VERSION 1.5

DATE 5/31/2024

NVIDIA JETSON-BASED OCR SYSTEM

SUMMER INTERNSHIP PROGRAM 2024

PRESENTED BY: ANTONIO GONZALEZ
A2E TECHNOLOGIES
10949 TECHNOLOGY PL SUITE B, SAN DIEGO, CA 92127

PROJECT VISION

This project aims to develop and implement an Optical Character Recognition (OCR) system by leveraging the capabilities of the NVIDIA Jetson Orin Nano Development Platform. The primary objective is to create a functional OCR application capable of accurately recognizing and identifying characters from a live video feed. The vision of the project is to harness the power of advanced neural networks and embedded GPU technology to deliver a robust, real-time OCR solution suitable for various real-world applications.

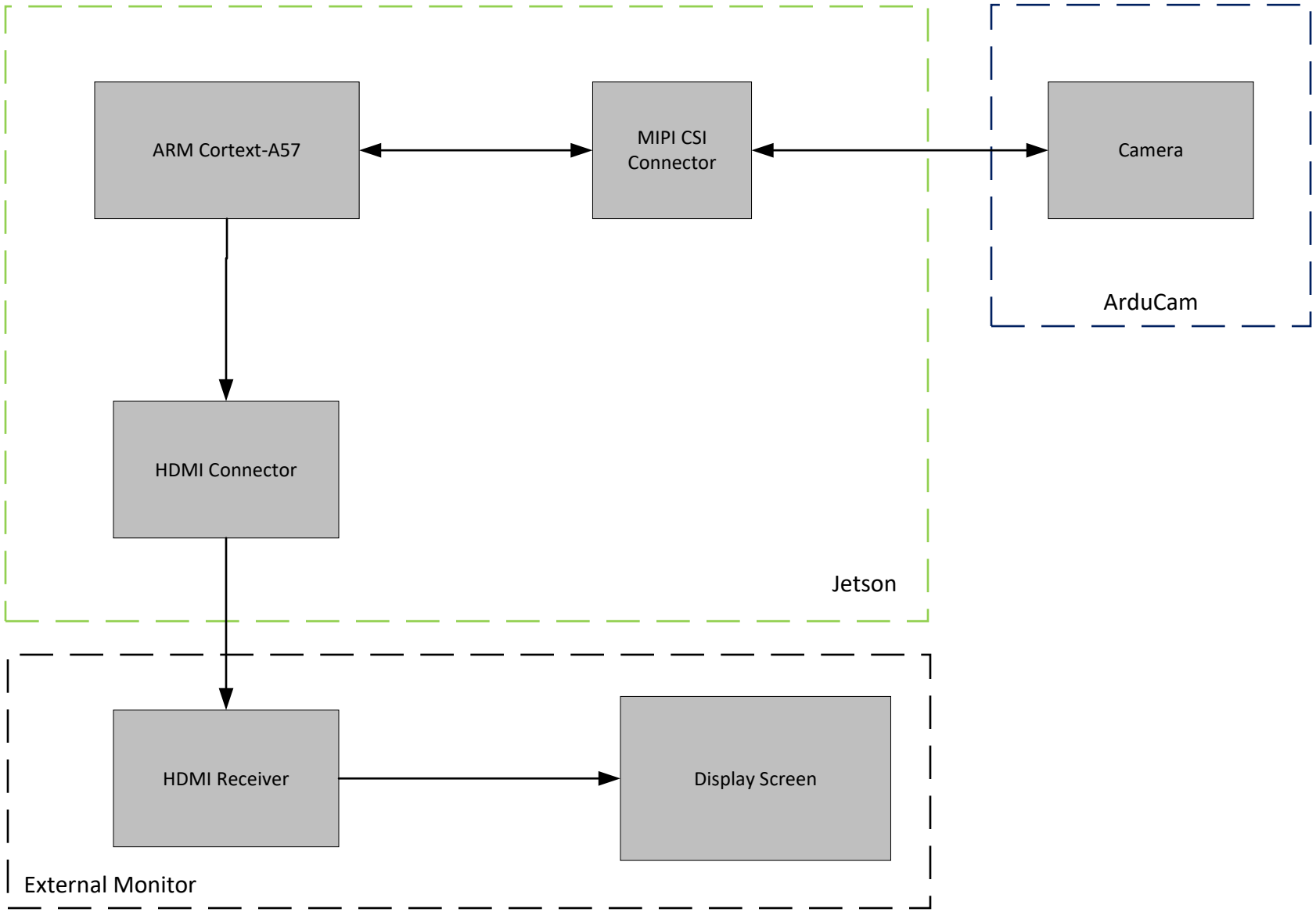
PROJECT DEVELOPMENT TIMELINE

The project would be split into the following phases:

1. **Research implementation and determine neural network architecture**
    a. Identify the most suitable neural network architecture for OCR tasks, considering factors such as accuracy, complexity, and computational efficiency. Review current literature and existing models to inform this decision.
    b. Examine performance between a classic N layer network and Convolutional Neural Networks (CNNs) for purpose of this project.
2. **Test camera hardware and determine capabilities**
    a. Connect the 12MP ArduCam camera module to the Jetson Nano and conduct a series of tests to assess its performance, resolution, and frame rate. Determine the optimal settings for capturing high-quality video suitable for the project.
    b. Evaluate character resolution by measuring pixel dimensions and distance from the sensor to the page. Calculate theoretical resolution and compare it with actual findings.
3. **Create a custom 3D-modeled camera stand**
    a. If time permits, design and create a 3D model of a stand tailored to maintain consistent camera distance and environmental conditions, ensuring reliable and reproducible results for image capture and analysis.
4. **Collect training data based on camera test findings**
    a. Gather a diverse dataset of labeled images reflecting the findings from camera tests, ensuring the dataset covers the ideal camera resolution and conditions.
5. **Refine neural network model using training data**
    a. Utilize TensorFlow to train and extract the weights of the neural network model, continuously evaluating and adjusting based on performance metrics derived from the collected training data.
6. **Implement initial C++ program on Jetson**
    a. Develop a C++ program to run inference on the Jetson, initially focusing on reading one still image to verify the functionality and accuracy of the neural network.

7. **Evaluate performance and optimize using CUDA**
    a. Convert the C++ program to CUDA to optimize neural network inference on the Jetson Nano, harnessing the parallel processing power of the GPU.
    b. Adapt the program to accommodate live video input, enabling real-time OCR system performance assessment and adjustment on the Jetson Nano.
8. **Integrate OCR model with GStreamer Pipeline**
    a. Develop a GStreamer pipeline to handle live video feed from the camera and integrate the OCR model into this pipeline.
9. **Expand program to include real-time results**
    a. Enhance the program by incorporating functionality to generate an output frame overlay, providing visual feedback and analysis results directly on the captured video stream
10. **Refine output and prepare final presentation**

Nvidia Jetson-Based OCR System Hardware Diagram



Nvidia Jetson-Based OCR System Software Diagram