Assignment 06 : Visual Object Classification UTA027 : Artificial Intelligence

Raghav B. Venkataramaiyer

Mar '25

1 Dataset Preparation

- Use torchvision.datasets.VOCDetection to load the Pascal VOC 2007 or 2012 dataset.
- Extract image-level labels from annotations.
- Convert the dataset into PyTorch tensors.
- Split the dataset into training and validation sets (e.g., 80% training, 20% validation).
- Resize and normalize the image data.

2 Data Visualization

- Display at least 5 sample images with their bounding boxes and labels.
- Plot a bar chart showing the frequency of each object class.
- Plot a pie chart of the top 5 most common classes.

3 Build the Classification Model

- Choose a pre-trained model from torchvision.models (e.g., ResNet, VGG).
- Replace the final layer with a new fully connected layer to output predictions for 20 classes.

• Use ReLU activation in hidden layers and softmax (implicitly handled by loss function) for the output.

4 Train the Model

- Define a suitable loss function (e.g., CrossEntropyLoss).
- Choose an optimizer (e.g., Adam or SGD).
- Train the model for a fixed number of epochs (e.g., 1020 epochs).
- Track and display training and validation loss per epoch.

5 Evaluate the Model

- Compute accuracy, precision, and recall on the validation set.
- Display a confusion matrix to visualize classification performance.