Dwight Look College of
ENGINEERING
TEXAS A&M UNIVERSITY

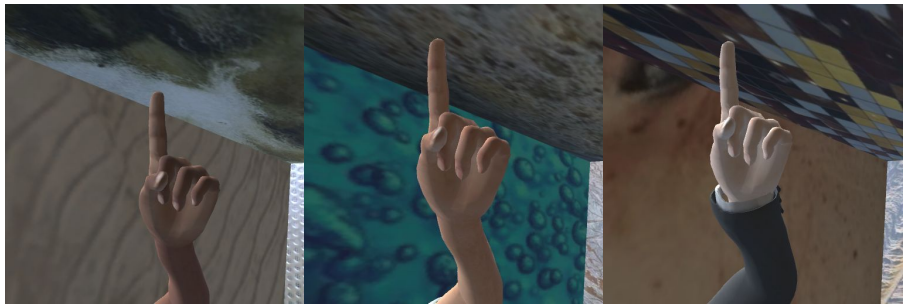
Team 72: Hand Gesture Recognition
Bi-Weekly Update 2

Samuel Oncken, Steven Claypool
Sponsor: Stavros Kalafatis
TA: Pranav Dhulipala

Project Summary

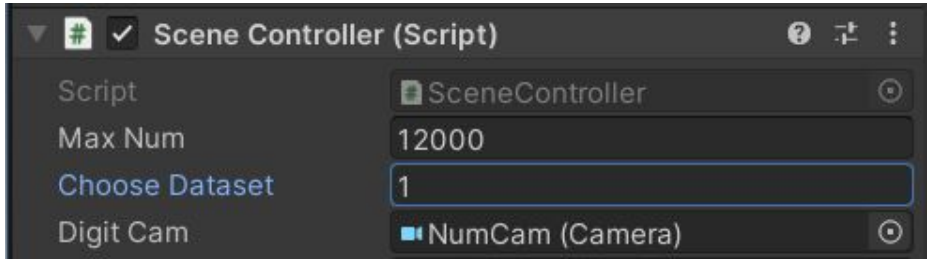
Problem Statement

- Collecting the large amounts of data required to properly train a gesture recognition NN is **time consuming**, **resource intensive**, and **difficult to scale**



Our Objective

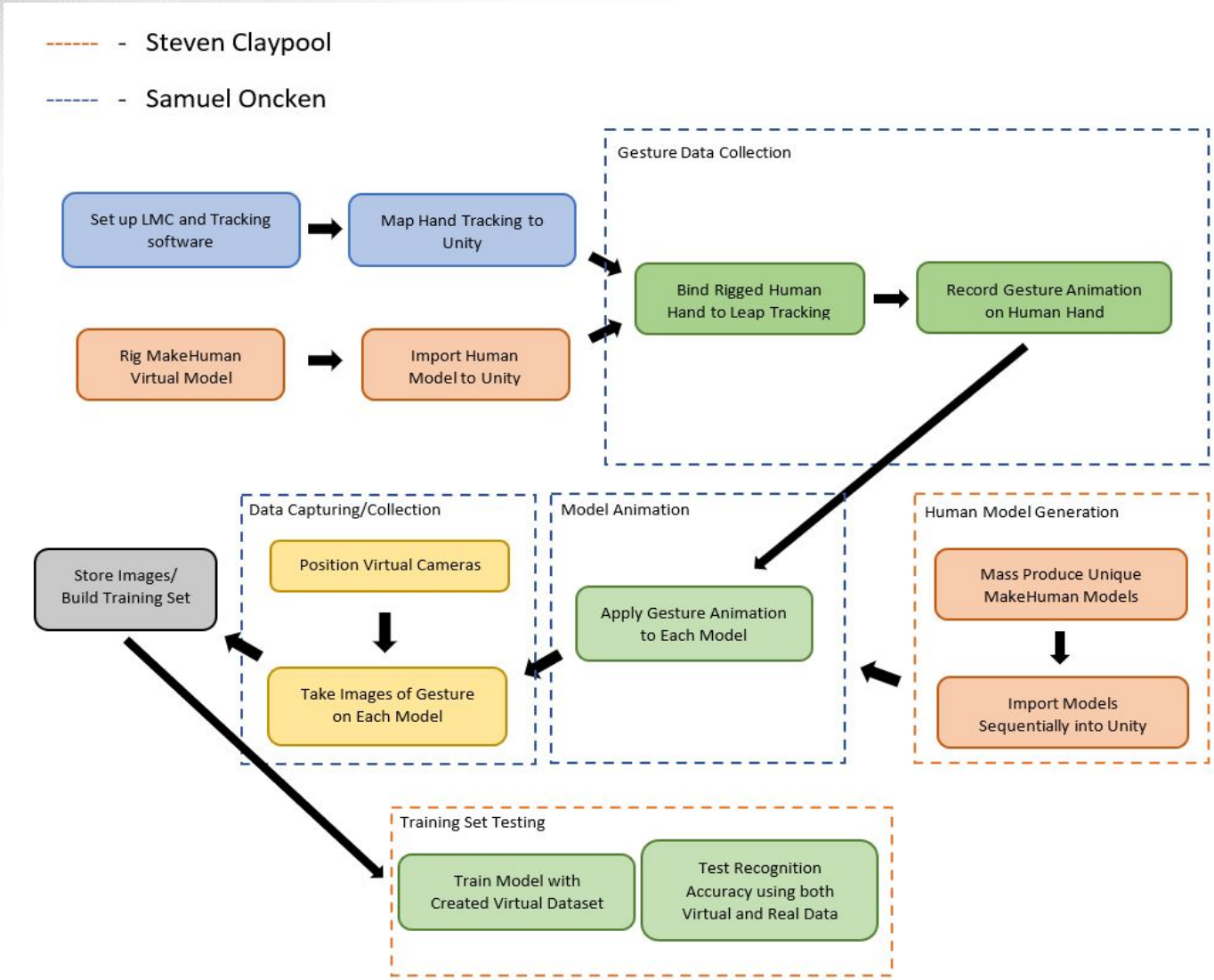
- Test viability of using **virtual data** to train a hand gesture recognition neural network
- Provide large amounts of **diverse data** (skin tones, hand sizes, accessories, etc.)
- Achieve **similar recognition accuracy** when tested against real, benchmark datasets



Why?

- Using a virtual environment, **scalability is easy** (# of images, # of participants, lighting conditions, background conditions, etc).

Project/Subsystem Overview



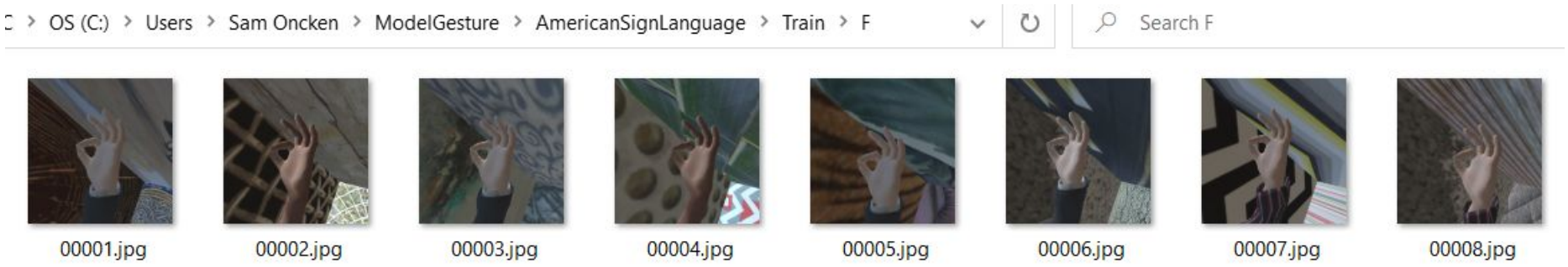
Project Timeline

Create Virtual Training Set Generation Unity Environment (Completed 11/20)	Generate Alphabet and Numbers Datasets – 12,000 Images/Gesture (Completed 12/12)	Complete Gesture Classification Analysis using Generated Datasets (To be complete by 03/31)	Complete Object Localization/ Detection Analysis using HANDS Generated Dataset (To be complete by 03/31)	Complete Robotic Application Testing/Training Virtual Environment (To be complete by 04/16)	Demo and Final Report (To be complete by 04/28)

Virtual Environment System Progress

Samuel Oncken

Accomplishments since 403 .5 hours of effort	Ongoing progress/problems and plans until the next presentation
- Plan for HANDs dataset application into environment for Object Localization work	- Waiting for ML analysis results to evaluate changes that might need to be made.



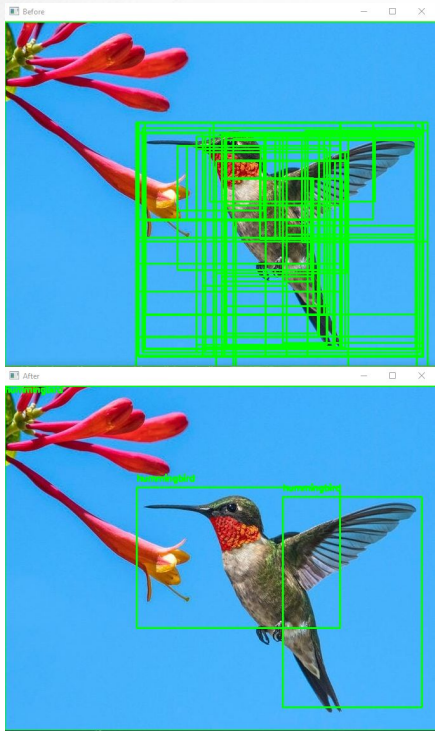
ML Analysis Progress

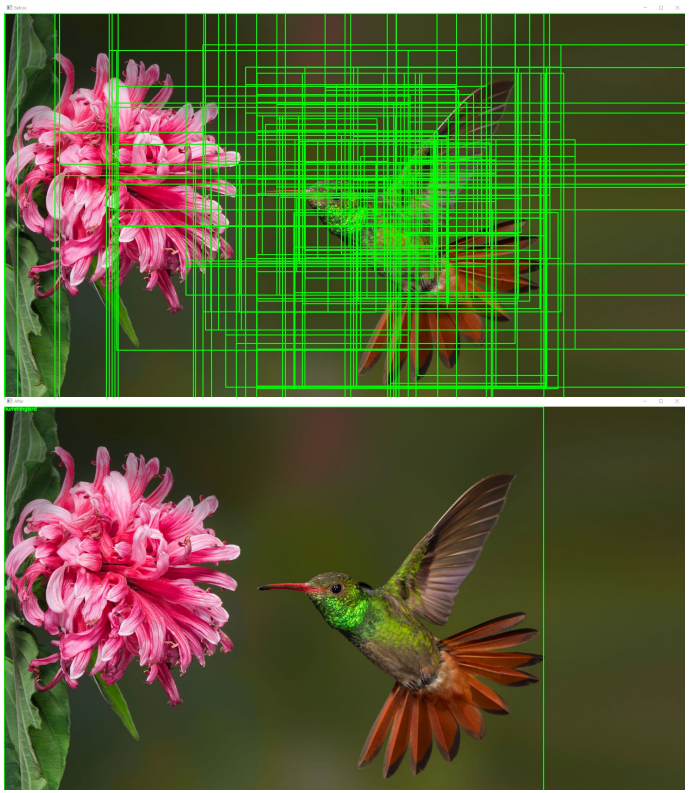
Pranav Dhulipala, Steven Claypool

Accomplishments since 403 8 hours of effort	Ongoing progress/problems and plans until the next presentation
- RCNN setup is done	- Code isn't setup for our use case - Adjusting RCNN for HANDS dataset - Add ground truth labels for accuracy measurements

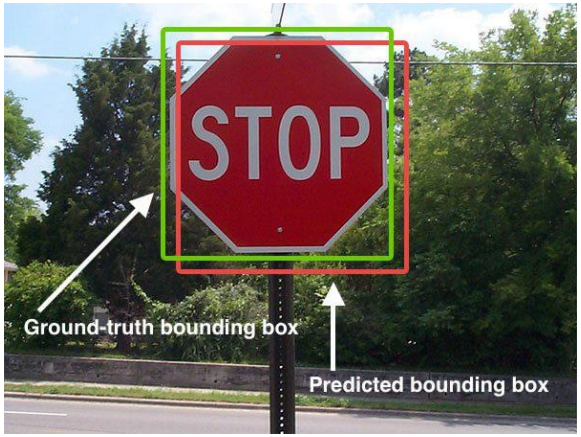
ML Analysis

Pranav Dhulipala, Steven Claypool





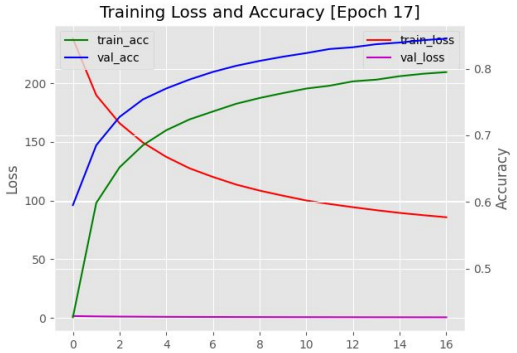
$$IOU = \frac{\text{area of overlap}}{\text{area of union}} =$$

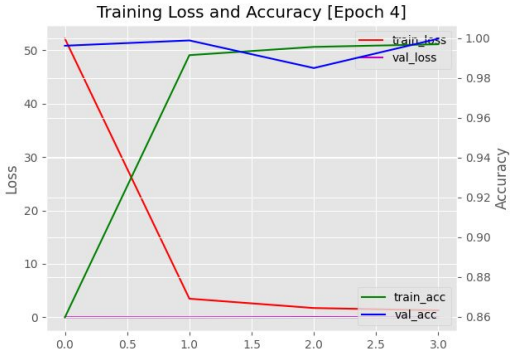


ML Analysis (cont.)

Pranav Dhulipala, Steven Claypool

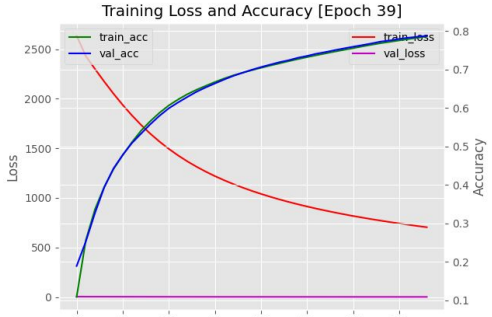
VGG16 Synthetic, Transfer (top) and scratch (bottom)





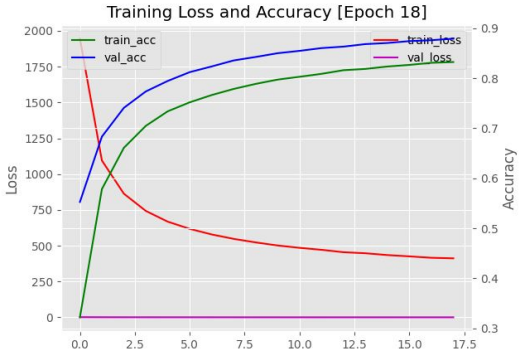
Alphabet Synthetic, Inception, Transfer, No Aug.

Alphabet Synthetic, ResNet50, Transfer, No Aug.



Alphabet Real, ResNet50, Transfer

Alphabet Synthetic, Inception, Transfer



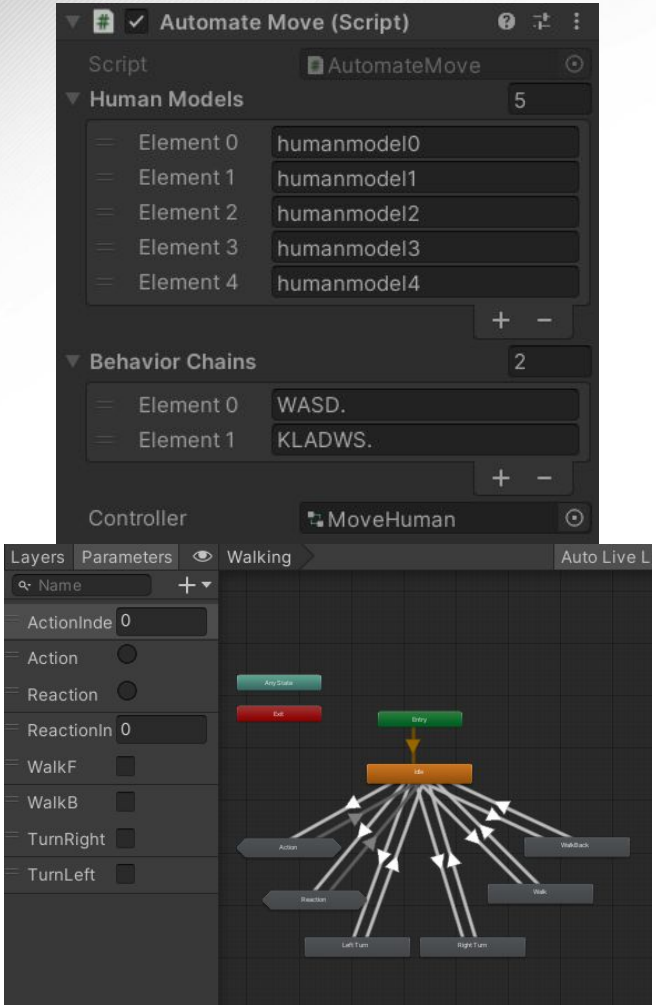
Robotic Application Extension Progress

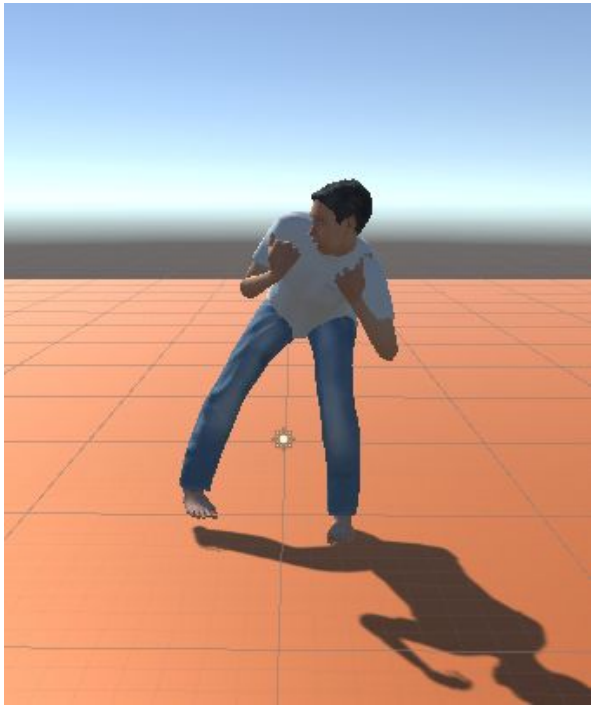
Samuel Oncken

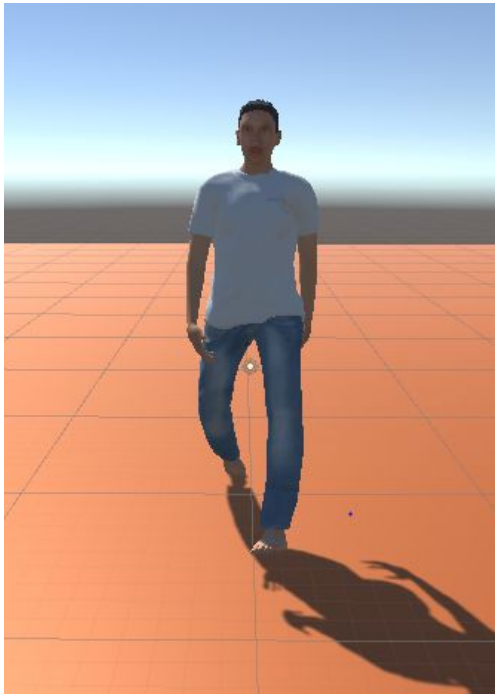
Accomplishments since 403 10 hours of effort	Ongoing progress/problems and plans until the next presentation
- Built Animation Controller to include sub-state machines for each behavior module - alterable by user and in script - Scripted the chaining of behavior module animations (using either keyboard inputs or public string variables) - Sequentially spawning character models to perform links of animations for test scenarios	- Randomize the location of spawned character models - Begin inclusion of central object (to replicate model arm) for spawned humans to interact with - Work on merging animation clips using scripted avatar masks (to be discussed with graduate sponsor)

Robotic Application Extension

Samuel Oncken



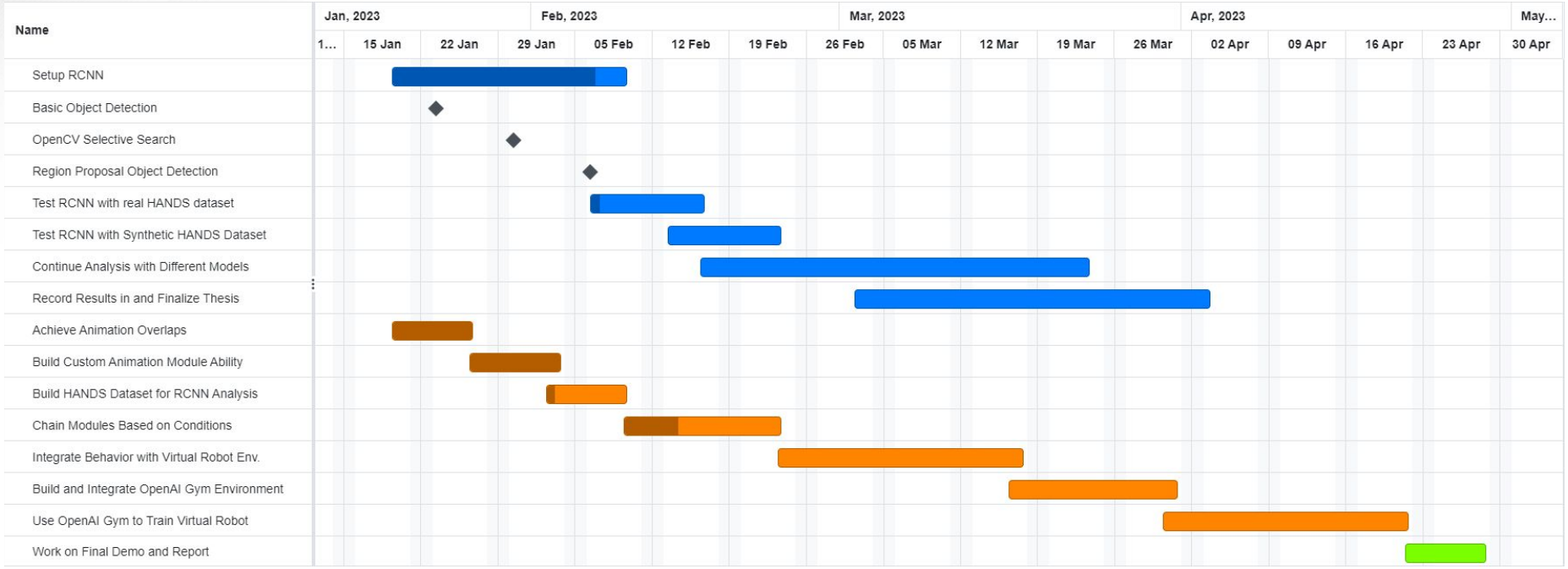




Execution Plan

Legend
- Analysis Extension
- Robotics Integration
- Final work

Darkened = Completed



Dwight Look College of ENGINEERING
TEXAS A&M UNIVERSITY

403 Validation Plan

Test Name	Success Criteria	Methodology	Status	Responsible Engineer(s)
Benchmark Dataset Training	Gesture recognition neural network can run on our home computer and train using the real dataset. Results quantified	Download the benchmark data set and the code for the CNN. Run the code and confirm similar accuracy to benchmark logs provided.	TESTED - Pass	Steven Claypool
Virtual Dataset Training	Gesture recognition neural network can train using our built dataset and provide accuracy results	Take a final virtual dataset modeled after a real benchmark dataset and use it to train the same CNN as the benchmark. Ensure similar accuracy results.	TESTED - Pass	Steven Claypool
Gesture Recognition Accuracy	Accuracy of gesture recognition is within 5% of benchmark accuracy using our virtual dataset	Train gesture recognition neural network using real and synthetic sets and compare accuracy	TESTED - Pass	Steven Claypool
Synthetic Data on Real Data Accuracy	Test real data with CNN trained on synthetic data and achieve accuracy similar to benchmark.	Train a CNN using different methods/compositions with synthetic and/or real data, and test using real data.	IN PROGRESS	Steven Claypool
Unity Hand Mapping	Real hand movement is mapped in Unity	Set up Unity, install Ultraleap plug-ins, map hand motion.	TESTED - Pass	Samuel Oncken
Import Rigged MakeHuman Model	A fully rigged MakeHuman model is imported into Unity	Import model into Unity and confirm appearance and functionality.	TESTED - Pass	Steven Claypool
Virtual Model Unity Hand Mapping	Map hand motion onto an imported MakeHuman model.	Use Hand Binder component/configure settings. Confirm natural motion.	TESTED - Pass	Samuel Oncken
Mounting Stability	Head mounted LMC remains in place during head motion	Mount LMC and plug the device into the computer. Rotate head in all directions and shake head left to right.	TESTED - Pass	Samuel Oncken
Apply Example Animation to Model	MakeHuman model is able to perform an imported full body gesture accurately.	Import an animation .fbx and apply the animation to the rigged human model. Confirm that motion is as expected.	TESTED - Pass	Samuel Oncken
Apply Recorded Gesture Animation to Model	Rigged MakeHuman model can perform a recorded gesture animation.	After recording an animation, apply it to an imported MakeHuman model using the Animator component.	TESTED - Pass	Samuel Oncken
Create and Import Rigged MakeHuman Models to Unity	Minimum 30 MakeHuman models can be generated and imported into Unity	Use MakeHuman “mass produce” function to generate unique character models, each fit with a “Default” rig, with 20% edge cases	TESTED - Pass	Steven Claypool
Data Capture Output and File Type	Virtual camera outputs image data as a .jpg files	Record images of gesture, validate that the data is stored, organized, and is of the desired file type.	TESTED - Pass	Samuel Oncken
Final System Validation	With the press of a button, a large, diverse virtual training set is produced	Run system and validate in output files that each gesture has at least 500 images of gesture performance on differing human models from numerous angles	TESTED - Pass	Samuel Oncken

404 Validation Plan

RCNN Validation	Create a functional RCNN that takes images as an input and outputs the image with correct bounding boxes	Using an image classifier with ImageNet weights, feed sample images into the RCNN and verify correct output of images with bounding boxes	In Progress	Steven Claypool
RCNN Object Detection on Real Data	Test the RCNN with the HANDS dataset and reach acceptable accuracy metrics	Feed dataset into RCNN and measure accuracy using Intersection of Union (IoU) to compare the predicted bounding box and the ground-truth bounding box	Not Tested	Steven Claypool
RCNN Object Detection on Synthetic Data	Test the RCNN with a synthetic HANDS dataset and reach similar accuracy compared to the real dataset	Feed dataset into RCNN and measure accuracy using Intersection of Union (IoU) to compare the predicted bounding box and the ground-truth bounding box	Not Tested	Steven Claypool
Import and Use Mixamo Animations	Animation clips from Mixamo FBX imported files can be applied to MakeHuman virtual models,	Download FBX (for Unity) files from mixamo.com for animations we wish to use. Import them into Unity project folder. Make each prefab “Humanoid” and extract and test animation clip on MakeHuman models	TESTED - Pass	Samuel Oncken
Randomly Select “Behavior” Animation	After model is spawned, depending on user input for desired “behavior”, a randomly selected animation clip will play.	Use Unity C# to script the categorization of “behavior” specific animation clips and random selection. Use Animator on human model with an animation controller selected that includes all desired animation clips.	TESTED - Pass	Samuel Oncken
Ability to Chain “Behavior” Animations	While environment is running, user is able to change input to change animation clip between behaviors to form test scenarios	Use Unity C# to constantly check for user input changes and if a change is detected, immediately switch animation behavior to the one desired by the user	TESTED - Pass	Samuel Oncken
Virtual Robot Unity Environment Integration	Be able to apply created test scenarios to environment created by Pranav including robot model	Transfer all necessary code and information through GitHub and apply all portions to environment already created by Pranav with included Robotic Model.	Not Tested	Samuel Oncken
Human Model Interaction with Robotic Model	If criteria is met such as human models becoming too close to robot arm, they perform a randomized reaction and so does robot	Unity C# scripting for reaction animations under distinct criteria. Also involves integration of my test environment with Pranav’s robot model environment before this can take place.	Not Tested	Samuel Oncken
Use Virtual Environment to Train Robotic Arm	Use OpenAI gym to simulate different test scenarios and train robot. Observe outputs	Download OpenAI gym and determine how to train virtual models using Unity. Work In Progress	Not Tested	Samuel Oncken

Thank you. Questions?