Long Nguyen

1001705873

Basic CNN:

Architecture: The Basic CNN follows a standard CNN design that consists of convolutional layers, activation functions, pooling layers, and fully connected layers for classification.

Network Design:

1. Feature Extraction:
    - Two convolutional layers with 3×3 filters to extract spatial patterns.
    - ReLU activation is applied after each convolution to introduce non-linearity.
    - Max pooling layers reduce spatial dimensions, improving efficiency and reducing overfitting.
2. Classification:
    - The extracted feature maps are flattened into a 1D vector.
    - Two fully connected (dense) layers translate feature representations into class scores.
    - The final layer outputs logits for classification.

Training loss: 1.87083315849304 - 0.0590602494776248

Validation loss: 2.1952714920043945

Test accuracy: 0.6119745373725891

All Convolutional Model:

Architecture: The All-CNN modifies the traditional CNN by removing fully connected layers and replacing them with 1×1 convolutional layers. This makes the network fully convolutional, keeping spatial information intact.

Network Design:

1. Feature Extraction:
    - Multiple convolutional layers replace both feature extraction and classification layers.

- 1×1 convolutional layers replace fully connected layers, reducing the number of parameters.
- Global Average Pooling (GAP) aggregates spatial information, reducing overfitting.

2. Classification:
   - Instead of flattening, a final 1×1 convolutional layer directly produces class scores.
   - The output is passed to a softmax activation, yielding class probabilities.

Training loss: 2.04505228996276 - 0.407043665647506

Validation loss: 1.012251853942871

Test accuracy: 0.7090445756912231

All-CNN w/o regularization vs All-CNN w/ regularization using dropout (p=0.3):

Compared to the results above:

Training loss: 2.18779134750366 - 0.682248413562774

Validation loss: 1.026788592338562

Test accuracy: 0.7329936027526855

With regularization the model is a little more accurate than w/o regularization.

Transfer Learning:

Model used: All-CNN w/ regularization using dropout (p=0.3)

Validation loss: 0.525752604007721

Test accuracy: 0.8414000272750854