CptS 534 Neural Network Design and Application Project Proposal

Dheeraj Vurukuti (11810023)
Priyanka Ghosh Dastidar (11804812)

February 2023

Project Name: Human Activity Recognition using LSTM-CNN
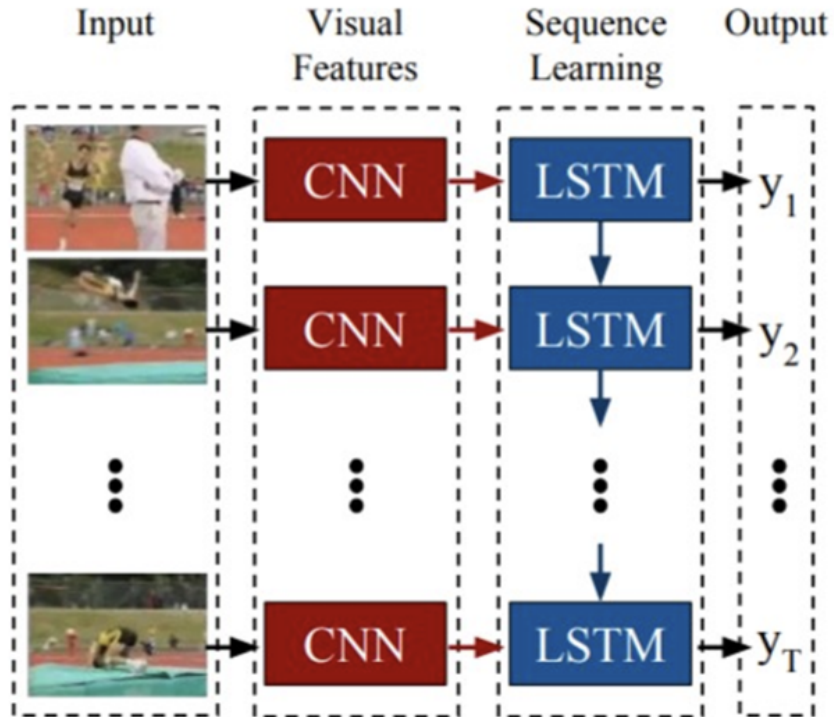
Abstract

Recent advancements in artificial intelligence (AI) have increased human curiosity about new research topics by allowing us to recognize objects, understand the world, analyze time series, and predict future sequences. AI researchers are becoming increasingly interested in neural networks, which have useful applications in speech recognition, language modeling, video processing, and time series analysis. One of the challenging questions in this fascinating area of AI that seeks solutions is the recognition of human behavior, also known as "human activity recognition" (HAR). A wide range of real-world applications are also covered by HAR, including those in the areas of security, gaming, personal fitness, and healthcare. The advancement of Human Computer Interaction (HCI) technology has become more popular for capturing behaviors using sensors like accelerometers, magnetometers, and gyroscopes. HAR can be accomplished with sensors, images, smartphones, or videos.

Introduction

In this project, we will discover how to use a Convolutional Neural Network along with a Long-Short Term Memory Network to accomplish human action recognition on videos. TensorFlow will be used twice, with two distinct structures and methods. Then we'll use the model that performs the best to make predictions about YouTube videos.

In the input sequence (video), a CNN will be used to extract spatial features at a specific time step, and an LSTM will be used to determine the temporal relationships between the frames.

Input   Visual Features   Sequence Learning   Output

CNN → LSTM → y_1

CNN → LSTM → y_2

⋮   ⋮   ⋮   ⋮

CNN → LSTM → y_T

Methodology

Here, we would use Convolutional Neural Networks (CNN) which are great for image data, and combine them with Long-Short Term Memory (LSTM) networks, which works great when working with sequence data.
Below are the steps we would perform in delivering the project:

- Step 1: Download the data and display it with its labels.

- Step 2 : Pre-process the dataset

- Step 3 : Divide the data into a train set and a test set

- Step 4: Use the ConvLSTM method to create and train the model. Along with it, draw the loss and accuracy curves for the model.

- Step 5: Implement the LRCN Method and create, compile, and train the model; and draw the loss and accuracy curves of the model.

- Step 6: Test the best-performing model on YouTube videos

Final Product

Trained model to gain higher accuracy in video classification and to perform human action recognition on videos by utilizing the temporal as well as spatial information of the data.