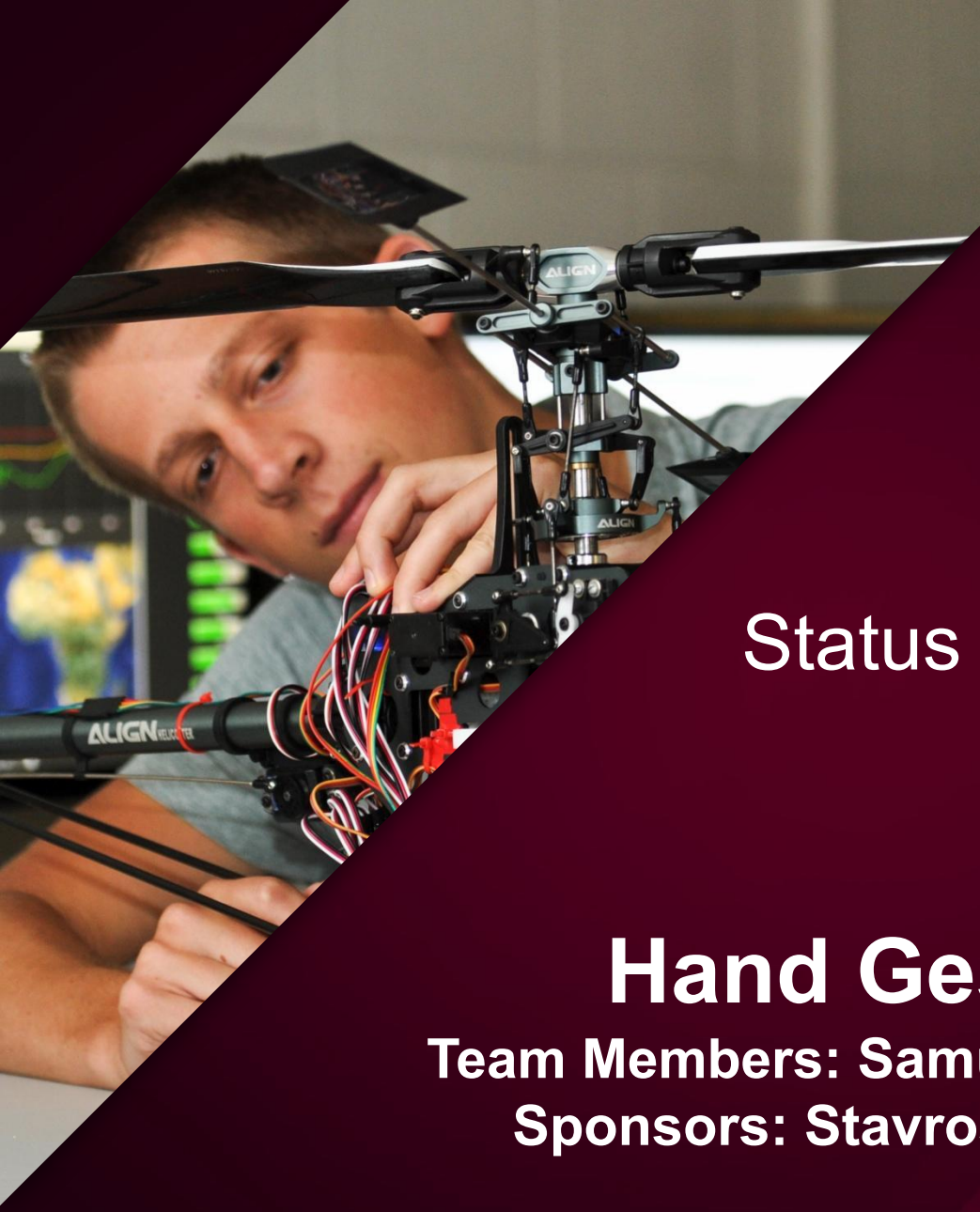


Dwight Look College of
ENGINEERING
TEXAS A&M UNIVERSITY

Status Update Presentation

Hand Gesture Recognition

Team Members: Samuel Oncken, Steven Claypool
Sponsors: Stavros Kalafatis, Pranav Dhulipala

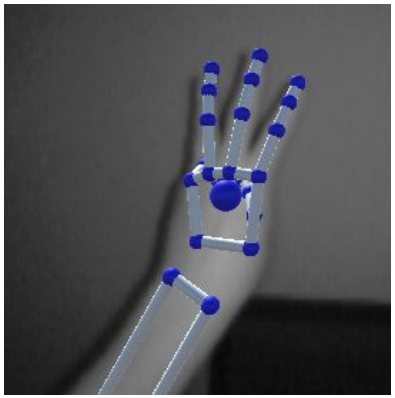
Project Summary

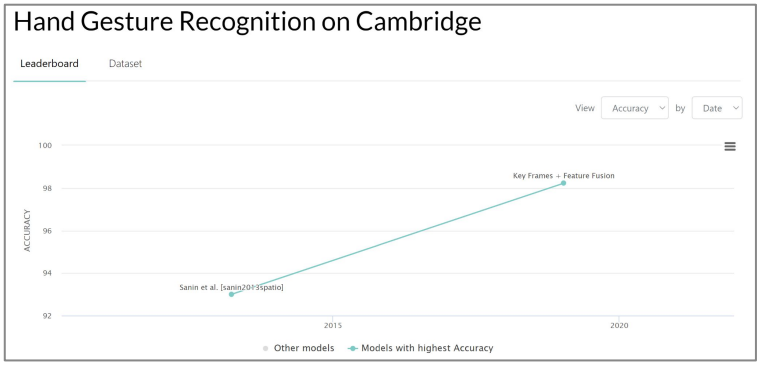
What is our goal?

- 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

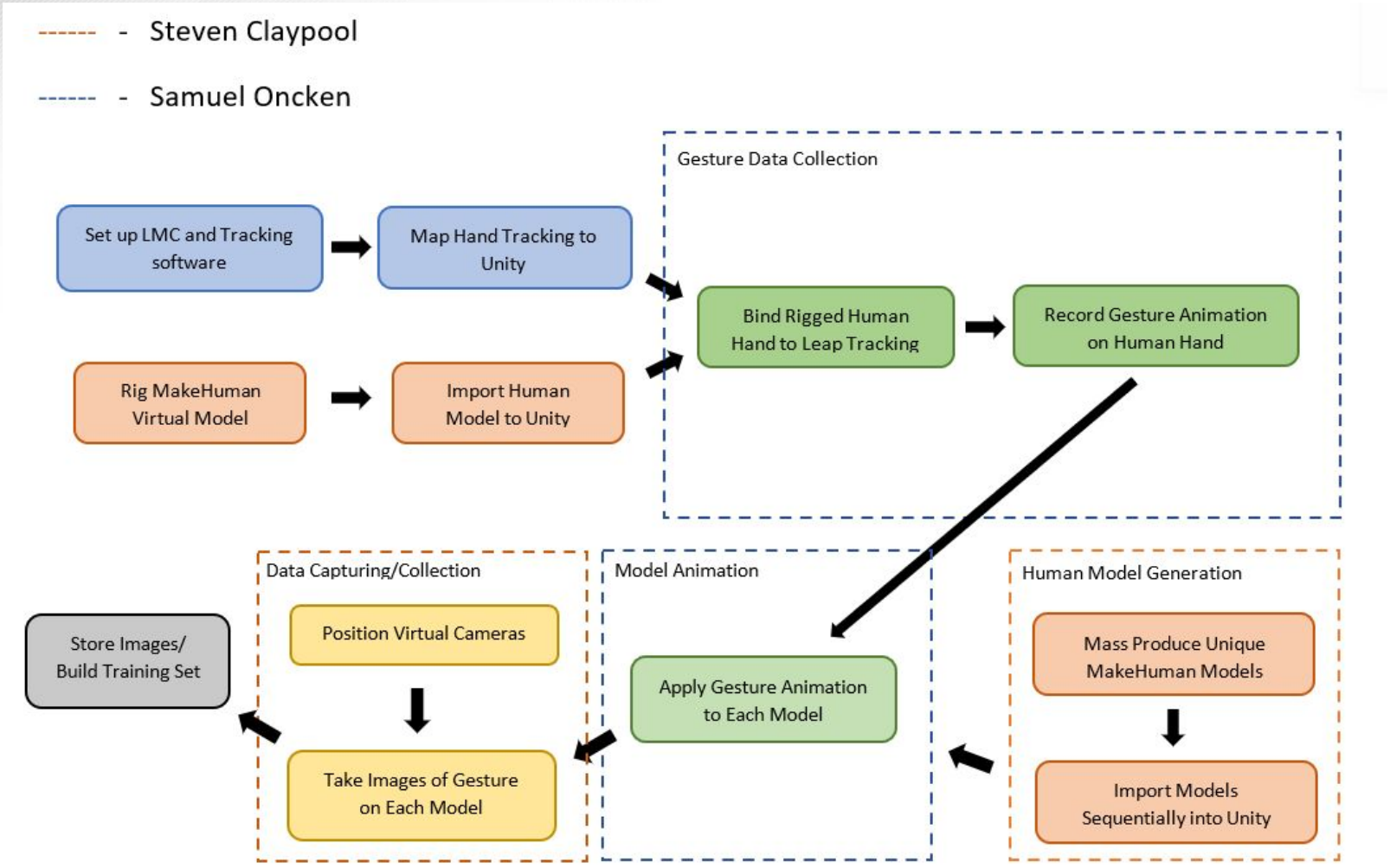
How?

- **Replicate** multiple existing real datasets using the Leap Motion Controller (LMC) and Unity software
- **Test** hand gesture recognition neural network with purely real benchmark data, purely virtual data, and various compositions of each





Subsystem Overview



Gesture Data Collection Subsystem - Update

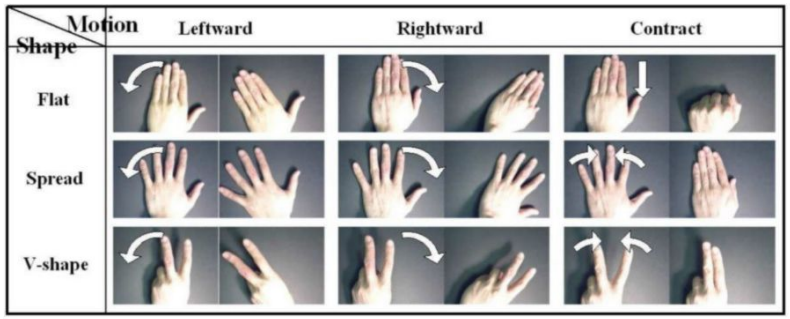
Last presentation I discussed the process of recording a gesture

- Leap Service Provider script, “Hand Binder” script, Unity Recorder, FBX Exporter

Progress:

- Determined real datasets we will be replicating
- Completed Cambridge hand gesture dataset replication
- Completed ASL for numbers replication
- ASL Alphabet replication in progress

- In order to achieve high recognition accuracy when testing a model, we must have accurate recordings.



Gesture Data Collection Subsystem - Update

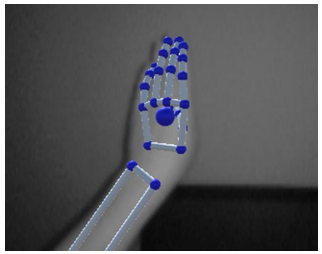
Problem:

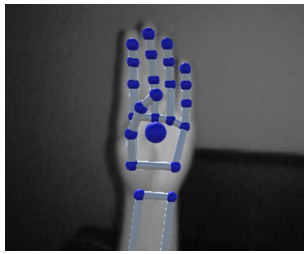
- Tracking accuracy from LMC
  - ➔ Very miniscule details such as thumb placement alter the meaning of a gesture in ASL

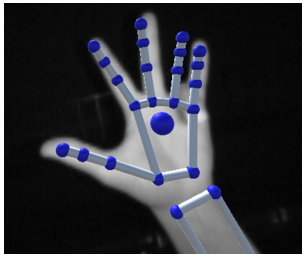
Solution ideas:

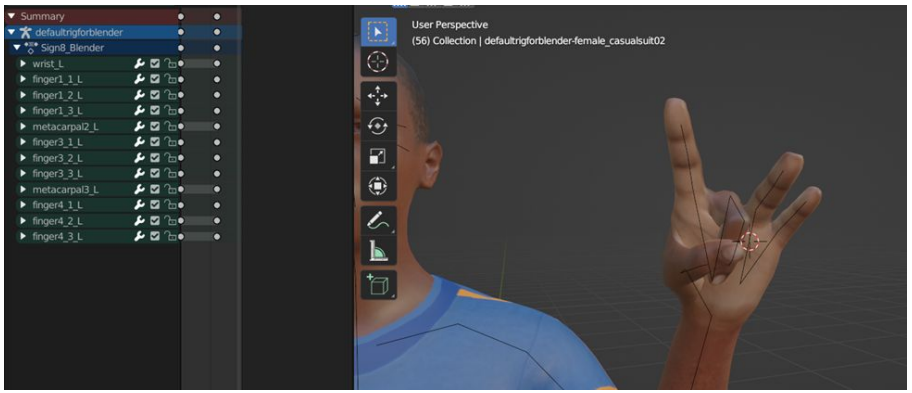
- Change mounting position of the LMC
- Build animation in Blender by positioning fingers manually and recording frames

E









Human Model Generation Subsystem - Update

Problem:

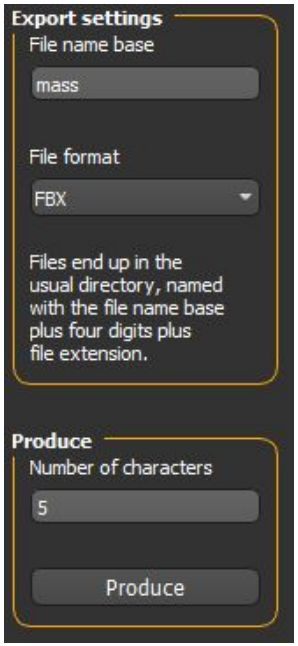
- MakeHuman's Massproduce plugin lacked customization for assets, rigging, and exporting

Progress:

- Downloaded and added additional assets for hand variation (gloves, nails, etc.)
- Currently trying to add default rig to models and choose the export location

Next Steps:

- Script sequential spawning to Unity environment
- Finish altering Makehuman code



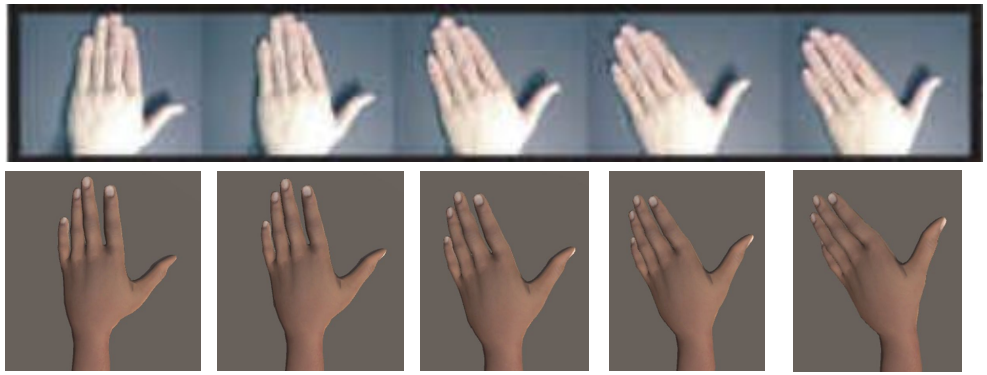
Allowed full body clothes:

	Clothes	Mixed	Female	Male
11	fingernails elegant	☒	☒	☐
12	fingernails elegant	☒	☒	☐
13	gloves hand	☒	☒	☒
14	gloves hand	☒	☒	☒
15	longfingernails bluebutterfly	☒	☒	☐
16	longfingernails bluebutterfly	☒	☒	☐

Model Animation Subsystem - Update

Last presentation I discussed how each animation is applied to a fresh MakeHuman model in Unity

- Animator, animation controller with proper animation clip selected. Scripting to automate this.



Progress:

- Testing each recorded animation clip on freshly imported model
- Began looking into C# scripting for placement of components on each imported model

- Before I can move forward, it is necessary for me to complete the data collection subsystem.

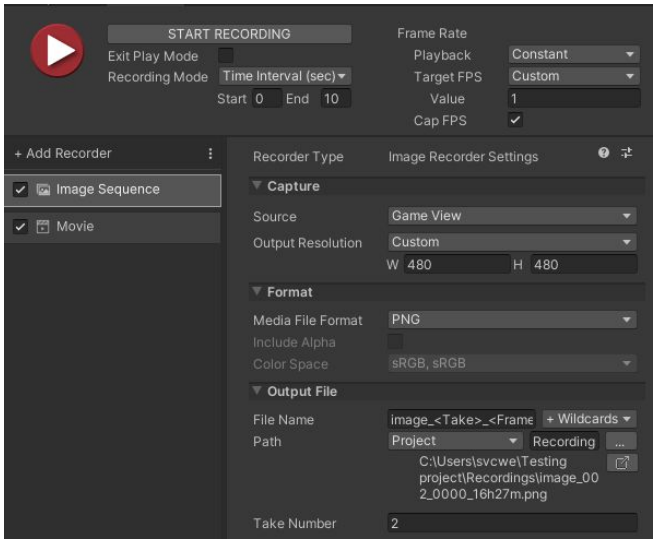
Data Capturing/Collection Subsystem - Update

Progress:

- Taken images at regular intervals through Unity camera

Next Steps:

- Script the recording to start when an animation is loaded onto a model
- Script organizing and exporting of taken images or videos



Training Set Testing - Update

Progress:

- Trained CNN with benchmark dataset for ASL digits dataset

Next Steps:

- Test ASL digits dataset on other models
- Download and test CamGes Dataset with benchmark CNN
- Test with virtual dataset and different dataset compositions after completing the system

```
Epoch 1/10
413/413 [==============================] - 16s 39ms/step - loss: 1.6755 - accuracy: 0.4170 - val_loss: 0.7570 - val_accuracy:
0.7685
Epoch 2/10
413/413 [==============================] - 15s 37ms/step - loss: 0.6110 - accuracy: 0.8066 - val_loss: 0.4503 - val_accuracy:
0.8606
Epoch 3/10
413/413 [==============================] - 15s 37ms/step - loss: 0.3719 - accuracy: 0.8806 - val_loss: 0.3055 - val_accuracy:
0.9039
Epoch 4/10
413/413 [==============================] - 18s 43ms/step - loss: 0.2666 - accuracy: 0.9146 - val_loss: 0.2347 - val_accuracy:
0.9258
Epoch 5/10
413/413 [==============================] - 16s 38ms/step - loss: 0.1903 - accuracy: 0.9401 - val_loss: 0.2078 - val_accuracy:
0.9330
Epoch 6/10
413/413 [==============================] - 16s 39ms/step - loss: 0.1433 - accuracy: 0.9560 - val_loss: 0.1913 - val_accuracy:
0.9394
Epoch 7/10
413/413 [==============================] - 15s 37ms/step - loss: 0.1150 - accuracy: 0.9633 - val_loss: 0.1909 - val_accuracy:
0.9433
Epoch 8/10
413/413 [==============================] - 15s 36ms/step - loss: 0.0943 - accuracy: 0.9700 - val_loss: 0.1921 - val_accuracy:
0.9442
Epoch 9/10
413/413 [==============================] - 15s 36ms/step - loss: 0.0817 - accuracy: 0.9742 - val_loss: 0.1700 - val_accuracy:
0.9494
Epoch 10/10
413/413 [==============================] - 15s 36ms/step - loss: 0.0729 - accuracy: 0.9761 - val_loss: 0.1686 - val_accuracy
0.9533
```

```
Epoch 8/10
13200/13200 [==============================] - 20s 2ms/sample - loss: 0.0985 - accuracy: 0.9
664 - val_loss: 0.2310 - val_accuracy: 0.9391
Epoch 9/10
13200/13200 [==============================] - 19s 1ms/sample - loss: 0.0918 - accuracy: 0.9
695 - val_loss: 0.1604 - val_accuracy: 0.9597
Epoch 10/10
13200/13200 [==============================] - 19s 1ms/sample - loss: 0.0704 - accuracy: 0.9
771 - val_loss: 0.3413 - val_accuracy: 0.9206
```

Execution Plan

	October 12th	October 26th	November 9th	November 23rd	November 30th
Gesture Data Collection	- Choose real datasets to replicate - Complete hand mapping configurations	- Complete gesture animation recording according to chosen datasets to replicate			
Human Model Generation	- Produce 6 models into Unity for animation testing	- Mass produce models into Unity environment sequentially			
Model Animation	- Apply recorded test gesture to a single model	- Randomly apply a recorded gesture to any model			
Data Capture/Collection	- Place at least 5 virtual cameras in Unity environment to face model hand from multiple angles	- Record and store images of gesture performed on any model			
Training Set Completion/ Testing		- Completion of training set creation system. - Testing process begins	- Preprocess each training set - Create new training sets ranging in composition of real and synthetic data	- Train each neural network with new training sets and record the metric used in the real dataset paper for proper comparison	- Evaluate results after comparison and prepare system and outcomes for final presentation.

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.	UNTESTED	All
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	UNTESTED	All
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.	IN PROGRESS	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.	UNTESTED	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
Mass Produce Rigged MakeHuman Models to Unity	Minimum 100 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	IN PROGRESS	Steven Claypool
Data Capture Output and File Type	Virtual camera outputs image data as a .png file or video data as mp4 (TBD from neural networks used).	Record images/videos of gesture, validate that the data is stored, organized, and is of the desired file type.	UNTESTED	Steven Claypool
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 performance on differing human models from numerous angles	UNTESTED	All