ASL to English Text Transcription

Overview

This code file implements a Convolutional Neural Network (CNN) model for American Sign Language (ASL) to English text transcription. It utilizes PyTorch and torchvision for deep learning functionalities. The code includes data preprocessing, model creation, training, and testing procedures.

Dependencies

Be sure you have the following libraries installed:

torch
torchvision
Scikit-learn

Configuration

data_path: Path to the folder containing subfolders for each class (ASL sign).
num_classes: Number of classes (adjust according to your ASL dataset).
Transformation pipeline (transform): Resize images, convert to tensor, and normalize.
Training hyperparameters: Batch size, learning rate, number of epochs.

Dataset Splitting

The dataset is split into training and testing sets using stratified sampling to maintain class distribution.

Model Architecture

The CNN model is defined with three convolutional layers, max-pooling, and fully connected layers. The output layer has a size corresponding to the number of classes.

Training

The model is trained using the Adam optimizer and CrossEntropyLoss. The training progress is printed, showing loss every 100 mini-batches.

Testing

The accuracy of the trained model is evaluated on the test dataset, providing insights into its performance.

Example Usage

The code includes an example with 10 training epochs and subsequent testing.