**Selecting 2D+GRU over 3D cnn due to following reasons:-**

**1. Number of parameters lesser for 2D+GRU  
for 3D CNN  
(**Total params: 1,805,797

Trainable params: 1,805,413

Non-trainable params: 384)

**for 2D+GRU**

Total params: 475,521

Trainable params: 475,329

Non-trainable params: 192

**2. Low categorical loss value  
3. Comparable accuracy in both cases however 2D+GRU required more number of epochs  
4. Please refer to the following tables for the results:-**

**3D CNN model building**

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Starting with large model on sample data**  **No. of image in input seq = 18**  **Conv3D**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Epochs -20**  **(following network topology of well established architectures in field of image recognition)** | **Model overfits on small data and 20 epochs**  **Model train\_ accuracy ~ 97%**  **~ loss – 6.87**  **Model val\_accuracy ~ 81%** | **Training on the complete data to check the model performance** |
| **2** | **No. of image in input seq = 18**  **Conv3D**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Epochs -50** | **Training accuracy drops to ~73%**  **Validation accuracy drops to ~68%**  **~loss 5.45** | **Decreasing batch size to 16** |
| **3** | **No. of image in input seq = 18**  **Conv3D**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Epochs -50**  **Batch size - 16** | **Training accuracy ~73%**  **Validation accuracy ~71%**  **Loss ~4.7** | **Increase number of images in input sequence** |
| **4** | **No. of image in input seq = 30**  **Conv3D**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Epochs -50**  **Batch size - 16** | **Training accuracy ~93%**  **Validation ~75%** | **Adding dropout at last hidden layer as the model overfits** |
| **5** | **No. of image in input seq = 30**  **Conv3D**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Dropout(0.25)**  **Epochs -50**  **Batch size – 16** | **Training accuracy ~83.3**  **Validation accuracy ~82.13**  **Loss ~4.5** | **Adding Dropout at layer 3** |
| **6** | **No. of image in input seq = 30**  **Conv3D**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Dropout(0.25)**  **Layer4 -64**  **Dropout(0.25)**  **Epochs -50**  **Batch size – 16** | **Best Training and validation accuracy ~82.5** |  |

**Conv 2D + GRU**

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Conv2D + GRU**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Dense-128**  **GRU-18**  **Epochs -20**  **Batch size = 20**  **(following network topology of well established architectures in field of image recognition)** | **Model overfits on small data and 20 epochs** | **Checking accuracy on entire training data** |
| **2** | **Conv2D + GRU**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Dense-128**  **GRU -18**  **Epochs -20**  **Batch size = 20**  **(following network topology of well established architectures in field of image recognition)** | **Model underfits on the training data** | **Removed image cropping. Resized images to 120\*120** |
| **3** | **Conv2D + GRU**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Dense-128**  **GRU-18**  **Epochs -50**  **Batch size = 20** | **Training accuracy ~70%**  **Validation accuracy ~79%** | **Decreasing batch size to 16** |
| **4** | **Conv2D + GRU**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Dense-128**  **GRU-18**  **Epochs -50**  **Batch size = 20** | **Training accuracy ~56%**  **Validation**  **Accuracy ~60** | **Keeping batch size = 20**  **Adding more samples in the input sequence as the model underfits** |
| **5** | **Conv2D + GRU**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Dense -128**  **GRU-18**  **Epochs -50**  **Batch size = 20** | **Model overfits**  **Training accuracy ~ 85**  **Validation accuracy ~69** | **Adding dropout since model overfits** |
| **7** | **Conv2D + GRU**  **Layer1 -32**  **Layer2 -32**  **Layer3- 64**  **Layer4 -64**  **Dense -128**  **GRU-18**  **Epochs -100**  **Batch size = 20** | **Best model accuracy**  **Training accuracy ~81**  **Validation accuracy ~78**  **Loss ~0.56** |  |

**Note:- The file paths have been reverted back to those originally provided in the starter code**

train\_path = '/notebooks/storage/Final\_data/Collated\_training/train'

val\_path = '/notebooks/storage/Final\_data/Collated\_training/val'