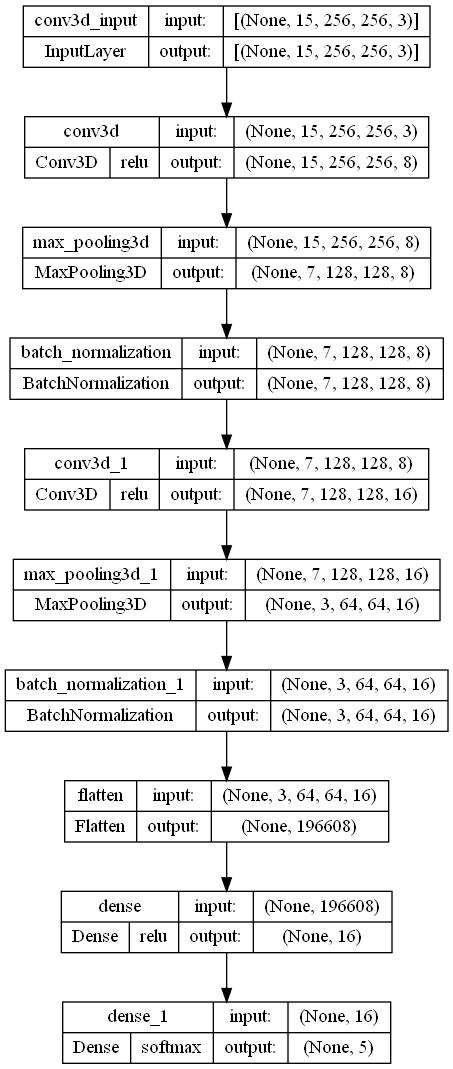
Gesture recognition architectural analysis and evaluation

Experimentation 1



Architecture type-: CNN3D

3D Convolution layer-: 3

Batch size -: 50

Kernel shape -: (3,3,3)

Pooling shape -: (2,2,2)

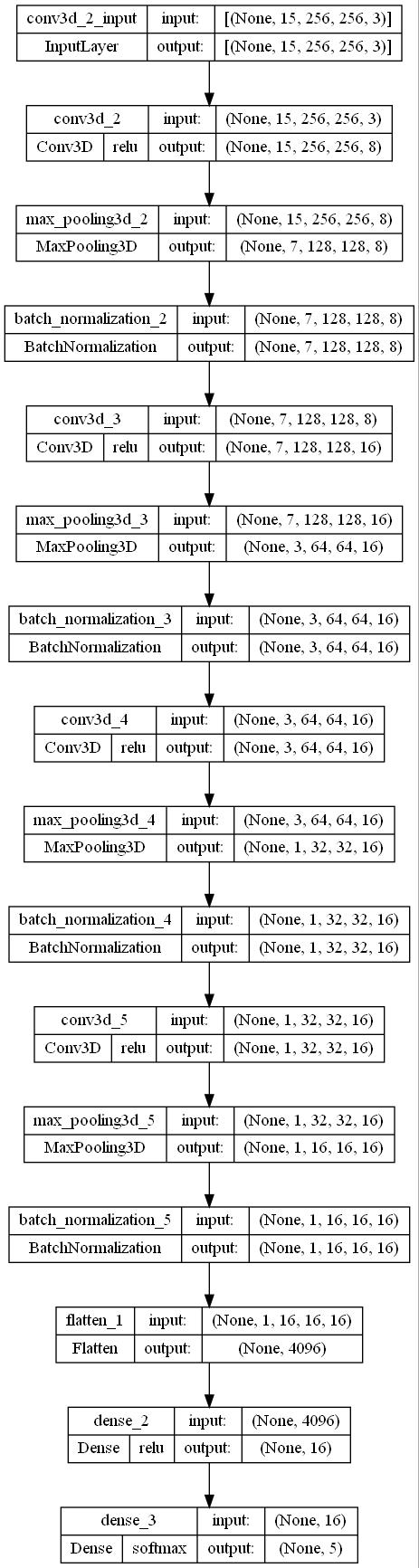
Highest Validation categorical accuracy -: 33% at 11/25 epoch

Status-: Underfitting (Highest training categorical accuracy -: 25% at 1/25 epoch)

Causation-: Lesser number of convolution layers, unable to learn the pattern sequentially.

Approached solution -: Building a deeper convolutional 3d network

Experimentation 2



Architecture type-: CNN3D

3D Convolution layer-: 5

Batch size -: 50

Kernel shape -: (3,3,3)

Pooling shape -: (2,2,2)

Highest Validation categorical accuracy -: 52% at 24/25 epochs

Status-: Overfitting (Highest training categorical accuracy -: 100% at multiple epochs)

Causation-: No regularization.

Approached solution -: Building a regularized convolutional 3d network

Experimentation 3

Architecture type-: CNN3D

3D Convolution layer-: 7

Batch size -: 50

Kernel shape -: (3,3,3)

Pooling shape -: (2,2,2)

Highest Validation categorical accuracy -: NONE

Status-: NONE

Causation-: ResourceExhaustedError

Approached solution -: Decreasing batch size to be able to fit images in GPU.

Experimentation 4

Architecture type-: CNN3D

3D Convolution layer-: 7

Batch size -: 32

Kernel shape -: (3,3,3)

Pooling shape -: (2,2,2)

Highest Validation categorical accuracy -: 32% at 19/20 epochs

Status-: underfitting (Highest training categorical accuracy -: 42.87% at 18th epoch)

Causation-: Might be because of lesser batch size model is unable to generalize the pattern.

Approached solution -: Approached solution should be down scaling the image even further, which can enable us to use a higher batch size as our GPU memory is limited, but we will keep this solution for later, first we will Experiment with CNN2D +RNN network to find out whether that perform better at training and validation accuracy.

Experimentation 5

Architecture type-: CNN2D (time distributed) +RNN

Layers-: 6 convolution 2d +1 LSTM layer

Batch size -: 32

Kernel shape -: (3,3)

Pooling shape -: (2,2)

Highest Validation categorical accuracy -: 52% at 19/20 epochs

Status-: Overfitting (Highest training categorical accuracy -: 84.62% at multiple epochs)

Causation-: Might be because of lesser batch size model is unable to generalize the pattern. Or the network architecture is not sufficiently deeper to learn the sequential pattern.

Approached solution -: use transfer learning for reusing the trained weights and biases from resnet 50, while allowing the top 10 layers of neurons to train its parameters for leering our specified purpose

Experimentation 6

Architecture type-: Resnet50 +RNN

Layer-: 40 layers of Resnet50 + 10 trainable layers of Resnet50 +1 LSTM layer

Batch size -: 32

Kernel shape -: Controlled by Resnet50 architecture

Pooling shape -: Controlled by Resnet50 architecture

Highest Validation categorical accuracy -: None

Status-: Overfitting (Highest training categorical accuracy -: None

Causation-: ResourceExhaustedError

Approached solution: - Decreasing batch size to be able to fit images in GPU.

Experimentation 7

Architecture type-: Resnet50 +RNN

Layer-: 40 layers of Resnet50 + 10 trainable layers of Resnet50 +1 LSTM layer

Batch size -: 10

Kernel shape -: Controlled by Resnet50 architecture

Pooling shape -: Controlled by Resnet50 architecture

Highest Validation categorical accuracy -: 27% at 24/25 epoch

Status-: Underfitting (Highest training categorical accuracy -: 21.42%

Causation-: Might be because of lesser batch size model is unable to generalize the pattern.

Approached solution: - Down scaling the image even further while generating the data batches, which can enable us to use a higher batch size as our GPU memory is limited.

Experimentation 8

Architecture type-: Resnet50 +RNN

Layer-: 40 layers of Resnet50 + 10 trainable layers of Resnet50 +1 LSTM layer

Batch size -: 20

Kernel shape -: Controlled by Resnet50 architecture

Pooling shape -: Controlled by Resnet50 architecture

Highest Validation categorical accuracy -: 87% at 18/25 epoch

Status-: Good fit (Highest training categorical accuracy -: 100% 24/25)

Causation-: Changed the resizing size to (75,75) which lead to lesser usage of GPU for storing image for each batch which enable us to give data in higher batch size, and we changed the cropping to none as, it seemed like cropping is losing necessary data from image.