NEURAL NETS GESTURE RECOGNITION CASE STUDY

Submission by – Kanika Raheja , Abhilash Dixit

**\*Please note the changes mentioned in the 4th column below are the only parameters that were changed rest were same as the previous step.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Conv3D** | **OOM Error** | **Created base Conv3D model with input & 3 hidden layers with 16,16,32,32 neurons & a dense layer (512 neurons) , filter size (3,3,3) which had 38.5 million trainable parameters while running with batch size of 20, 20 frames, 120x120 image size and 30 epochs** |
| **2** | **Conv3D** | **Accuracy: 0.25**  **Trainable parameters 38,588,021** | **Ran above model with 15 epochs.** |
| **3** | **Conv3D** | **Accuracy: 0.35 on training data, 0.38 on val data. Model is underfitting. Trainable parameters 1,812,341** | **To Reduce the number of trainable parameters- Reducing the number of neurons in each layer (8,16,32,64), adding another dense layer(with 216 neurons) in network, batch normalization with each layer & dropouts. Using Adam optimizer, epochs = 60** |
| **4** | **Conv3D** | **Accuracy: 0.44 on training data, 0.46 on val data. Model is underfitting.**  **Trainable parameters 1,386,357** | **Tried another model with SGD optimizer instead of Adam, but accuracy was very low & model was unstable.**  **So Reducing image size to 100x100 for next iteration.** |
| **5** | **Conv3D** | **Accuracy: 0.51 on training data, 0.63 on val data. Model is underfitting.**  **Trainable parameters**  **979,829** | **Reducing image size further to 80x80** |
| **6 (Best Accuracy received with Conv3D model)** | **Conv3D** | **Accuracy: 0.79 on training data, 0.70 on val data.**  **Trainable parameters**  **470,901** | **Changing the model- reducing dense layer neurons (256,128). Removing Dropouts & Adding Learning rate of 0.001. & Increasing Batch size to 30 from 20** |
| **7** | **Conv3D** | **Accuracy: 0.86 on training data, 0.63 on val data. Trainable parameters**  **470,901. Model is overfitting** | **Experimenting by reducing the learning rate to 0.0002 from 0.001** |
| **8** | **Conv3D** | **Accuracy: 0.68 on training data, 0.67 on val data. Trainable parameters**  **477,861** | **Experimenting by adding another hidden layer of 16 neurons with batch normalization, keeping learning rate 0.0002** |
| **9** | **Conv3D** | **Accuracy: 1.00 on training data, 0.83 on val data. Model is overfitting. Trainable parameters**  **470,901.** | **Removing added layer & setting LR to automatic as they are not making model stable. Reducing the batch size to 25 from 30.** |
| **10** | **CNN+RNN Model** | **Accuracy: .97 on training data, .72 on val data. Model is overfitting. . Trainable parameters**  **498,965.** | **Experimenting with Conv-GRU model. Image size 120,120, Frames 30, batch size 10. Dense layers (128,64). Filters (8,16,32,64), GRU layer with 128 neurons, Dropout value 0.25 & LR = 0.001. Epochs=30** |
| **11** | **CNN+RNN Model** | **Accuracy: .94 on training data, .75 on val data. Model is overfitting. Trainable parameters**  **302,357.** | **Reducing image size to 80x80**  **Increased the filter size and increased dropout t0 0.3** |
| **12** | **CNN+RNN Model** | **Accuracy: .86 on training data, .77 on val data. Not overfitting a lot but accuracy can be increased** | **Increasing neurons in each layer (16,32,64,128), increased dense layer neurons (256,128), increased batch size 20** |
| **12** | **CNN+RNN Model** | **Accuracy: 1 on training data, .74 on val data. Model overfitting again.**  **Trainable parameters**  **498,965.** | **Increased the dropouts and reduces learning rate from .001 to .0002** |
| **13** | **Conv2d with GRU – Transfer Learning** | **Accuracy: .98 on training data, .87 on val data.**  **Trainable parameters**  **16,421,701** | **Experimenting with transfer learning using VGG16 Architecture &Imagenet weights. Image size 100x100, GRU layer with 512 neurons & Dense layer with 256 neurons. LR = 0.0002, Frames=20, Batch Size =25, Epochs=35** |
| **14** | **Conv2d with GRU – Transfer Learning** | **Accuracy: .99 on training data, .89 on val data.** | **Reducing image size to 80x80** |
| **15 (Best Model with Transfer Learning)** | **Conv2d with GRU – Transfer Learning** | **Accuracy: .95 on training data, .94 on val data.**  **Trainable parameters**  **4,226,373** | **Experimenting with transfer learning using MobileNet Architecture &Imagenet weights. Image size 100x100, GRU layer with 256 neurons & Dense layer with 128 neurons & Dropout 0.25. LR = Auto, Frames=20, Batch Size =15, Epochs=40** |
| **Final Model** | **Conv2d with GRU – Transfer Learning** | **Accuracy: .95 on training data, .94 on val data.**  **Trainable parameters**  **4,226,373** | **Experimenting with transfer learning using MobileNet Architecture &Imagenet weights. Image size 100x100, GRU layer with 256 neurons & Dense layer with 128 neurons & Dropout 0.25. LR = Auto, Frames=20, Batch Size =15, Epochs=40** |