Model architecture:

- Resnet 34 with basic block
- Modification: the first conv2d layer's kernal size is 3, stride is 1, and padding is 1; added embedding to the forward.

Loss function

- CrossEntropyLoss

Hyper parameters

learning_rate = 0.1

weight_decay = 5e-5

batch_size = 128

n_epochs = 70

Data Augmentation

```
data_transform = transforms.Compose([
  transforms.RandomHorizontalFlip(),
  transforms.RandomRotation(20),
  transforms.ColorJitter(brightness=0.5, contrast=0.5, saturation=0.2),
  transforms.ToTensor()])
```

Optimiser and Scheduler

```
optimizer = torch.optim.SGD(network.parameters(), lr=learning_rate,
weight_decay=weight_decay, momentum=0.9)
scheduler = torch.optim.lr_scheduler.ReduceLROnPlateau(optimizer, mode="min", factor=0.8,
patience=3, verbose=True)
```

Steps

Classification

1. Run all the cells before the cell that downloads test data.
2. Finding the model with the highest accuracy and uncomment the line to load that model (this line is indicated in the file).
3. Use that model to run all the cells before the verification part.

Verification

1. Run the rest cells.