant impact on training time | Accuracy dropped significantly and |
| Experiment 7 | Conv3D (3 Conv3D + 2 Dense Layer) | Dropout 0.4, 128 neurons in Dense layers | Improved  (90.0%) | Improved  (88.0%) | Better model | Significant improvement in accuracy |
| Experiment 8 | Conv3D + LSTM (3 Conv3D + 1 LSTM Layer +2 Dense Layer) | Add LSTM layer | Improved  (95.1%) | Improved  (92.0%) | Better model | Gradual increase in efficiency |
| **Experiment 9 {Final Model}** | **Conv3D + LSTM (3 Conv3D + 1 LSTM Layer +2 Dense Layer)** | **Dropout 0.3, L2 regularization** | **Improved**  **(99.1%)** | **Improved**  **(96.0%)** | **Better model** | **Gradual increase in training efficiency** |
| Experiment 10 | Conv3D + LSTM (3 Conv3D + 1 LSTM Layer +2 Dense Layer) | Epochs: 40 | Slightly Reduced  (98.3%) | Dropped  (89.0%) | Unstable | Increasing Epochs has Negative impact of accuracy and performance |

**Experiment Summary and Conclusion**

After conducting 10 experiments, we conclude that Model 9 (Experiment 9) is the best model as it provides consistent training and validation accuracy

**Conclusion of Model 9:**

* Training Accuracy: 99.1%
* Validation Accuracy: 96.0%
* Model Structure:
  + Conv3D + LSTM (3 Conv3D Layers + 1 LSTM Layer + 2 Dense Layers)
  + Dropout: 0.3
  + L2 Regularization Factor: 0.005
  + Epochs: 25
  + Batch Size: 8
  + Num Frames: 15
  + Image Size: 60 x 60

**Model 9** has proven to be the most effective with consistent high accuracy for training and validation. Further increasing the number of epochs in Experiment 10 resulted in a decrease in accuracy, confirming Model 9's parameters as optimal.