Deep Learning Models Can Detect Muscular Dystrophy from Video-Derived Kinematic Data

Sydney Covitz[1], Parker Ruth[1,2], Tian Tan[1], Constance de Monts[5], Shelby Vogt-Domke[5], Nicholas Bianco[1], Sarah Ismail[5], Lin Karman[5], Julie Muccini[1], John Day[5], Tina Duong[5], Scott L. Delp[1,3,4]

Depts. of [1]Bioengineering, [2]Computer Science, [3]Mechanical Engineering, [4]Orthopaedic Surgery, [5]Neurology, Stanford University, Stanford, CA

Email: covitz@stanford.edu

Summary

We have developed models that can identify the presence of muscular dystrophy—a group of neuromuscular disorders characterized by progressive weakness and degeneration of skeletal muscles—using motion data extracted from video. Our current models obtain test accuracies of 94.0%—fully connected neural network trained on a set of clinically-informed features—and 88.2%—transformer-based architecture trained on kinematic time series data.

Introduction

Measurements of human movement are critical for diagnosing and tracking neuromuscular disorders, yet current clinical practices lack sensitive tools to detect subtle weakness patterns early presentations of these disorders can cause. While researchers have developed methods to identify movement signatures linked to disease, these often rely on expensive technology or time-intensive tests that are not always feasible to conduct in a clinic [1]. OpenCap—a smartphone application that performs video-based, markerless motion capture—can help address this gap by enabling scalable, accessible, out-of-lab biomechanical data collection [2]. Our team has shown that a set of engineered features derived from OpenCap data can identify disease-specific movement characteristics that traditional timed functional tests (TFTs) cannot [3]. We are now developing models that perform the binary classification task of detecting muscular dystrophy.

Methods

We used OpenCap to collect marker positions and joint kinematics for 129 individuals performing nine different upper and lower extremity-based TFTs, including a 10-meter walk, a timed-up-and-go, and a five-time sit-to-stand. Our cohort contains 28 patients with facioscapulohumeral muscular dystrophy (FSHD), 58 patients with myotonic dystrophy (DM)—both subclasses of muscular dystrophy—and 43 healthy controls. OpenCap automatically generates the time series of 3D joint positions and angles, which were sampled at 120 or 60 Hz depending on the task. We built two models to detect muscular dystrophy from video. In the first, we trained a neural network with three fully connected hidden layers on a set of 35 clinically informed, engineered features [3] that we extracted from each trial. In the second, we built a transformer-based model designed to run directly on the raw kinematic time series data. In both models, we applied an 80:10:10 percentage split of the input dataset for training, validation, and test sets, respectively. We applied z-score normalization to the inputs, employed a class weights dictionary to account for class imbalances, and used a binary cross entropy loss function to train each model.

Results and Discussion

Both models are able to learn movement signatures associated with muscular dystrophy, and the transformer architecture does so from the raw kinematic data with minimal preprocessing required, which indicates potential for future uses of minimally preprocessed, video-based movement analysis as a diagnostic tool in neuromuscular populations. The fully connected neural network trained on the clinically informed features set achieves higher test accuracy and recall, and the transformer achieves slightly higher test precision (Table 1). Additionally, our transformer produces consistently higher precision than it does recall, indicating it is better at minimizing false positives than it is at ensuring no true positives are missed. This is consistent with the observed patterns in our dataset, as some of our muscular dystrophy patients have not yet progressed to having severe movement impairments, rendering their disease presence harder to detect.

Table 1: Performance of different architecture and input combinations. The precision and recall values reported are weighted averages between the two classes (disease vs. no disease).

Architecture (input)	Training Accuracy	Test Accuracy	Test Precision	Test Recall
Fully Connected NN (features set)	97.1%	94.0%	91.7%	91.7%
Transformer (kinematic time series)	100.0%	88.2%	92.2%	88.2%

Conclusions

Deep learning architectures trained on video-derived data can differentiate between individuals with muscular dystrophy and healthy controls. Many patients with neuromuscular disorders endure a prolonged diagnostic odyssey; video-based disease detection, especially using minimally preprocessed kinematic time series, could help augment existing clinician-graded assessments of function, remote monitoring, and diagnostic tools. Future work includes further hyperparameter tuning of our transformer model to improve test accuracy, adapting our models for multi-class neuromuscular disorder detection, and continuing to expand our input dataset to produce a larger, more diverse set of training examples.

References

[1] Jaspers et al. (2011). *Gait & Posture*, **34:** 227-233.
[2] Uhlrich SD et al. (2023). *PLOS Comp. Bio*, **19.**
[3] Ruth P et al. (2024). BioRxiv. Under review at *New Eng. J. Med*