GymbrAIn: Artificial Intelligence at the Service of Your Muscles

Casali Cristian Flotta Aldo Lamberti Francesco

University of Modena and Reggio Emilia, Italy

November 7, 2024

Abstract

- Problem: Increasing prevalence of incorrect exercise form in gyms
- Solution: GymbrAIn - AI system for real-time form correction
- Key Features:
  - Advanced pose estimation
  - Self-attention mechanisms
  - Real-time feedback
- Results: Competitive accuracy in identifying form discrepancies

Introduction

- Social media's impact on fitness trends
- Challenges:
  - Limited access to professional guidance
  - Cost barriers
  - Risk of injury from improper form
- Solution Overview:
  - AI-based computer vision technology
  - Real-time exercise analysis
  - Immediate form feedback

Related Work

- Previous Research [4] (ICIT):
  - Comparison of neural network architectures
  - GRU: 97.27% accuracy in training
  - LSTM: 95% accuracy in training
- Our Innovations:
  - Enhanced GRU architecture
  - Self-attention layers
  - Optimized dropout parameters

Dataset

- Source: Kaggle [1]
- Data Augmentation:
  - Horizontal translations ($\pm$100 pixels)
  - Mirror transformations
- Dataset Composition:
  - 112 bench press videos
  - 104 barbell biceps curl videos
  - 90 triceps push down videos
  - 94 lat pull down videos

Methods

- Preprocessing:
  - Gaussian filter for noise reduction
  - Sharpening filter for clarity
- Pose Estimation:
  - YOLO POSE [3] implementation
  - Real-time skeletal tracking
- Classification Architecture:
  - GRU-based processing
  - Self-attention mechanisms
  - 256 neurons, 0.25 dropout rate
- Video retrieval and geometric analysis

Preprocessing Stage

Two-Phase Filtering Process

- **Gaussian Filter**
  - Suppresses random noise
  - Preserves essential movement information
- **Sharpening Filter**
  - Enhances edge definition
  - Improves human figure clarity

Pose Estimation

YOLO POSE [3] Implementation

- Real-time skeletal tracking
- High-precision anatomical landmark identification
- Temporal consistency maintenance
- Mobile-optimized computation

Exercise Classification

Neural Network Architecture

- GRU-based sequential processing
- Self-attention mechanisms
- 256 neurons in hidden layers
- Dropout rate of 0.25
- 4-neuron output layer with softmax activation

Video Retrieval & Geometrical Analysis

Process Flow

1. **Movement Analysis**
   - Velocity minima detection
   - Identification of critical phases
2. **Reference Video Retrieval**
   - L2 norm comparison
   - Expert form matching

Video Retrieval & Geometrical Analysis (cont.)

Process Flow

3. **Comparative Analysis**
   - Joint angle extraction
   - Temporal alignment
   - Direct comparison
4. **Feedback Generation**
   - Angular difference calculation
   - Threshold-based correction

Test Classification Results

Architecture	Neurons	Dropout	Self-Attention	Accuracy (%)
GRU	256	0.5	No	77
GRU	512	0.5	No	67
GRU	512	0.5	Yes	35
GRU	256	0.5	Yes	77
GRU	256	0.25	Yes	80

Training Results - Original Architecture

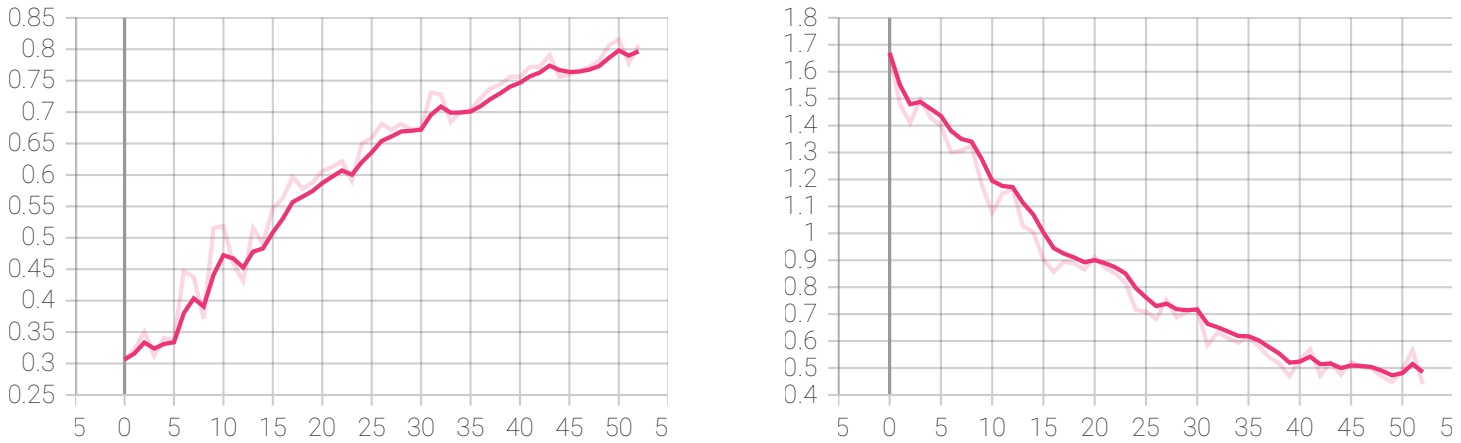

Figure: GRU with 256 neurons and dropout = 0.5

Training Results - Final Architecture

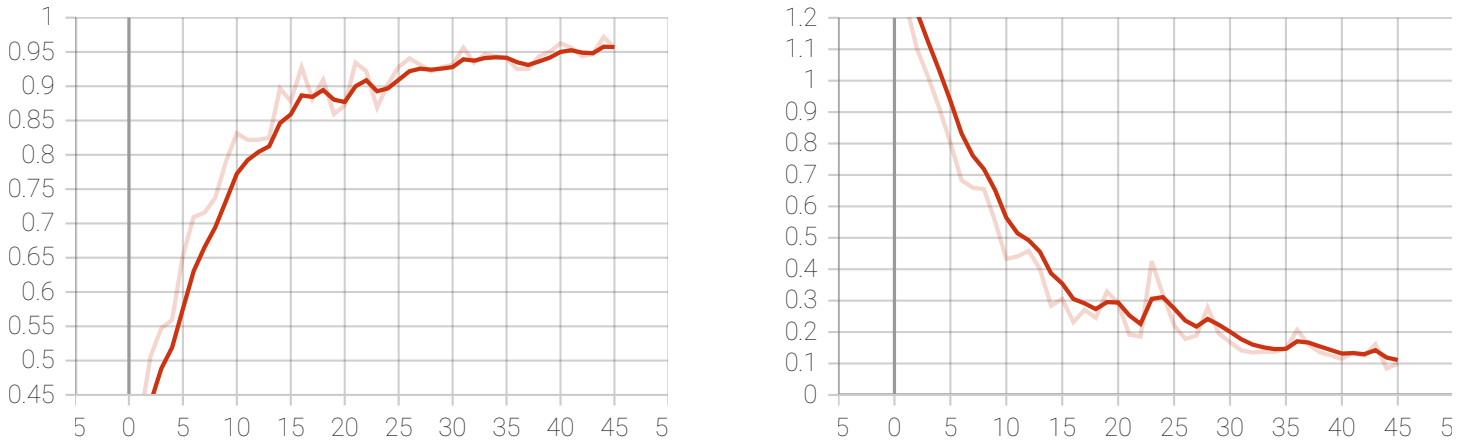

Figure: GRU + Self-Attention with 256 neurons and dropout = 0.25

Example: Reference Bench Press

Figure: Reference Form - Starting and Ending Positions

Example: Incorrect Form Detection

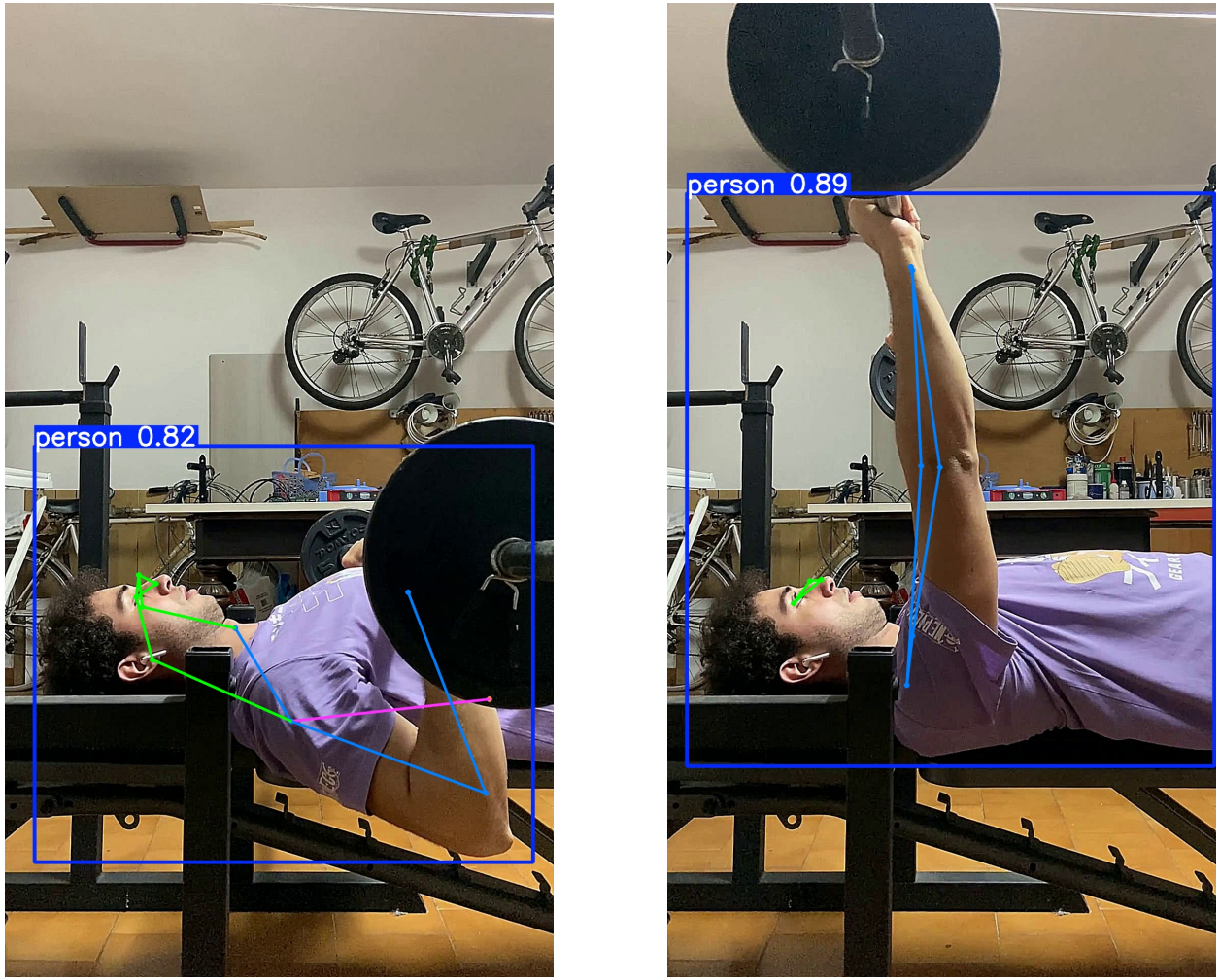

Figure: Incorrect Form Detection (27° difference from reference)

Example: Correct Form Detection

Figure: Correct Form Execution

Limitations & Future Work

- Current Limitations:
  - Manual camera calibration [6]
  - Dataset size
  - Classification accuracy
- Future Improvements:
  - Automated calibration (DeepFocal [5], DeepCalib[2])
  - Dataset expansion
  - Mobile optimization
  - Clinical applications

Conclusion

- Achievements:
  - Functional exercise analysis system
  - 80% classification accuracy
  - Real-time form correction
- Impact:
  - Accessible fitness guidance
  - Potential for rehabilitation applications
  - Foundation for future development

References I


Hasyim Abdillah.
Kaggle.
https://www.kaggle.com/datasets/hasyimabdillah/workoutfitness-video.


Oleksandr Bogdan, Viktor Eckstein, Francois Rameau, and Jean-Charles Bazin.
Deepcalib: a deep learning approach for automatic intrinsic calibration of wide field-of-view cameras.
In *Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production*, CVMP '18, New York, NY, USA, 2018. Association for Computing Machinery.


Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi.
You only look once: Unified, real-time object detection, 2016.

References II

Sijie Shang, Rong Jin, and Kevin Desai.
A study of human fitness pose classification using artificial neural networks.
In *2023 International Conference on Information Technology (ICIT)*, pages 250–255, 2023.

Scott Workman, Connor Greenwell, Menghua Zhai, Ryan Baltenberger, and Nathan Jacobs.
Deepfocal: A method for direct focal length estimation.
In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 1369–1373, 2015.

Z. Zhang.
A flexible new technique for camera calibration.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330–1334, 2000.

Thank You

Questions?

{284312, 282955, 272361}@studenti.unimore.it