CS5480
Final Presentation

CS21BTECH11061
Vishal Vijay Devadiga

Challenge

- Multiclass Image Classification
  - 50 classes of different types: Animals, Fishes, Human Activities, Insects, Food .etc.
  - 65000 training images
  - 38366 test images

Final Submission

- A neural net that uses **Squeezenet**(self-implemented) as one of the layers.
- Submission:
  - Data Modification:
    - Resizing images to 128x128
    - Randomly flipping images vertically and horizontally
    - Grayscaling the images
  - Model:
    - Convolutional Layer with 1x1x3 kernels
    - Pooling Layer (2x2)
    - Squeezenet
    - GAP layer
    - FC layer (10 to 50)

Final Submission

- ○ Model Parameters:
    - ■ Number of Epochs = 25
    - ■ Loss Function : Cross Entropy Loss
    - ■ Batch Size = 32
    - ■ Optimizer: Adam with learning rate 0.0005

Public Leaderboard Accuracy: 25.786

Private Leaderboard Accuracy: 16.082

Preprocessing

- Most images had either 512 height/width.

  - Tried both 256x256, 224x224 , 128x128 with varied results

- Tried both grayscale and colored images

  - In hindsight, should have gone with colored images

- Randomly flipping images vertically and horizontally to increase the variability in images and dampen over-learning

- Normalized the images(values of tensor) with a mean and standard deviation

Neural Net

- Tried a small CNN with 3-4 layers without preprocessing to test the waters and got a public accuracy of 25.626

- Tried ViT submitted in Assignment-3, but the model was not able to learn the weights required.
  - Most probably due to the positional encoding

- Tried a large CNN but kept getting lower accuracy compared to the first attempt

- Tried squeezenet, and kept using it for the remainder of submissions

Hyperparameters

- Optimizers, Batch Size and Learning Rate:
    - Tried both SGD(batch = 1) and Adam
    - With Adam, tried learning rates varying from 0.00001 to 0.1
    - Tried batch sizes varying from 1 to 128
- Kept using validation data, until submission, where entire data was used to learn the weights

Conclusion/Things that could have been done differently

- Could have tried more known models instead of sticking to squeezenet

  - Using ResNet gave around 50% accuracy without any preprocessing

- Could have tried more robust, simpler CNN rather than using a known model

- Could have tried manipulating the parameters of ViT to fix the issue of zero learning

- Could have gone through data more extensively and tried to extract some features

  - Should have tried 64x64 sized images since most images in the test folder are of that height/width.

Thank You