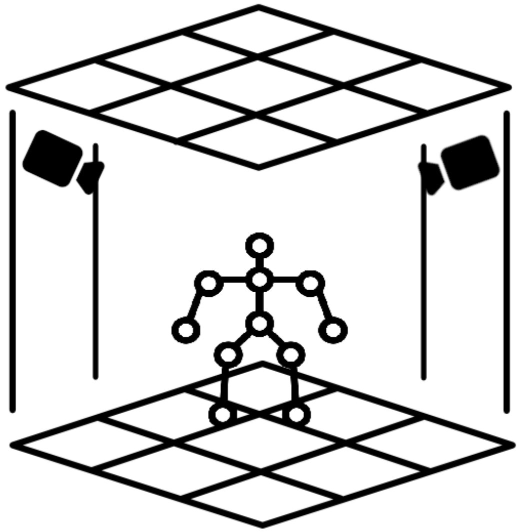


# 3D Non-Contact Whole Body Musculoskeletal Exercise Dataset

Aidan Kimberley | Ethan Matzek | Ava Megyeri

# Table of Contents



ALTEC

# Introductions

Aidan Kimberley

- McGill University
- B.Eng. in Mechanical Engineering
- Altec Intern from 5/5 – 8/15
- Data Analysis | Biomechanics

Ethan Matzek

- Wright State University
- M.S. in Computer Science and Engineering
- Altec Intern from 5/12 – 8/8
- Computer Hardware & Software

Ava Megyeri

- Wright State University
- Ph.D. in Computer Science and Engineering
- Altec Intern from 5/27 – 8/15
- Algorithm Development

# Study Objectives

- Capture a high-fidelity dataset across muscular-skeletal injury (MSKI) rehabilitation exercises to:

1. Design and validate custom 3D body tracking algorithms
2. Compare IR, RGB, and point-cloud data quality between different camera alternatives
3. Perform validation assessment of new solutions with respect to a reference motion capture system

- The Microsoft Kinect is discontinued and has become antiquated. Finding a suitable replacement for Kinect Body Tracking is the first step in moving away from depreciated hardware

# MEDIC System

**M**USCULOSKELETAL **E**VALUATION **D**EVICE FOR **I**NJURY **C**ONTROL

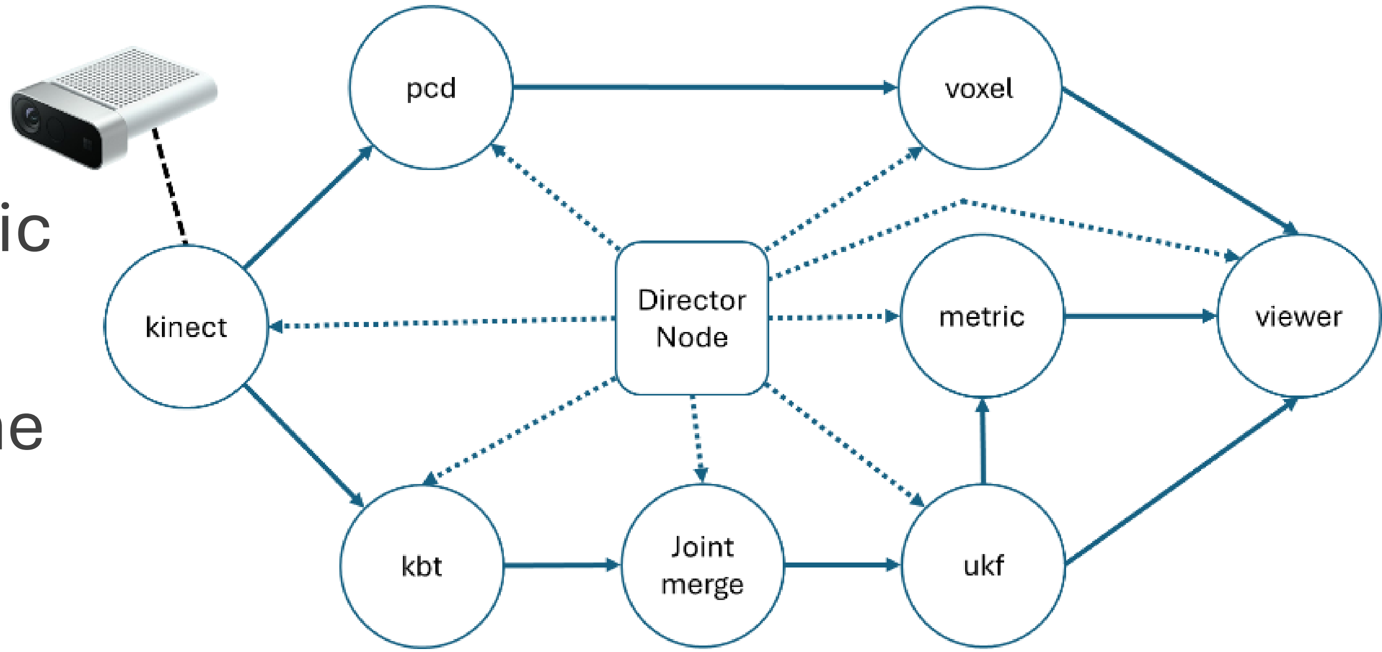
SECTION 2

# Purpose

- MSKIs are the leading cause of disabilities among warfighters
- A majority of MSKIs are non-battle related, but rather stem from poor technique and excessive joint loading
- MEDIC was developed for exercise assessment, injury screening, developing injury resilience, and providing tailored biofeedback to users

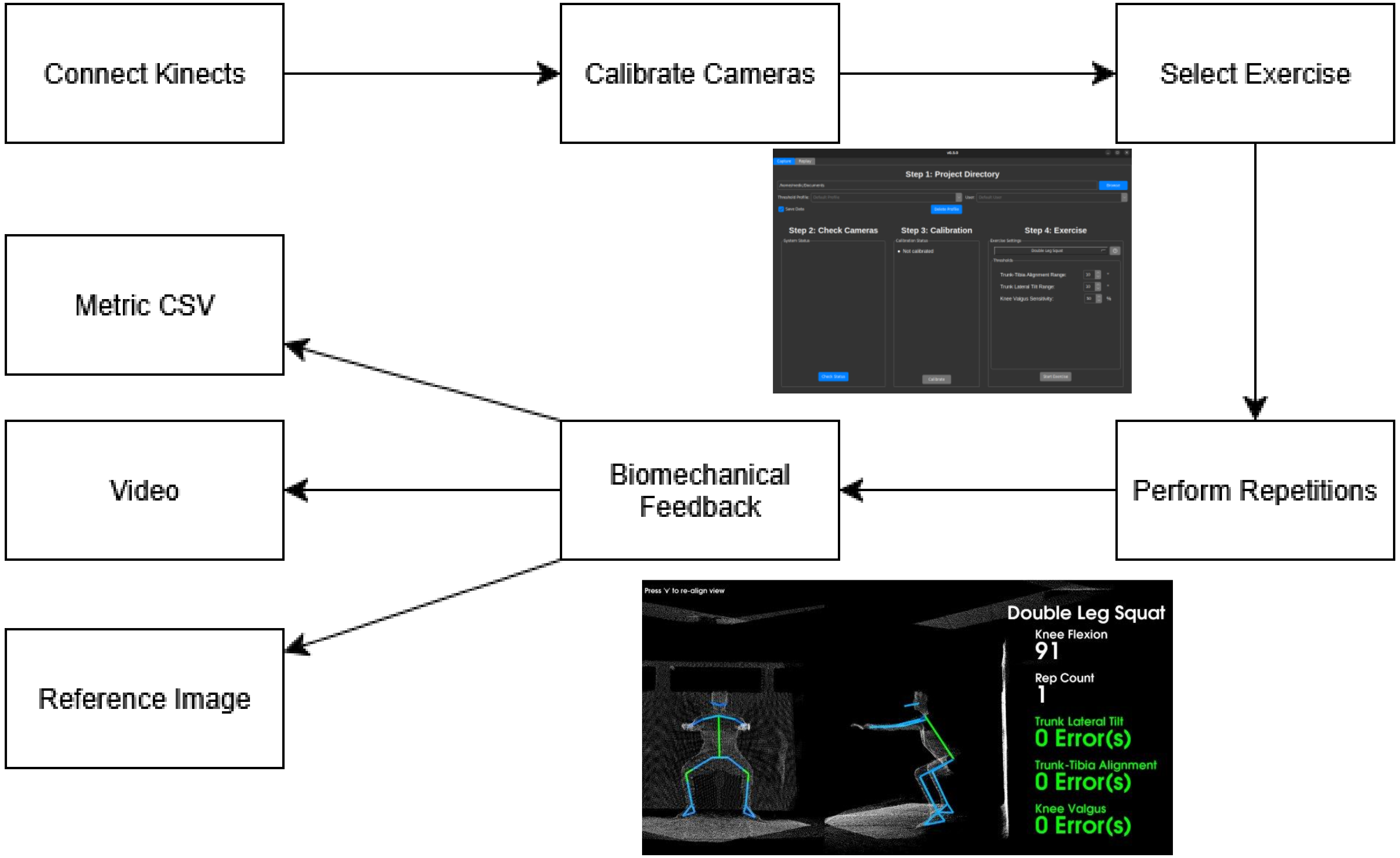
# System Architecture

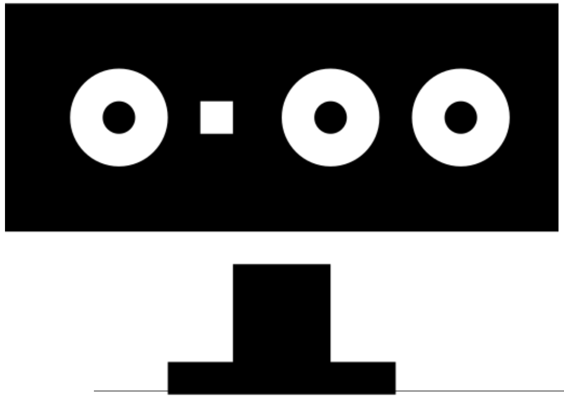
- MEDIC uses a node-based architecture
- Each node performs a specific function
- A master node commands the other nodes
- Nodes communicate asynchronously, so they can run on separate PCs



# User Experience

- MEDIC offers 9 common military battery exercises
  - 6 High exercises (3’)
  - 3 Low exercises (1’6”)
- User interacts with the system through a GUI





# 16-Kinect System

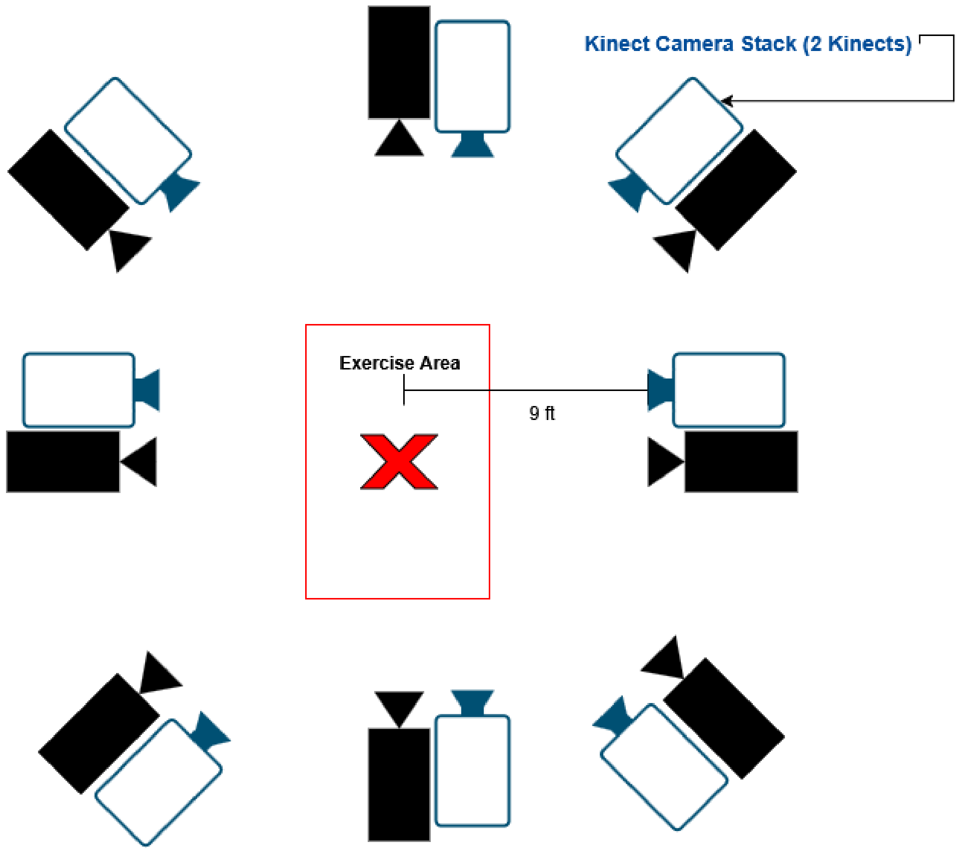
SECTION 3

# Purpose

- Evaluate current MEDIC biomechanical outcomes versus Vicon (reference standard)
- Establish baseline feasibility of custom 3D human pose estimation algorithms
- Set foundation for evaluating the trade-off between accuracy and number of cameras in pose estimation

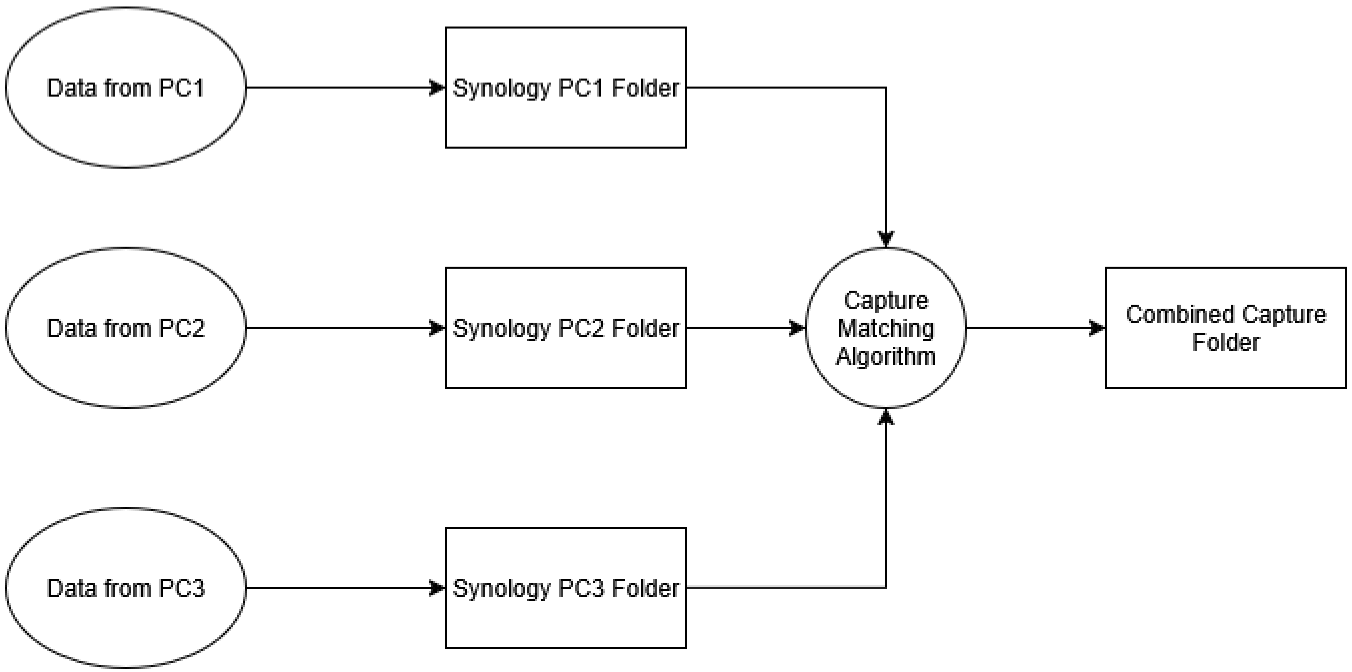
# Hardware Requirements

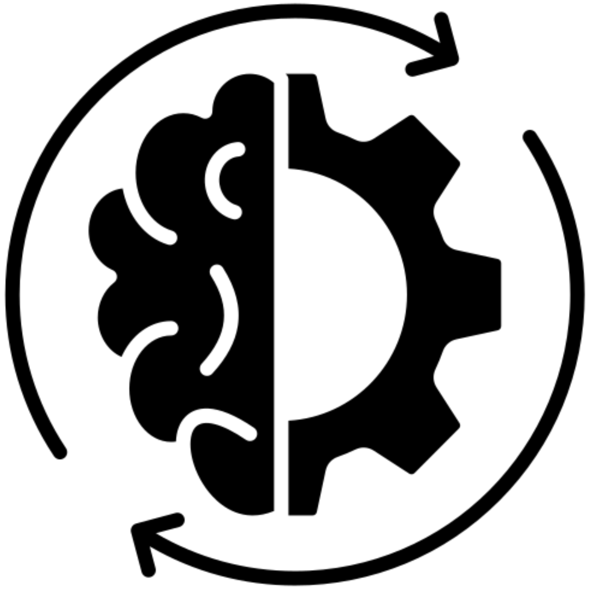
- 16 Kinects
  - RGB, Depth, and IR images
  - Native 30 FPS
- Stacked cameras with time offset emulate 60 FPS capture
- Cameras positioned in a circle
  - 45° spacing
  - 9' radius
- 2 additional PCs required to run the system



# Software Requirements

- Increasing output to 16 devices necessitates file management and bulk storage
- Each capture session records ~1TB of data
- A 10–gigabit ethernet switch connects the capture PCs to a central datacenter
- Capture matching algorithms are used to transfer and combine data on the server





# Sapiens Pipeline

SECTION 4

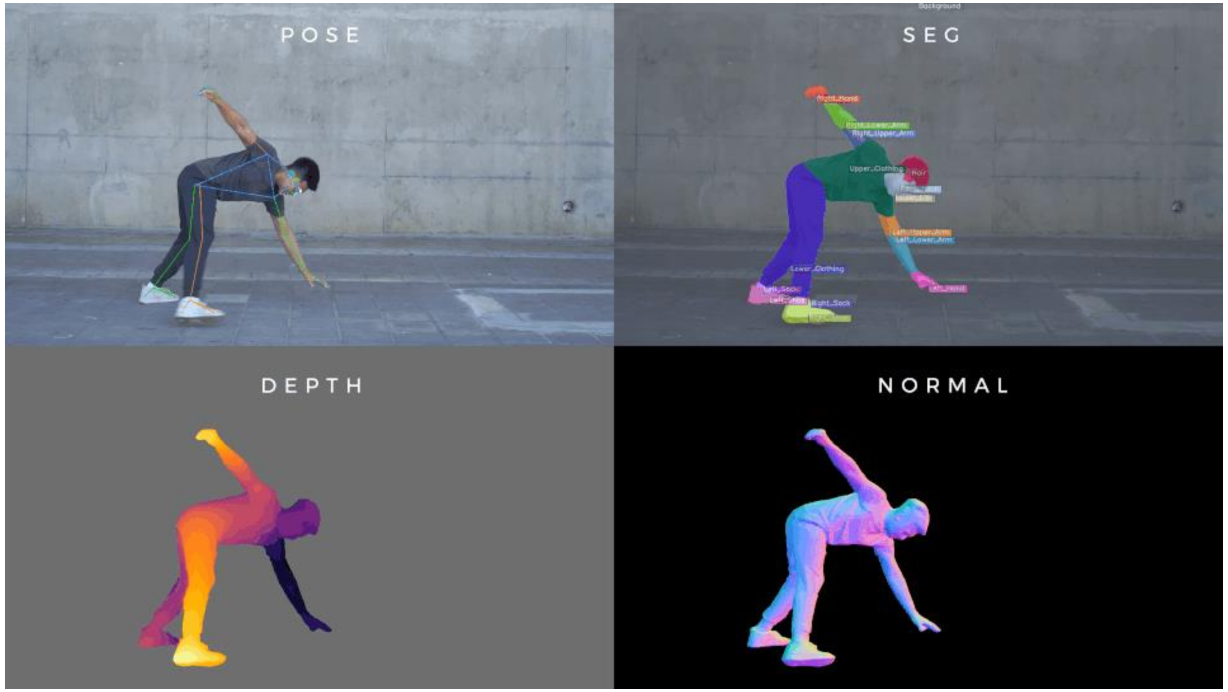
# Purpose

- Retain MEDIC functionality when retiring Kinect Body Tracking
- Create a model agnostic pipeline to find the best performing option

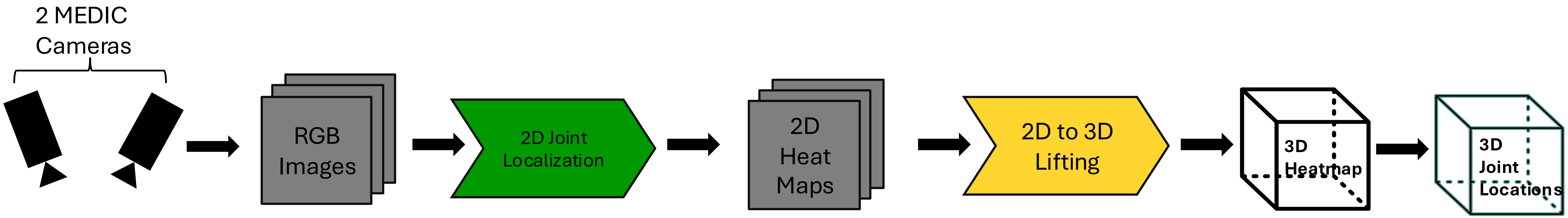
| 2D Pose Estimation Algorithm | Model Metadata |
| --- | --- |
| Real-Time Multi-Person 2D and 3D Whole-body Pose Estimation | Real-time, 133 keypoints |
| Real-Time Multi-Person Pose Estimation based on MMPose | Real-time, 17 or 26 keypoints |
| Sapiens 0.3 Billion | Real-time, 17 or 133 keypoints |
| Sapiens 0.6 Billion | Post, 17 or 133 keypoints |
| Sapiens 1 Billion | Post, 17 or 133 keypoints |
| Sapiens 2 Billion | Post, 17,or 133 keypoints |
| OpenPose | Real-time, 135 keypoints |
| MediaPipe | Real time, 33 keypoints |

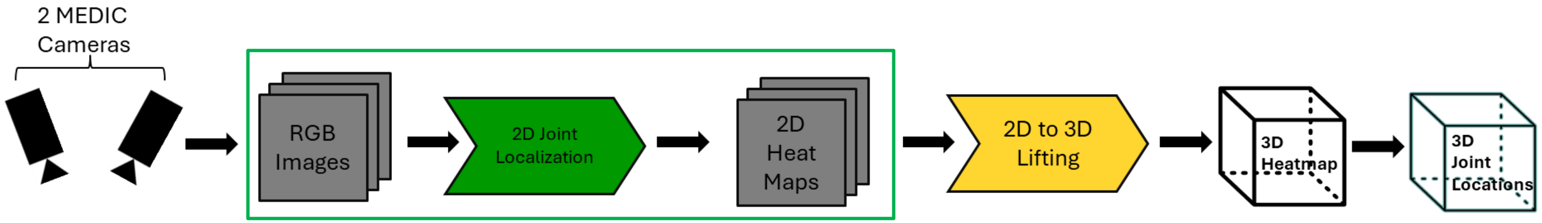
# Sapiens Algorithm Overview

- Sapiens is a foundation model for human-centric vision tasks
  - 2D pose estimation
  - Body-part segmentation
  - Depth estimation
  - Surface normal prediction



# Overview of the Pipeline

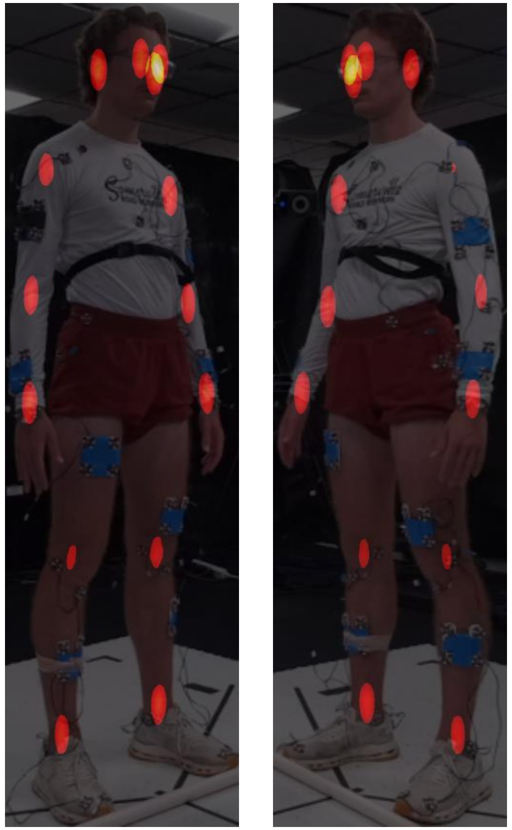




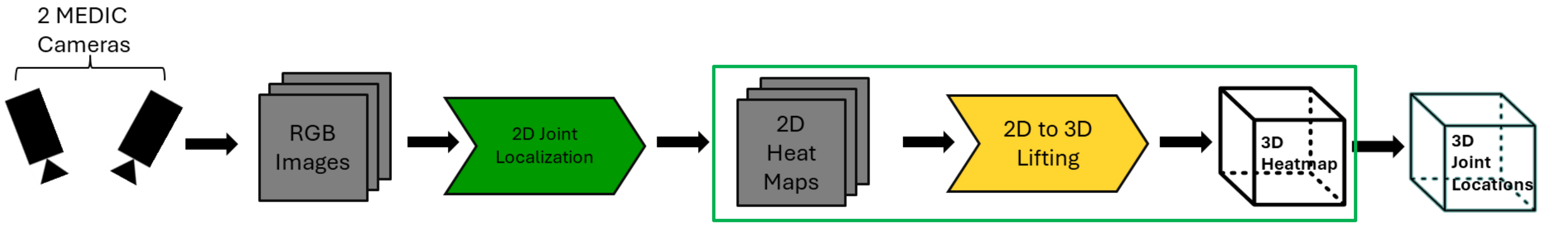

# Sapiens Algorithm Overview

Input RGB Images

YOLO Cropping

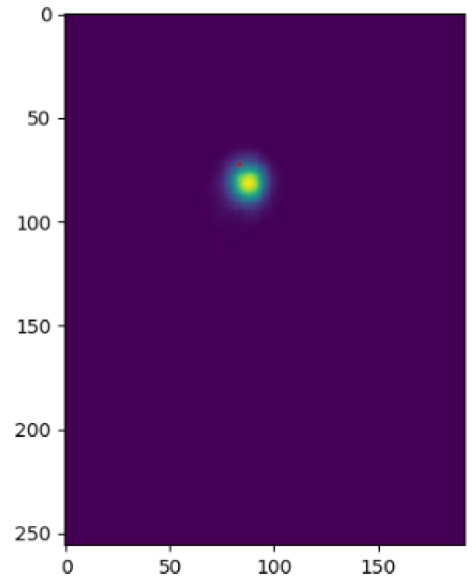


2D Heatmaps

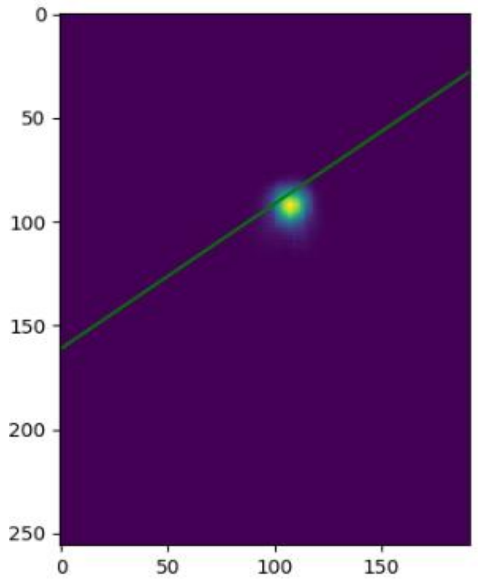



# 2D to 3D Lifting

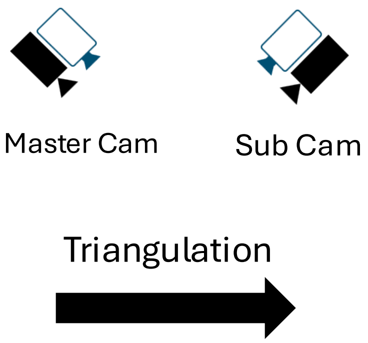
- 2 cameras viewpoints.
- Create the fundamental matrix to relate the points from one viewpoint to another
- Triangulate the points to get the 3D points

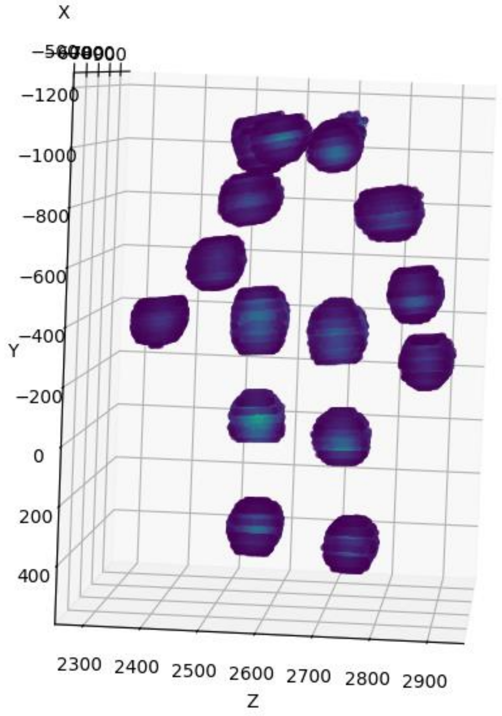


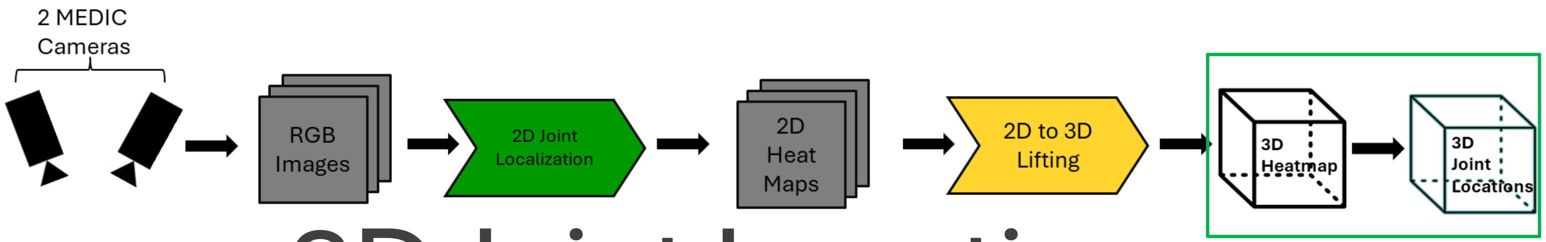

Master Cam




Subordinate Cam

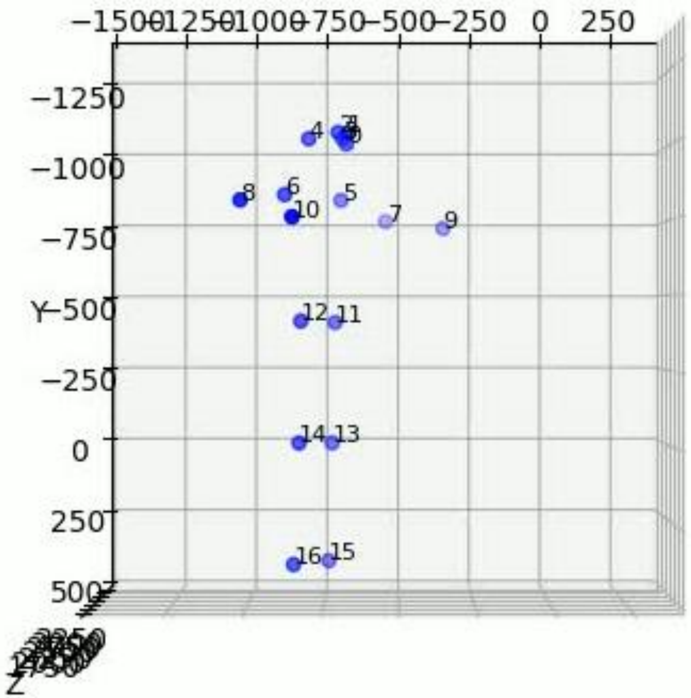


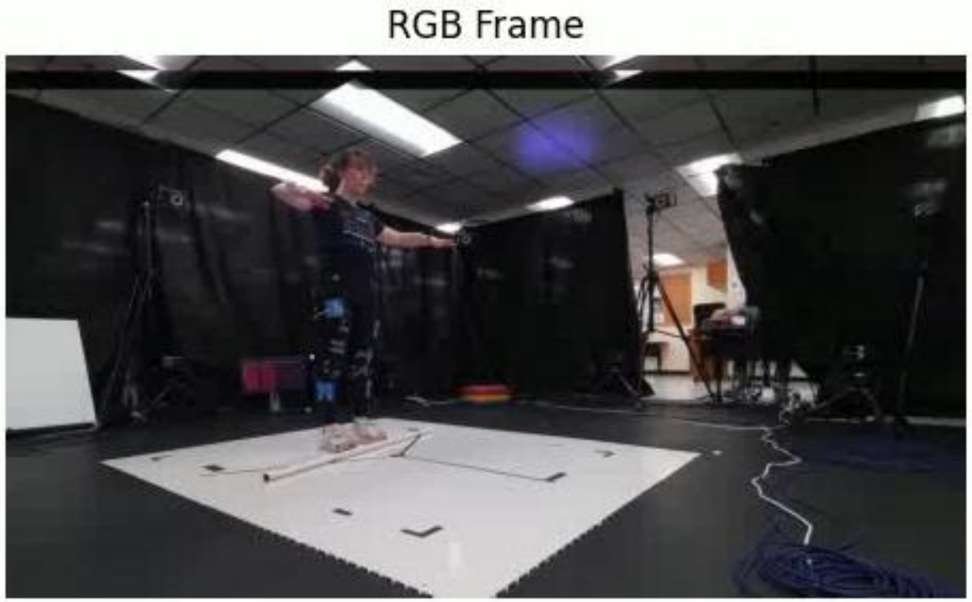






# 3D Joint Locations

- Calculate the weighted mean of all points, using the probabilities as the weights



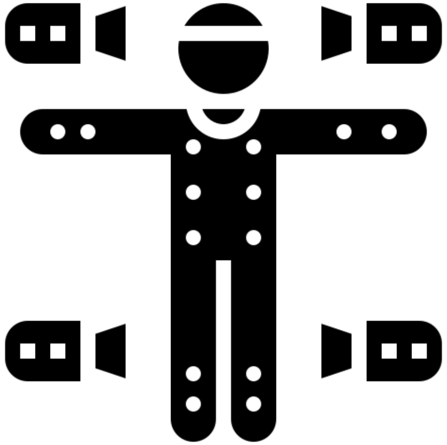



# Feasibility Assessment

- Pipeline is model agnostic
  - All Sapiens model variations were confirmed to work
  - Open-source body tracking models use standardized inputs
    - Any off-the-shelf model can be used in the pipeline

# Capabilities of the Pipeline

- Adding musculoskeletal constraints could improve the outputs
- Different viewpoints can achieve more accurate 3D joints
- Finetune the algorithm for our use-case to improve accuracy

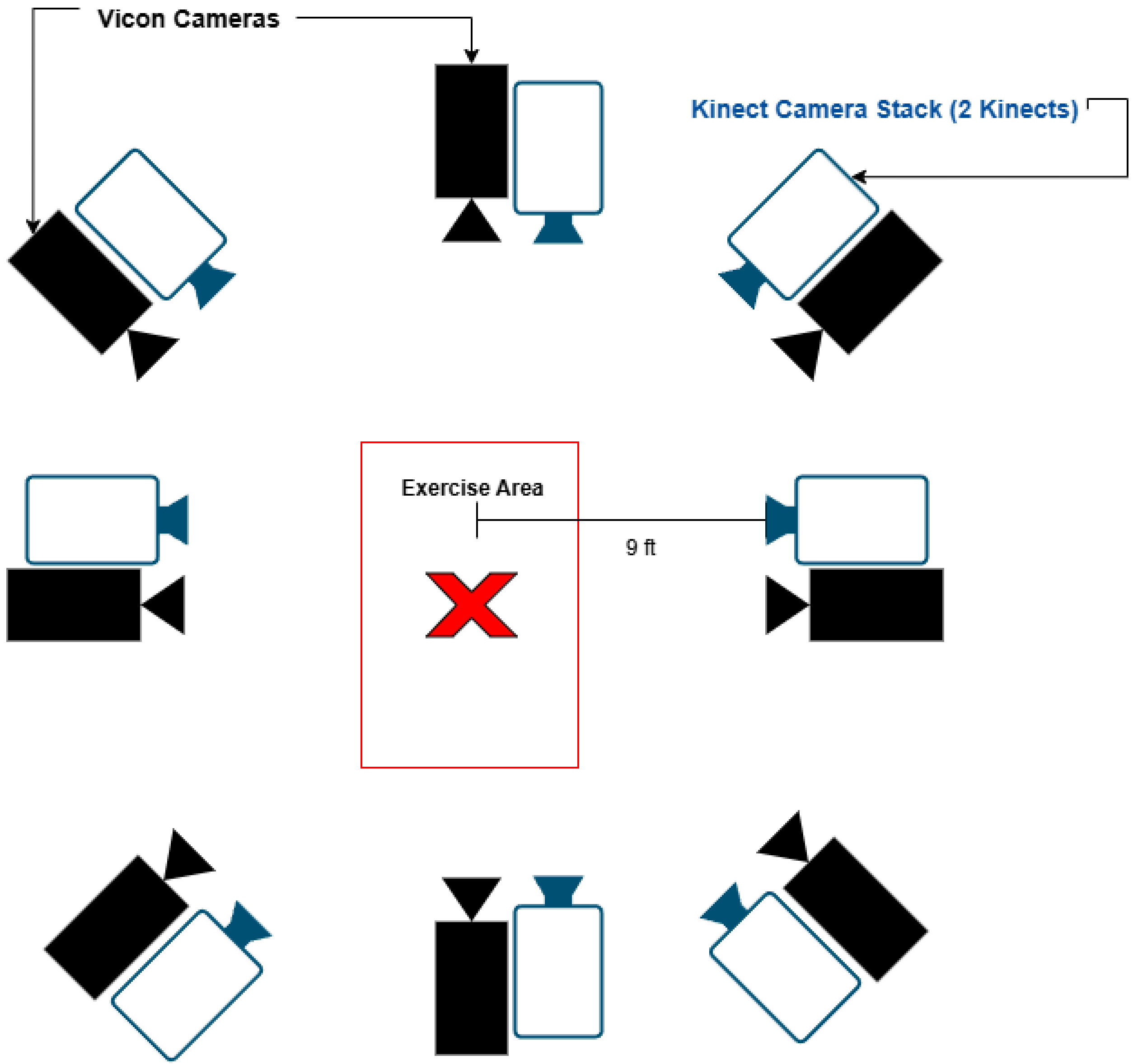


# Vicon System

SECTION 5

# Vicon System Overview

- Used Vicon as ground truth for joint position estimation
- Had 8 Vicon T160 cameras:
  - Greater horizontal FOV than vertical. 54 vs 39.7 deg
  - Placed at optimal heights and angles to minimize occlusions
  - Standard Operating at 790nm
- Marker labelling done in Nexus 2.16 software
- Used built in interpolation functions to estimate occluded marker positions

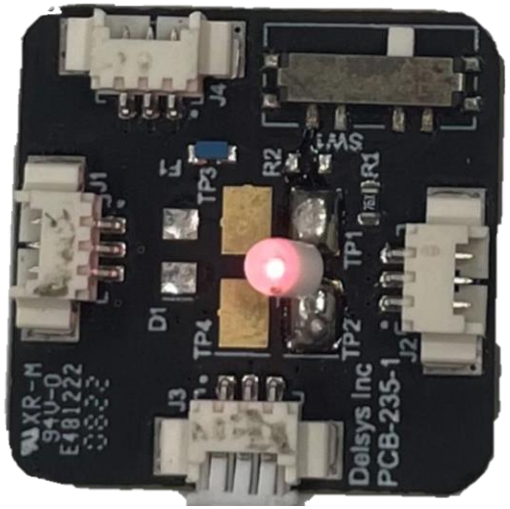



ALTEC

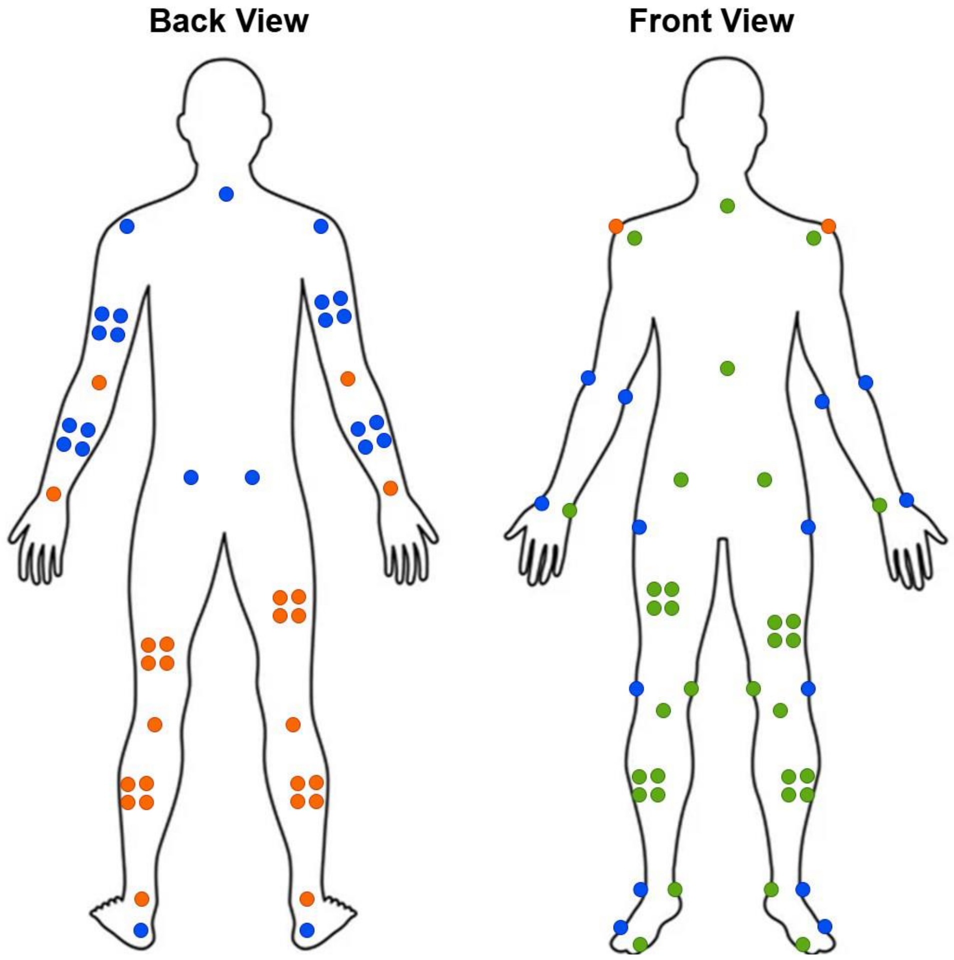
# Vicon Marker Placement & Joint Estimation

- Markers placed according to International Society of Biomechanics (ISB) standards
- Using marker group centroids to find joints
- Marker clusters attached to improve interpolation accuracy
- Used custom 740nm active LED markers to avoid interference with Kinect TOF sensor
  - Kinect operates at 850nm
  - Vicon T160 operates at 790nm
- Two marker sets used, optimized for visibility during different exercises

*Active Marker*




*Marker set Diagram*



SECTION 6

# Data Collection

# Study Overview

- Subjects:
  - 15 non-military controls with no MSKIs in the past 12 months
  - Even M/F proportions
- Each subject performed 27 different exercises
- Exercises repeated at 2 camera heights
- Goal:
  - Assess body tracking systems' performance across varied body types, exercises, and camera heights using Vicon as the reference
  - Validate real-time Medic outcomes
  - Test Sapiens as an alternative
  - Obtain a corpus of data to be used for future studies

| Exercises (MEDIC Exercises) | Outcomes Tracked |
|---|---|
| Single Leg Balance | Trunk Lateral Tilt, Trunk Forward Tilt, Knee Abduction |
| Double Leg Squat | Trunk Lateral Tilt, Trunk Forward Tilt, Hip Flexion, Knee Flexion, Foot Height |
| Forward Lunge | Trunk Lateral Tilt, Trunk Forward Tilt, Postural Sway, Balance Time |
| Prisoner Jump Squat | Knee Flexion |
| Side Hop To Balance | Trunk Lateral Tilt, Knee Abduction, Knee Flexion, Rep Count |
| Overhead Press | Knee Flexion, Knee Distance, Hip Flexion, Trunk Forward Tilt, Trunk Lateral Tilt, Trunk Tibia Alignment |
| Hex Bar Deadlift | Knee Flexion, Hip Flexion, Trunk Lateral Tilt, Trunk Forward Tilt, Leading Knee Past Ankle |
| Plank | Knee Flexion, Hip Flexion, Trunk Forward Tilt, Trunk Lateral Tilt, Knee Distance, Foot Height |
| Push Up | Trunk Lateral Tilt, Knee Flexion, Lateral Jump Distance |
| Double Leg Stance | Trunk Lateral Tilt, Trunk Forward Tilt, Pelvis Lateral Tilt |
| Tandem Stance | Knee Flexion, Hip Flexion, Trunk Forward Tilt, Trunk Lateral Tilt, Foot Height, Knee Distance, Leading Knee Past Ankle |
| Standing Knee Flexion | Trunk Lateral Tilt, Trunk Forward Tilt, Hop Distance |
| March In Place | Knee Flexion, Trunk Forward Tilt, Trunk Tibia Alignment |
| Ice Skater | Wrist Height Over Shoulder, Trunk Lateral Tilt, Trunk Forward Tilt, Elbow Flexion, Hip Flexion, Knee Flexion, Wrist Past Ankle |
| Single Leg Reach | Foot Height, Knee Flexion, Pelvis Stability, Trunk Lateral Tilt |
| L Hop | Knee Flexion, Trunk Forward Tilt, Trunk Lateral Tilt, Ankle Distance |
| Step Up | Knee Flexion, Trunk Forward Tilt, Trunk Tibia Alignment, Hip Flexion |
| Hurdle Step | Knee Flexion, Trunk Forward Tilt, Trunk Lateral Tilt, Trunk Tibia Alignment, Hip Flexion, Knee Distance |
| Y-Balance Test | Knee Flexion |
| Barbell Deadlift | Knee Extension, Rep Count |
| Heel Slide | Trunk Lateral Tilt, Rep Count, Task Time |
| Knee Lift | Hip Height, Pelvis Lateral Tilt |
| Sit To Stand | Single Leg Hip Flexion, Pelvis Lateral Tilt, Foot Height |
| Single Leg Bridge | Trunk Forward Tilt, Elbow Flexion, Hip-Center Displacement, Hip Flexion, Knee-Center Displacement |
| Straight Leg Raise | Knee Height, Elbow Shoulder Alignment, Trunk Lower body Alignment |
| Side Plank | Knee Flexion, Hip Flexion, Elbow Flexion, Hip-Center Displacement, Trunk Forward Tilt |
| Hand Release Push Up | Trunk Lower body Alignment, Elbow Flexion, Knee Height |

# Protocol Summary

**Briefing**
- Subjects familiarized with all exercises prior to recording

**Markers**
- 69 active LED markers applied to subjects
- Marker-set changed mid-collection

**Exercises**
- 5 reps per set
- 27 exercises, 56 total sets

**Finish**
- Total time ~3.5 hours

# Capture Timeline

- Data Collection every day for 3 weeks
- Processing data in-between collections

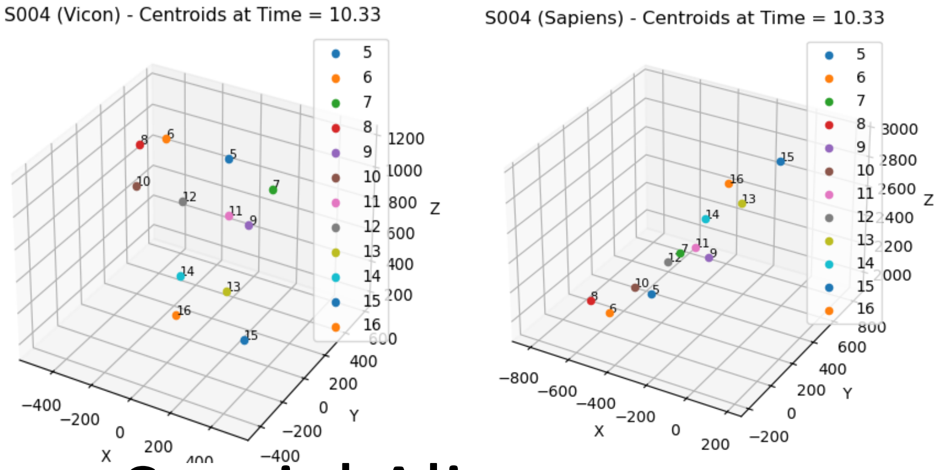
| Day # | Task | Time Estimate |
| --- | --- | --- |
| Day –1 | Familiarize subject | 15 min |
| Day 0 | Data Collection | 5 hours |
| Day 0 | Transfer data from PC to Candor DC-1 server | 5 hours |
| Day 0 to 1 | Kinect file sorting on Candor DC-1 server | 1 hour |
| Day 0 to 1 | Sapiens pipeline and lift3D | 10 hours |
| Day 0 to 1 | MEDIC pipeline | 30min – 1 hour |
| Day 1 | Vicon annotation | 2-4 hours |
| Day 2 | Biomechanics pipeline | 30min – 1 hour |

SECTION 7

# Data Analysis Pipeline

# Input Alignment Obstacles

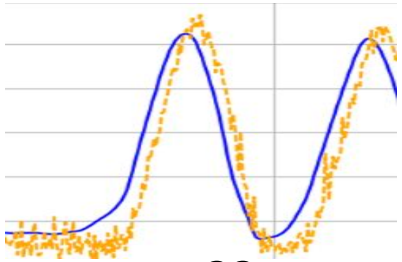
- Inputs
  - Vicon Marker Trajectories 100Hz
  - Sapiens Joint Trajectories 60Hz
  - Medic Timeseries Metric Outcomes 60Hz




**Challenges:**

## Sampling Rate

- Vicon: 100Hz
- MEDIC and Sapiens: 60Hz



## Time Offset

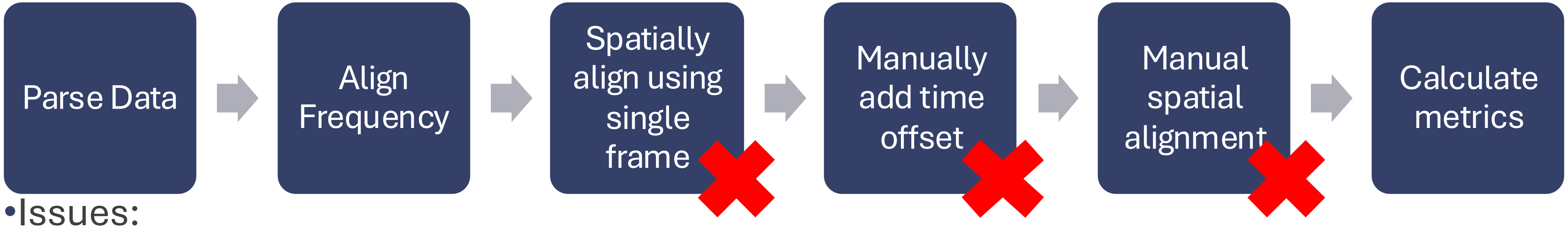
- Kinect and Vicon cameras controlled separately
- MEDIC capture starts after body scan

## Spatial Alignment

- Vicon and Kinect have different coordinate systems

# Errors in Existing Biomechanics Pipeline

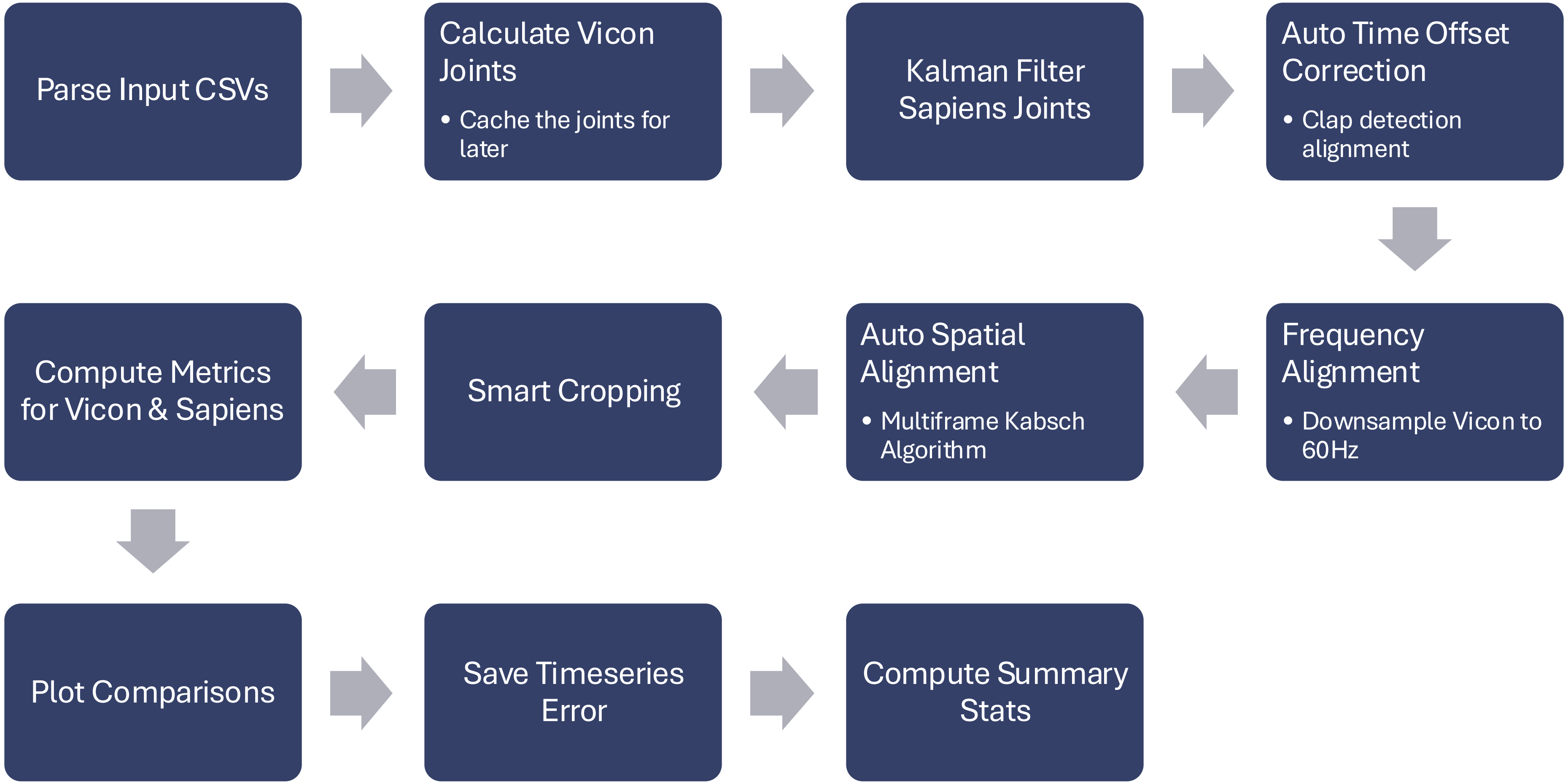
- Existing Biomechanics Pipeline Overview:




- Issues:
  - Auto-spatial alignment failed due to temporal misalignment
  - Extensive manual input -> prolonged processing time and constant human supervision
  - No applicability across different exercises & metrics
  - Used 'vicon planes' rather than body-defined planes for kinematics calculations

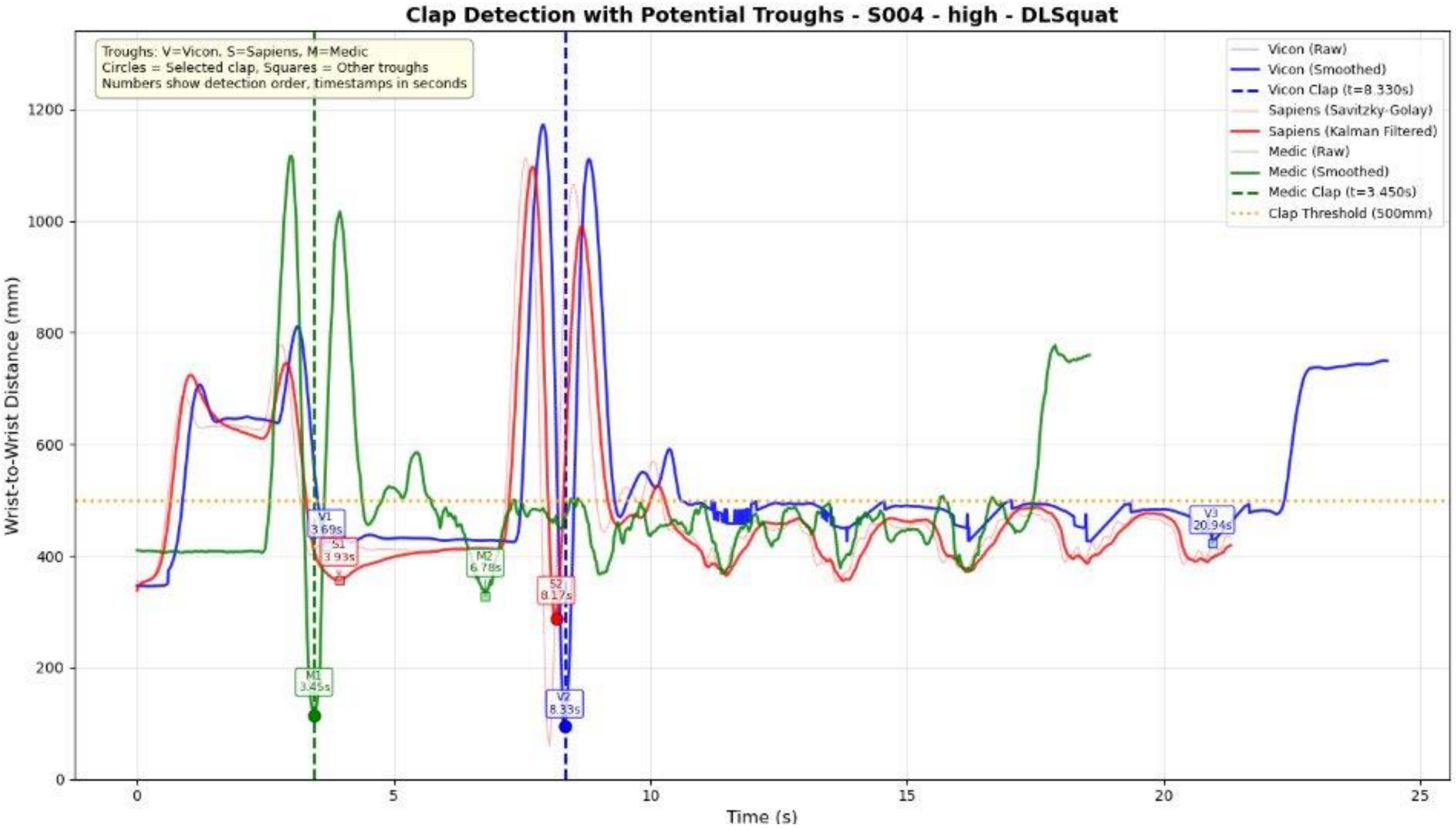
- Ended up  largely writing new pipeline from scratch

# Sapiens Medic Validation Pipeline Operations Overview



# Clap Based Time Alignment

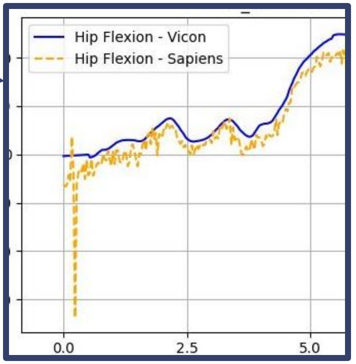
- Clap at exercise start
- Distinct clap events are detected
- Savitsky Golay filtering applied to get clear peaks
- Kalman filtering induces a time offset and flattens peaks:
  - Initial search window is identified using Savitsky Golay filtered data
  - Final clap detection refined with Kalman filtered data
- Potential troughs labelled




ALTEC

# Improved Spatial Alignment: Multi-frame Kabsch Algorithm

- Fully temporally aligned before spatial alignment
- Middle 20% of frames used for joint input
  - Using multiple frames reduced noise effects
  - Sapiens less accurate in initial frames. Middle is the most reliable
  - Alignment most important during the exercise (middle of the capture).
- Kabsch Algorithm
  - Minimizes RMSE between two paired sets of points
  - Stack frames as one tall matrix input
  - No scaling factor used – testing raw distance accuracy
  - If det(R)<0, flip the sign of the rightmost right singular vector ($V_{2.4T,3}$) and recompute R to prevent reflections




$$N = 12\ (\#\ joints), T = \#\ Frames, D = 3(x, y, z\ dims)$$

$$P = SapiensJoints \in \mathbb{R}^{(0.2 \cdot T \cdot N) \times 3}$$

$$Q = ViconJoints \in \mathbb{R}^{(0.2 \cdot T \cdot N) \times 3}$$

$$C = (P)^{\top} Q$$

$$C = U \Sigma V^{\top}$$

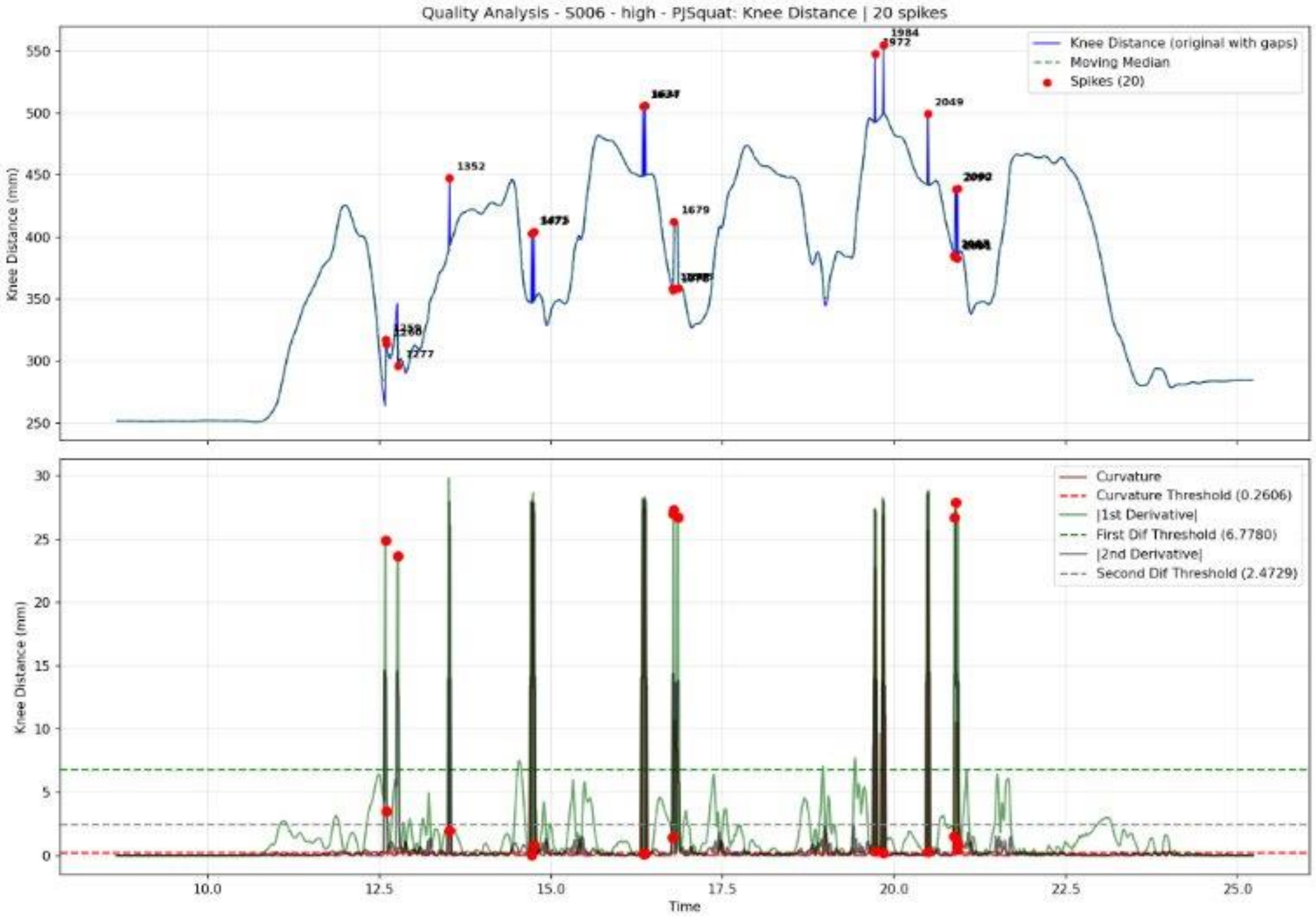
$$R = V^{\top} U \in \mathbb{R}^{3 \times 3}$$

$$t = \bar{q} - R\bar{p}$$

$$SapiensAligned = P_{aligned} = RP + t$$

# Vicon Quality Inspection

- The largest bottleneck of analysis was Vicon annotation
- Automates error detection process
- Detection methods:
    - 1st and 2nd Derivative
    - Curvature
    - Moving absolute median deviation
- Adaptive thresholds
- Indicates anomaly frames on plots
- Modular detection methods and thresholds

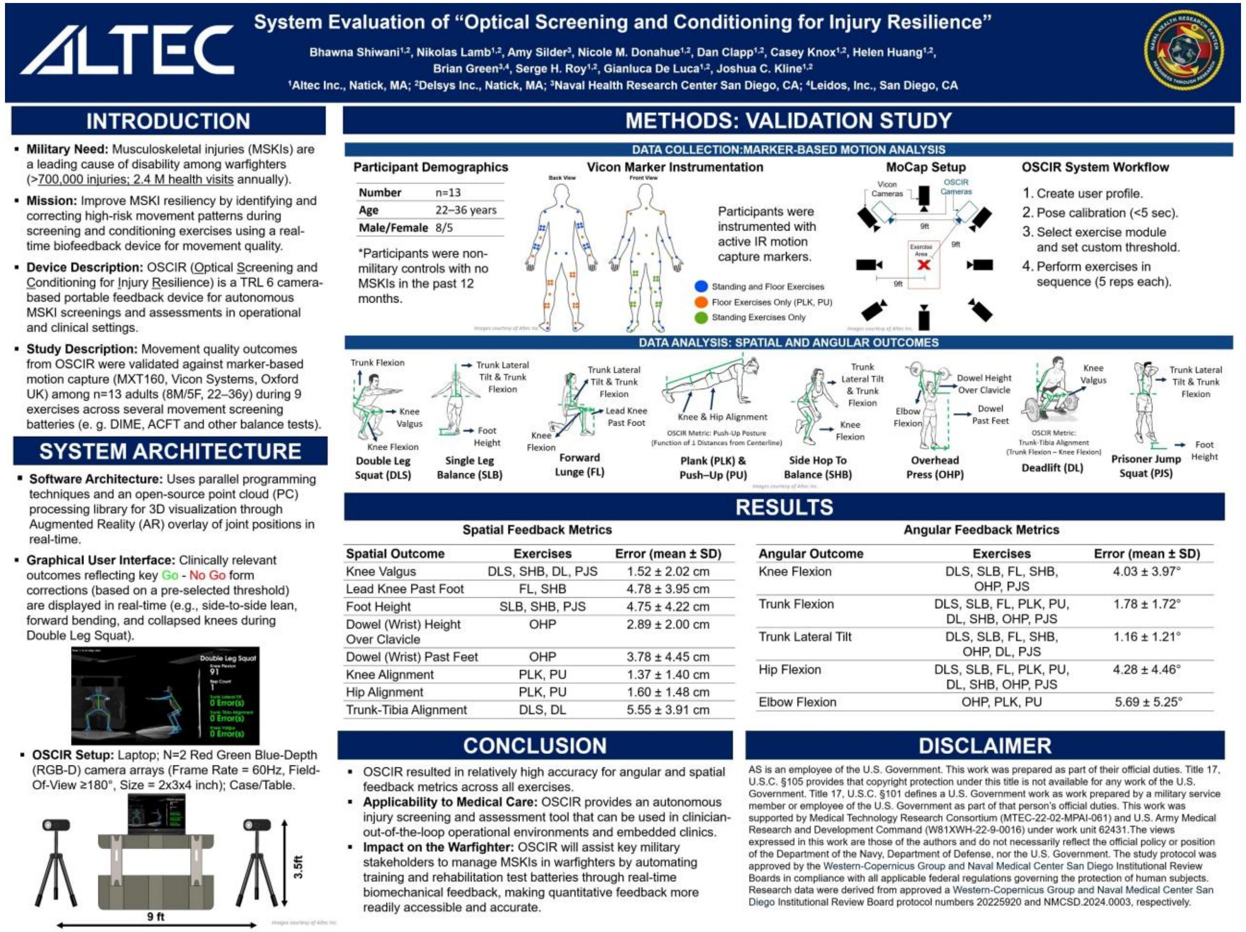


# Results

SECTION 8

# MHSRS Poster

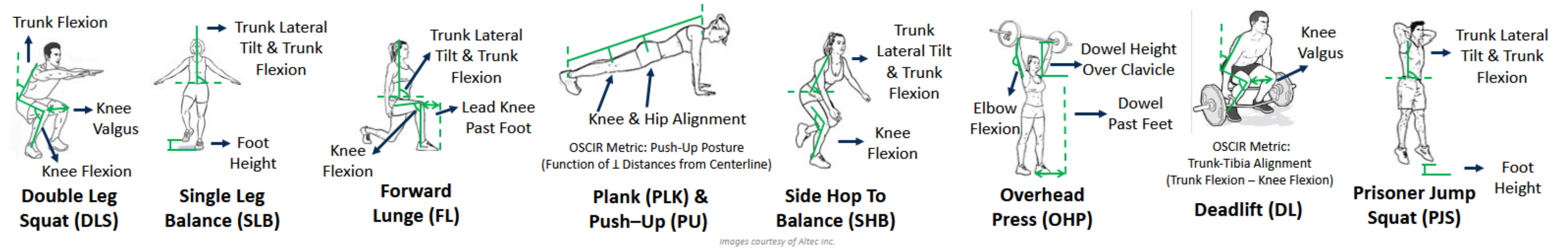
- Work from the study was presented in a poster format at the Military Health System Research Symposium (MHSRS) on August 4 – 7, 2025
- The poster detailed conclusions drawn about the validation of the MEDIC system's accuracy when computing spatial and angular feedback metrics as compared to Vicon ground truth



# Statistics

- Both spatial and angular metrics display relatively low error (< 5° or < 5 cm) across all exercises
- Both spatial and angular metrics display relatively low standard deviation (< 5° or < 5 cm)
- These results convey that MEDIC is accurate when producing outcomes

## DATA ANALYSIS: SPATIAL AND ANGULAR OUTCOMES




## RESULTS

### Spatial Feedback Metrics

| Spatial Outcome | Exercises | Error (mean ± SD) |
|---|---|---|
| Knee Valgus | DLS, SHB, DL, PJS | 1.52 ± 2.02 cm |
| Lead Knee Past Foot | FL, SHB | 4.78 ± 3.95 cm |
| Foot Height | SLB, SHB, PJS | 4.75 ± 4.22 cm |
| Dowel (Wrist) Height Over Clavicle | OHP | 2.89 ± 2.00 cm |
| Dowel (Wrist) Past Feet | OHP | 3.78 ± 4.45 cm |
| Knee Alignment | PLK, PU | 1.37 ± 1.40 cm |
| Hip Alignment | PLK, PU | 1.60 ± 1.48 cm |
| Trunk-Tibia Alignment | DLS, DL | 5.55 ± 3.91 cm |

### Angular Feedback Metrics

| Angular Outcome | Exercises | Error (mean ± SD) |
|---|---|---|
| Knee Flexion | DLS, SLB, FL, SHB, OHP, PJS | 4.03 ± 3.97° |
| Trunk Flexion | DLS, SLB, FL, PLK, PU, DL, SHB, OHP, PJS | 1.78 ± 1.72° |
| Trunk Lateral Tilt | DLS, SLB, FL, SHB, OHP, DL, PJS | 1.16 ± 1.21° |
| Hip Flexion | DLS, SLB, FL, PLK, PU, DL, SHB, OHP, PJS | 4.28 ± 4.46° |
| Elbow Flexion | OHP, PLK, PU | 5.69 ± 5.25° |

ALTEC

# Broader Goals

- Outside of the scope for our summer work, there are some pathways the team intends to explore:
  - **Optimal camera count**
    - How does accuracy degrade as viewpoints are lost? How many cameras are necessary for accurate capture?
  - **Relevance of musculoskeletal constraints**
    - How does imposing musculoskeletal constraints impact body tracking?
  - **Monocular 3D joint localization with a 2D camera**
    - Is it feasible to train a body tracking model to predict 3D joint locations from a single perspective?
  - **Finetune Sapiens**
    - Finetune the base model of sapiens for 3D pose estimation
  - **Kinect replacement**
    - How important is each Kinect data channel? Microsoft no longer supports the Kinect, so are there other options we could use instead?

# Skills Learned

Aidan Kimberley

- Handling large datasets
- Statistical analysis
- Biomechanics calculations
- Algorithm development
- Human motion capture studies

Ethan Matzek

- Managing large multimodal systems
- Bash Scripting
- Debian Packaging and Cmake
- Performing human studies
- Vicon Nexus

Ava Megyeri

- 3D Reconstruction
- Bash Scripting
- Computer Vision
- Algorithm Development

# Acknowledgements

# Thank you Delsys/Altec!

DELSYS

ALTEC