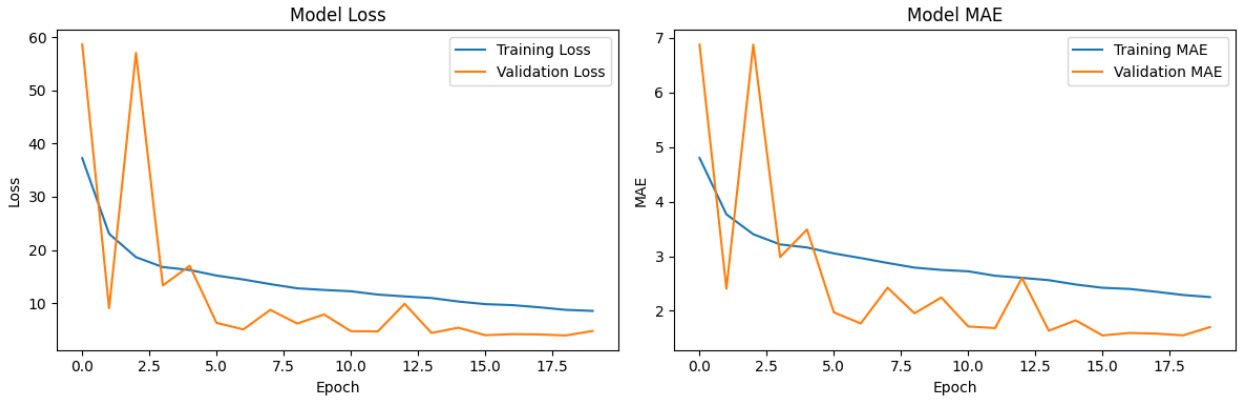
Training.ipynb

1. **Data Loading and Preparation**:
    - Combined datasets (data0.npy, data1.npy, data2.npy) and their corresponding labels (lab0.npy, lab1.npy, lab2.npy) into unified arrays.
    - Preprocessed the data:
        - Reshaped images to include a single channel dimension.
        - Normalized pixel values to fall within the range [0, 1].
2. **Model Design**:
    - Built a sequential CNN architecture with:
        - Three convolutional layers followed by batch normalization and pooling layers.
        - A flattening layer, followed by a dense layer with 64 neurons, a dropout layer for regularization, and a final dense output layer for sum prediction.
    - Used ReLU activation in the hidden layers.
3. **Compilation and Training**:
    - Compiled the model with:
        - Adam optimizer.
        - Mean Squared Error (MSE) as the loss function.
        - Mean Absolute Error (MAE) as the evaluation metric.
    - Trained the model on the preprocessed data for 20 epochs with a batch size of 32, reserving 20% of the data for validation.
4. **Evaluation**:
    - Visualized the training and validation loss and MAE over epochs.
    - Saved the trained model to digit_sum_model.keras for future inference.

Results:



Final Training Metrics:
Training Loss: 8.5387
Validation Loss: 4.7669
Training MAE: 2.2520
Validation MAE: 1.7024

Inference.ipynb

1. **Data Loading and Preprocessing**:
   - Loaded test data (data0.npy and lab0.npy).
   - Preprocessed images by reshaping them and normalizing pixel values.
2. **Model Loading**:
   - Loaded the trained model (digit_sum_model.keras) from the training phase.
3. **Prediction and Evaluation**:
   - Predicted the sum of digits for the test dataset.
   - Calculated:
     - Exact Match Accuracy: Proportion of predictions matching the true values exactly.
     - Accuracy within a tolerance of ±0.5: Proportion of predictions close to the true values.
4. **Visualization**:
   - Displayed sample test images with their true and predicted values.
   - Conducted a detailed analysis, reporting:
   - Total predictions, exact matches, and matches within the tolerance range.
   - Average absolute error.
5. **Test Metrics**:
   - Reported the test loss (MSE) and MAE, showcasing the model's ability to generalize.

Results:

Model Evaluation Results:
Exact Match Accuracy: 0.2171
Accuracy (within ±0.5): 0.2171
Test Loss (MSE): 3.4181
Test MAE: 1.4622

Detailed Prediction Analysis:
Total Predictions: 10000
Exact Matches: 2171 (21.71%)
Within ±0.5 tolerance: 2171 (21.71%)
Average Absolute Error: 1.462

Random Demo: