LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

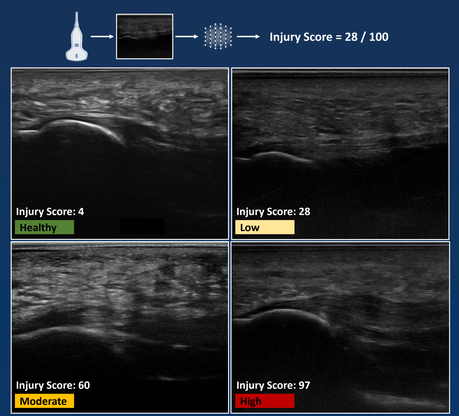
A Framework for Investigating Objective Injury Severity Modeling

Brian S. Baum[1], Benjamin W. Roop[1], Kevin J. Brady[1], Dan M. Cushman[2], Lars A. Gjesteby[1]

[1] Human Health & Performance Systems, MIT Lincoln Laboratory, Lexington, MA

[2] Department of Physical Medicine & Rehabilitation and Orthopedics, The University of Utah School of Medicine, Salt Lake City, UT

Contact:
Brian S. Baum, PhD
MIT Lincoln Laboratory
brian.baum@ll.mit.edu



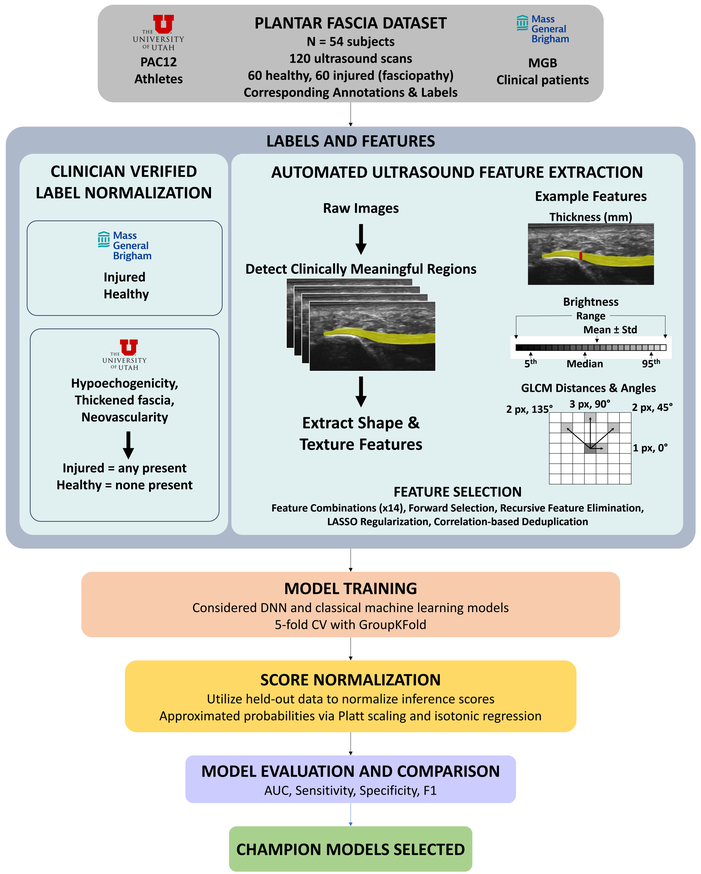
Concept of generating injury scores based on objective features from ultrasound images

Sponsored by DHA Research, Technology, and Innovation

Introduction

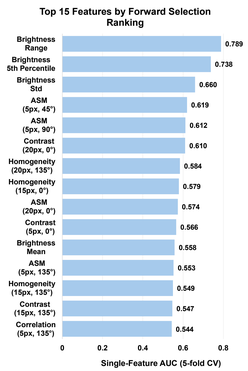
- Musculoskeletal injury assessment lacks standardized severity scales for some tissues, limiting risk stratification and return-to-duty (RTD) prediction
- Binary classification (healthy vs. injured) provides insufficient granularity for monitoring and clinical decisions
- An experimental framework for injury diagnosis can identify promising scoring scales and objective ultrasound features for modeling severity
- Experimental Framework
  - Develop machine learning (ML) models to assess injury from ultrasound images
  - Calibrate continuous scores with held out data to mitigate small dataset size problem
  - Characterize candidate feature importance towards severity prediction

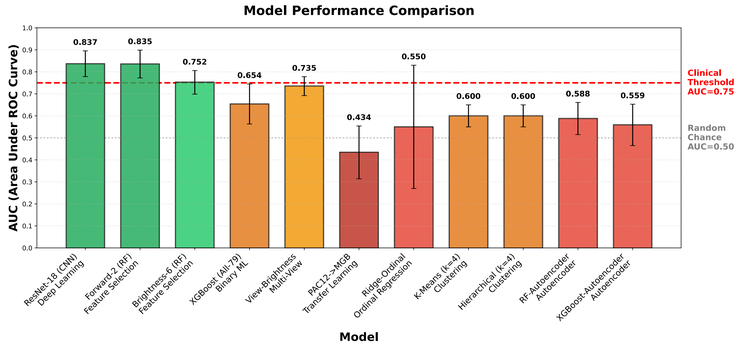
Method and Materials



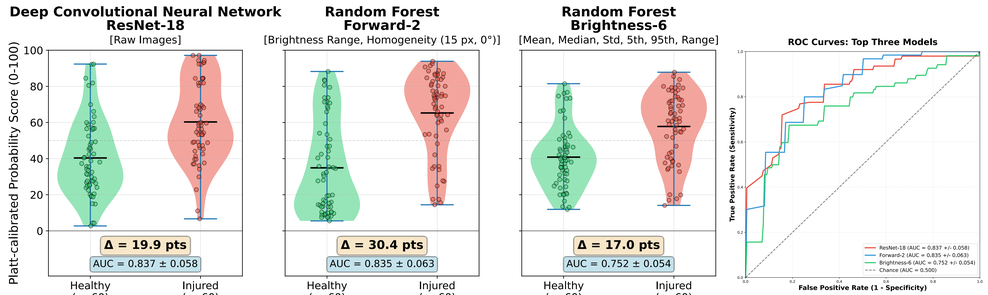
5th: 5th Percentile; 95th: 95th Percentile; ADL: Activities of Daily Living; ASM: Angular Second Moment; AUC: Area Under the Curve; CNN: Convoluted Neural Network; CV: Cross-Validation; DNN: Deep Neural Network; GLCM: Gray-Level Co-Occurrence Matrix; LASSO: Least Absolute Shrinkage and Selection Operator; LogReg: Logistical Regression; MLP: Multi-Layer Perceptron; MGB: Mass General Brigham; P: Probability; PAC12: Pacific Athletic Conference 12; PCA: Principal Component Analysis; px: Pixels; ResNet-18: Residual Network 18 layers; RF: Random Forest; RFE: Recursive Feature Elimination; RTD: Return to Duty; ROC: Receiver Operating Characteristic; Std: Standard Deviation; SVM: Support Vector Machine; XGBoost: eXtreme Gradient Boosting

Results





Forward selection ranking (left) produced the top performing model (exemplar data from Random Forest modeling shown). Comprehensive model comparisons (subset shown on right) identified three models that achieved clinically viable binary classification performance (AUC ≥0.75).



	ResNet-18	Forward-2	Brightness-6
Sensitivity:	63.3%	78.3%	70.0%
Specificity:	71.7%	71.7%	76.7%
F1 Score:	0.661	0.758	0.724

The top three models demonstrated meaningful separation between healthy and injured groups (Δ), calculated from the mean probability scores.

Discussion and Conclusions

Key Findings

- The proposed experimental framework provided insights for the injury score prediction problem
- Ultrasound features achieved meaningful separation for binary injury classification
- Objective brightness metrics successfully automated detection of echogenicity, which is used in clinical labeling, emerging as promising features in binary classification
- Weak correlations with FAAM scores ($\rho \leq 0.25$) highlight disconnect between objective structural pathology and subjective functional disability

Open Questions

- Will the features that predict binary injury classification successfully characterize severity?
- Do objective brightness metrics align with clinician determined echogenicity?
- What are benefits / limitations of additional patient-reported clinical inventories in characterizing injury diagnosis and severity?

Future Work

- Engage clinicians to develop definitions for non-standardized severity scales
- Investigate effectiveness of predicting injury severity from ultrasound objective measures using the experimental framework
- Extend MIT Lincoln Laboratory's deployable assessment framework
  - From binary injury diagnosis prediction to severity assessment
  - To monitor healing, predict risk stratification and RTD timelines