Video Capture and Streaming for Neural Network Training

Marco Infantado (infanm1@unlv.nevada.edu)| Mentor: Evangelos Yfantis
Department of Computer Science, University of Nevada Las Vegas, National Science Foundation

Introduction:

- Knowing how many pedestrians are waiting at an intersection can aid in managing traffic flow. If no pedestrian is currently waiting at an intersection, then the traffic light only needs to be controlled based on the flow of vehicles
- The system needs to be robust and intelligent enough to help pedestrians cross intersections quickly and safely no matter the weather conditions
- Pedestrians should be detectable even in poor lighting conditions, poor placement of camera, and poor video quality
- Captured video of crosswalks can be used to train a system with counting and detecting pedestrians using neural networks

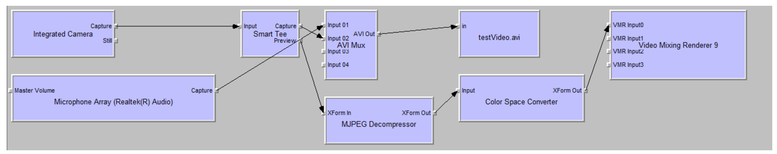
Purpose/Aim:

- The objective was to write a player for video capture-store-play for RGB cameras as well as multispectral LiDAR
- Captured data can be used to train and test NN capable of recognizing all the images in a frame and ordering them according to the distance from the capturing source and direction (path)

Methods:

- A video capture and playback software would also be created using WinAPI and DirectShow written in C++
- GraphEdit shows a simulation of how DirectShow filters and devices should be structured for video capture and preview. This simulation will then be converted to code (Figure 1)
- A version of the video player will be used for server-side purposes to store the captured footage and another version is used for client-side purposes to request and view footage stored on the server
- The server would be able to stream the live feed onto the client

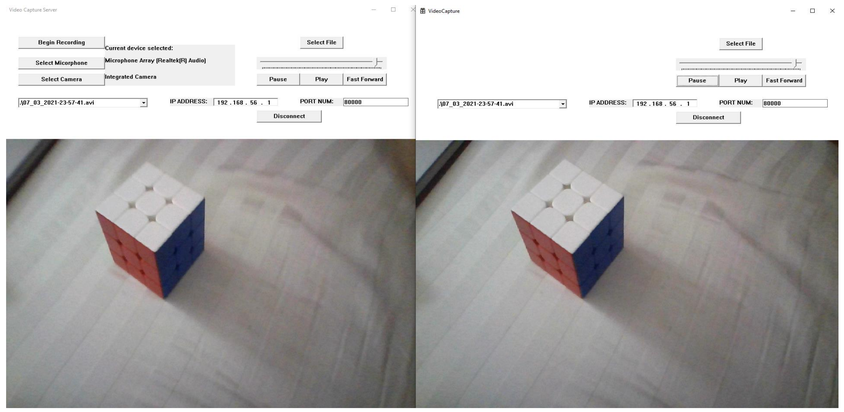
1) GraphEdit example filter graph for video capture and preview



Results:

- Using DirectShow, client will be able to view videos which have been previously recorded. The software can record continuously and store every 5-minute video into its own file (Figure 2)
- Files recorded are named based on time and date recorded
- User can choose which camera and microphone to use for recording
- A preview of the video is shown as it is recorded
- Video can be captured while another file is played simultaneously
- Videos are uncompressed and saved as Audio Video Interleave (.avi) files

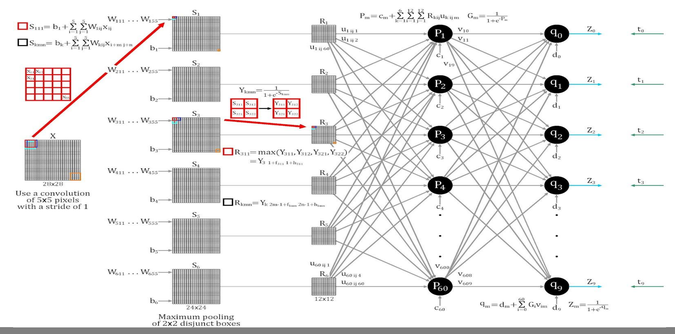
2) Video capture and playback software demonstrating transmission and receiving



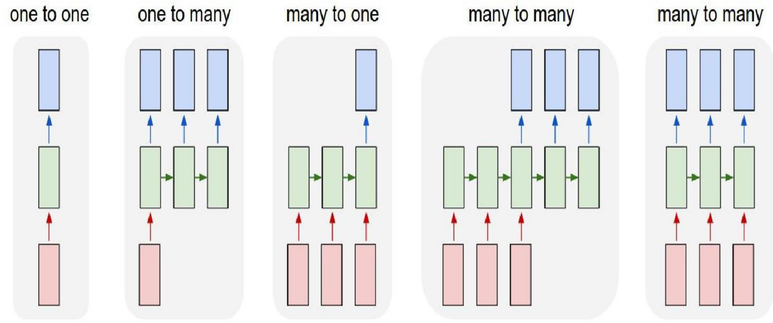
Conclusions & Future Research:

- The program should be converted to use Windows Media Foundation which was made to replace DirectShow
- The capture software would be used to record example footage of roads and pedestrians to train the neural network
- Captured video will be used to run the image recognition
- A convolutional neural network can be used for image classification and detecting certain objects in images (Figure 3)
- A recurrent neural network can be used to account for the motion between images in a video (Figure 4)

3) Convolutional neural network diagram



4) Recurrent neural network diagram



- The neural networks will be trained to classify or count the objects into video which could be used for counting pedestrians (Figure 5)

5) Example of expected results

Acknowledgements:

This project was funded by the National Science Foundation and is a part of their Research Experiences for Undergraduates program.

#1950872