UMUD: A Web Application for Easy Access to Musculoskeletal Ultrasonography Datasets

Paul Ritsche[1], Fabio Sarto[2], Francesco Santini[3,4], Christoph Leitner[5], Martino V. Franchi[2], Oliver Faude[1], Taija Finni[6], Olivier Seynnes[7], Neil Cronin[6,8]

[1] Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland
[2] Department of Biomedical Sciences, University of Padova, Padova, Italy
[3] Basel Muscle MRI, Department of Biomedical Engineering, University of Basel, Basel, Switzerland
[4] Department of Radiology, University Hospital of Basel, Basel, Switzerland
[5] Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland
[6] Faculty of Sport and Health Sciences, University of Jyvaskyla, Jyvaskyla, Finland
[7] Department for Physical Performance, Norwegian School of Sport Sciences, Oslo, Norway
[8] School of Education & Science, University of Gloucestershire, Gloucester, UK
Email: paul.ritsche@unibas.ch

Summary

We developed the Universal Musculoskeletal Ultrasonography Database (UMUD), a web platform for sharing and standardizing musculoskeletal ultrasound datasets. It indexes metadata from public datasets, enabling easy search and exploration.
So far, UMUD indexes 11 datasets with 75,569 images and 2,573 videos from 1,769 participants. It includes benchmark datasets for trainee training and validating image analysis algorithms. By centralizing data, UMUD promotes transparency, reproducibility, and innovation in ultrasound research.

Introduction

Ultrasonography is widely used to assess skeletal muscle and tendon properties, such as architecture, cross-sectional area, and tissue stiffness [1]. Despite its growing application in different scenarios, and the increasing call for open data access and sharing in clinical research, there remains a significant scarcity of public datasets in this field. This lack of accessible and standardized public datasets limits large-scale analysis algorithm studies, trainee training and the development of image analysis algorithms. To address this, we developed the Universal Musculoskeletal Ultrasonography Database (UMUD), a web application designed to facilitate access to these datasets and foster standardization and innovation in musculoskeletal ultrasonography imaging research.

Methods

UMUD is an online repository that aggregates and indexes metadata from publicly available musculoskeletal ultrasonography datasets hosted on platforms like the Open Science Framework and Zenodo. The web application (https://universalmuscledatabase.streamlit.app/) is built using a streamlit (v1.35.0) frontend and mongoDB for its database infrastructure. Standardized metadata descriptors, i.e., muscle group, ultrasound device, participant demographics, are implemented using a combination of pydantic models (v1.10.0) and json schemata to ensure reproducibility and ease of usage for contributing data. UMUD provides detailed instructions for community contributions, including tools for data anonymization.

Results and Discussion

So far, UMUD includes 11 Datasets from 10 studies, including 75.569 images (and/or 2.573 videos) derived from 1.769 participants. A total of nine (lower limb) muscles and one muscle-tendon junction (distal triceps surae and achilles tendon) assessed with different acquisition methods (static images, videos of muscle contraction or tendon movement, or three dimensional muscle reconstructions) are included. Datasets were acquired in different regions of the respective muscles, ranging from proximal to distal muscle parts. UMUD also includes datasets for trainee training and benchmark datasets for training, validating and comparing image analysis algorithms. These comprise a dataset containing multi-expert manual analyses of 60 muscle architecture and panoramic cross-sectional area images, a dataset containing 125 images with overlays of muscle geometry for trainee training, and two fully labeled datasets for image analysis algorithm training. Using these benchmark datasets, UMUD aims to compare available image analysis algorithms to rate their overall performance. Additionally, UMUD already lists state-of-the-art automated analysis algorithms, as a guide of usage for the community.

Conclusions

In conclusion, UMUD addresses relevant challenges in musculoskeletal ultrasonography by providing a centralized, standardized repository of datasets and tools. It promotes transparency and innovation in the field, supporting reproducible research and advancements in automated image analysis. Future developments include adding datasets, expand functionalities and introducing community-driven algorithm development challenges.

References

[1] Sarto F. et al. (2021). *Sports Med*, **51**:1151–1170