iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VI Month of publication: June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72783

AI-Based Health Risk Prediction Using Real-Time Pose Estimation and Deep Learning

Suhas Revansiddha Amane
UG Student, Department of Computer Science and Engineering, Brahmdevdada Mane Institute of Technology, Solapur, Maharashtra, India

Abstract: Preventive healthcare needs early health risk detection of conditions like obesity, malnutrition, and metabolic disorders. Conventional screening procedures are either expensive, demand trained staff, and physical visits, hence less accessible and scalable. This paper introduces a real-time, AI-powered health risk prediction system that employs computer vision and deep learning to gauge physical health with a webcam alone.
The system proposed utilizes MediaPipe and OpenCV to infer body pose landmarks from live video feed. Anthropometric features like height, BMI, shoulder width, arm length, and leg length, as well as waist-to-hip ratio, are automatically computed from these. These values, together with user-supplied information such as age and gender, are input into a deep learning model that has been trained to predict an individual as being in one of four categories: healthy, risk of obesity, malnourishment, or risk of metabolic disorder. All processing is carried out in real time and aggregated into a web-based interface, providing instant feedback, graphical insights, and health suggestions to the user. The backend is implemented with Python, TensorFlow/Keras, and Flask for rapid processing and model inference. This system illustrates the promise of affordable, non-invasive technologies for initial health testing. It can be used in telemedicine, fitness programs, and distant health screening, increasing preventive care to be more scalable and affordable.
Keywords: Health risk prediction, deep learning, pose estimation, BMI, MediaPipe, OpenCV, real-time screening, body measurements

I. INTRODUCTION

The worldwide epidemic of non-communicable diseases (NCDs) like obesity, metabolic conditions, and malnutrition poses a significant challenge to public health systems. These conditions are most commonly linked with modifiable lifestyle risk factors and, significantly, are preventable through early screening and timely intervention. Nonetheless, traditional health assessments are usually based on costly equipment and skilled clinical personnel, which severely hampers their scalability and accessibility—particularly in rural or under-resourced settings.

Emerging computer vision and artificial intelligence techniques have unlocked new horizons in non-invasive healthcare monitoring. By virtue of the ubiquitous presence of webcams and smartphones, human pose estimation (HPE) has emerged as a viable technique for the extraction of bodily and biomechanical information from video streams [1][12][13]. Methods like MediaPipe and OpenPose enable precise real-time identification of skeletal landmarks, which can be translated to anthropometric characteristics such as height, limb lengths, waist-to-hip ratio, and shoulder width [2][4][7]. These measurements are commonly utilized in clinical practice to assess obesity risks, malnutrition, and metabolic syndrome risks [6][8][9].

Concurrently, deep learning has exhibited robust predictive power in biomedical domains, such as BMI prediction [7][14][16], body fat measurement [19], and health inference based on gait [5][6][15]. Experiments have established that convolutional and fully connected networks can accurately classify physical health states from well-structured features like body ratios and demographic features like age and gender [17][22][23]. In addition, recent pose-based systems have been demonstrated to reliably detect ergonomic and postural hazards from low-level RGB inputs [10][11][18].

This work suggests a comprehensive, web-based health evaluation system that integrates pose estimation with deep learning to forecast physical health threats in real time. Utilizing an ordinary webcam, the system extracts skeletal landmarks, calculates important anthropometric measures, and employs a learned deep neural network to predict users into health categories such as obesity threat, malnourishment, metabolic disorder threat, or healthy. In contrast with customary methods, this platform does not demand any physical tools, guaranteeing cost-adequacy and applicability across diverse environments.

The research makes a contribution to accessible health screening by showing how low-cost vision systems can provide personalized health testing, particularly for telemedicine, wellness monitoring, and initial health diagnostics [3][21][25].

II. RELATED WORK

Our work is based on major advances in computer vision, deep learning, and biomedical data analysis. The project studies various areas that contribute to its foundation, such as pose estimation systems, anthropometric measurement methods, and health prediction models. The literature below is in favor of developing our AI-based system for health risk assessment.

A. Advances in Deep Learning for Health Prediction

Past research in healthcare has shown the potential of deep learning models to predict chronic diseases. Pantanowitz et al. (2019) predicted Body Mass Index (BMI) from silhouette pictures with the help of convolutional neural networks (CNNs), providing a consistent, image-based screening tool [7]. Ribeiro et al. (2024) described an in-depth review on the application of deep learning for diagnosing obesity and its possibility in both prediction and tailored treatment [8]. Islam et al. (2022) employed a hybrid deep model on visual and structured information for the prediction of overall physical and mental health hazards, exemplifying the robustness of neural models in biomedical applications [22].

B. Pose Estimation and Extraction of Body Measurements

Pose estimation systems like MediaPipe and OpenPose enable precise real-time landmark detection from 2D video [1][12]. Škorvánková et al. (2021) showed that pose information can be employed to derive anthropometric measurements such as height, shoulder breadth, and waist-hip ratio from webcam video with acceptable precision [1]. Fadllullah et al. (2024) proposed a deep learning system that predicts human height from camera input and linear regression methods, highlighting the increasing relevance of camera-based biometric assessment [17].

C. Anthropometry and Risk Indicators of Health

Anthropometric measures like BMI, waist-hip ratio, and limb ratio have been utilized over the years to determine risks for obesity, malnutrition, and metabolic syndrome. Research by Kim et al. (2023) indicated the viability of calculating the percentage of body fat from 2D camera inputs, a useful procedure for screening health without medical equipment [19]. Nguyen et al. (2025) suggested 3D body measurement through stereo vision to enhance precision in physical danger estimation [18]. Such studies support the clinical significance of body form and shape in anticipating health condition.

D. Gait and Postural Analysis Using Vision

Gait analysis is becoming increasingly accepted as a measure of physical health. Li et al. (2022) created a pose-guided model that was able to predict health markers including BMI and age from the gait patterns seen on video [5]. Markerless gait monitoring has been employed to identify muscular imbalance and postural symmetry, which are associated with musculoskeletal hazards and chronic fatigue [11][15]. These applications confirm the effectiveness of vision-based biomechanical analysis in preventive medicine.

E. Real-Time Monitoring and User-Centric Interfaces

The capability to make real-time predictions from a webcam alone opens up new avenues for health screening. Heliyon (2024) showed how pose estimation might be employed for real-time feedback in movement training [9], whereas Wahid et al. (2023) investigated applying AI vision systems to ongoing physical health monitoring [11]. Both of these systems are closely aligned with the objectives of the present work, providing affordable, non-invasive means of health assessment and personalized advice.

III. SYSTEM ARCHITECTURE

The system architecture of the suggested AI-powered health risk forecasting platform is intended to support non-invasive, real-time health measurement through a webcam. It combines computer vision, anthropometric measurement extraction, prediction based on deep learning, and user-oriented interfaces. It prioritizes modularity, scalability, and real-time responsiveness so users can self-evaluate physical health risks without clinical equipment or appointments.

A. Overview

The system has the following primary components:

- Real-Time Pose Capture through Webcam
- Landmark Extraction and Feature Engineering

- Deep Learning-Based Health Prediction Engine (DNN Model)
- Frontend Web Interface (HTML + Bootstrap + JS)
- Backend Server (Flask with TensorFlow/Keras)
- Visualization Dashboard for Results
- Optional Email Report Delivery through SMTP

These elements communicate through RESTful APIs, enabling asynchronous and modular communication between frontend and backend layers.

B. Real-Time Pose Acquisition

1) Webcam Video Input

Users enable their webcam via the web interface. The webcam records the user standing upright, preferably within a calibrated boundary to optimize exposure of important body joints. Pose tracking is performed directly in the browser by MediaPipe's Pose model [1][12].

2) Landmark Detection and Body Measurement

MediaPipe extracts 33 skeletal landmarks (e.g., shoulders, hips, knees, ankles), from which the system calculates:

- Height (approximate from foot-to-head distance)
- Shoulder width
- Arm and leg lengths
- Waist-to-hip ratio (horizontal landmark distances)
- Gait stability (when movement occurs)

Previous research has established the viability of obtaining robust anthropometric information from pose estimation models [1][2][3].

C. Feature Engineering and Preprocessing

After detecting body landmarks, important health-related features are derived and formatted in a model-friendly manner:

- Estimated height and proportions of limbs
- Derived BMI (if weight is known)
- Waist-to-hip ratio
- Age and gender (keyed manually)

The preprocessing pipeline involves:

- Normalization of numerical features
- One-hot encoding of categorical variables (e.g., gender)
- Validation of landmark completeness
- Management of non-compulsory inputs (manual override weight/height)

These features are the input vector to the prediction model [4][7][9].

D. Prediction Engine: Deep Neural Network (DNN)

A deep neural network (DNN) fully connected is employed for health risk classification:

- Input Layer: Anthropometric and demographic attributes
- Hidden Layers: 2–3 layers with ReLU activation and dropout
- Output Layer: Softmax layer generating 4-class output:
- Healthy
- Obesity Risk
- Malnourishment
- Metabolic Disorder Risk

Training was conducted with a labeled dataset of more than 2,000 health records, employing 80/20 train-test split, categorical cross-entropy loss, and Adam optimizer.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

Model Performance:
- Accuracy: 92.4%
- F1-Score: 0.91

Low overfitting as a result of regularization techniques [7][8][16].

E. Frontend and Client Interface
The frontend is made with HTML, JavaScript, and Bootstrap and includes:
- Live Posture Following Window
- User Input Board (age, sexual orientation, weight)
- Health Hazard Outline (result cards with hazard level and suggestions)
- Visual Criticism (highlighted posture joints and measurements)

It gives responsive plan for desktop and portable get to, and employments Bootstrap classes to actualize grid-based styling and layout.

F. Mail Notice through SMTP
The clients can select to be sent a outline report by e-mail after running the expectation. The backend sends the result through SMTP, which includes:
- Health chance category
- Actual/estimated body measurements
- Personalized suggestions and future actions

This gives post-evaluation availability, especially helpful for wellbeing following over time [20].

G. Comes about Dashboard and Visualization
The dashboard gives visual criticism within the shape of charts and symbols for:
- Real-time hazard classification
- Confidence scores
- Comparison with ordinary anthropometric ranges
- BMI vs age/gender diffuse visual (optional)

The plan guarantees interpretability, especially for clients with negligible specialized foundation [6][13][21].

This design empowers clients to conduct moment wellbeing screening with fair a webcam, closing the crevice between preventive care and reasonableness. The usage of computer vision coupled with prepared AI models ensures exactness and openness in surveying wellbeing risks.

IV. METHODOLOGY
This framework takes a vision-based methodology to distinguish wellbeing dangers through the extraction of real-time anthropometric and postural characteristics from webcam video. The strategy is based on a consecutive pipeline that incorporates information collection, preprocessing, highlight designing, demonstrate preparing, forecast, and end-user result transmission.

A. Information Collection
The framework collects biometric and statistic data through two primary channels:
- Webcam Video Feed: MediaPipe Posture is utilized to return 33 skeletal points of interest from a live webcam bolster. A few of these obvious joints are the eyes, shoulders, hips, knees, and lower legs. These points of interest are utilized as the establishment for calculating body estimations such as bear width, waist-hip proportion, and tallness [1][2].
- User Inputs: Manual inputs are age, sexual orientation, and alternatively weight or tallness. These include exactness to the forecast demonstrate by giving pose-derived gauges with genuine information on the off chance that available.

The past inquire about has confirmed the need of combining pose-based estimation with statistic inputs for exhaustive wellbeing modeling [3][7].

B. Information Preprocessing and Include Extraction

After posture points of interest and client inputs are gotten, the framework does the following: • Calculates anthropometric features:

- Estimated tallness (from y-distance between noteworthy joints)
- Shoulder width (cleared out and right bear distance)
- Waist-hip proportion (from pixel distances)
- Limb length (arm and leg fragment approximations)
- cleans and normalizes data:
- Missing areas are ascribed with default values or omitted
- Features are Z-score normalized
- Gender is one-hot encoded

This stable include set is prepared for profound learning classification [4][9].

C. Show Development

A Profound Neural Organize (DNN) was chosen since it can capture complicated designs in biometric features.

- Input Highlights: Anthropometric information + statistic input
- Model Sort: Completely associated feedforward neural network
- Output Classes:
- Healthy
- Obesity Risk
- Malnourishment
- Metabolic Clutter Risk

The show was prepared on a well-structured dataset with more than 2,000 labeled tests. Preparing included categorical cross-entropy misfortune, Adam optimizer, and dropout regularization to avoid overfitting. 10-fold cross-validation was utilized to degree execution with 92.4% precision and 0.91 F1-score [7][8].

D. Backend Implementation

Everything model-related rationale and deduction is exhausted Python with TensorFlow/Keras. The backend:

- Processes include vectors gotten through REST API
- Pitches information into the prepared DNN show for prediction
- Returns comes about in JSON arrange back to the frontend
- Optionally makes and sends out a wellbeing report by SMTP [20]

The backend workflow of the framework was motivated by AI-based biometric frameworks utilized in telehealth considers [21].

E. Client Criticism and Interaction

After expectation, comes about are displayed through:

- On-Screen Visualization: The dashboard shows chance levels anticipated, highlight examination, and wellbeing suggestions in genuine time.
- Email Rundown Report (optional): Users are sent a arranged report by mail with wellbeing categorizations and prescribed actions.

This dual-feedback approach guarantees clients can translate their wellbeing results in both prompt and follow-up settings [13][25].

V. IMPLEMENTATION

The architecture is designed as a web application that combines computer vision and machine learning for real-time prediction of health risks. The overall framework includes three central layers: the frontend for user interaction, the backend for model computation and data interpretation, and an optional admin dashboard for performance monitoring and analytics.

A. Frontend Development

The frontend is built on HTML5 and JavaScript, with Bootstrap being used for responsive design. This enables the interface to adapt from desktops to mobile devices. The main interaction is through a live video stream where users are asked to stand in front of the webcam. MediaPipe's pose detection framework runs right in the browser, taking skeletal landmarks from the image of the user.

The main user interface components are:

- Live Pose Feed: Skeletal overlay in real-time viewed on the video screen.
- Input Fields: Fields for capturing user-entered information like age, gender, and optionally height and weight.
- Output Display: The system displays health classification output along with corresponding body measurements and confidence levels.

The frontend is lightweight and browser-based, and no installation is required, making it easily accessible.

B. Backend Integration

The backend is written in Python, where Flask is used as the web framework for routing API calls and facilitating communication between the interface and the machine learning model. Upon detecting a pose and computing body features, the system publishes this information to the server for processing.

Backend duties include:

- Feature Computation: Calculating metrics such as height approximation, waist-hip ratio, and limb lengths from landmark coordinates.
- Model Inference: An optimized deep neural network (DNN) implemented in TensorFlow classifies the processed data and provides a health risk category.
- Optional Email Report Generation: On demand, the report of the user's classification and measurements is emailed via a secure SMTP protocol.

The backend provides real-time processing while ensuring data security and scalability.

C. Administrative Dashboard

There is an optional Power BI dashboard that can be used in administrative purposes. It displays trends and usage behavior, providing metrics such as:

- Health risk category distribution by users
- Measurement averages and demographic splits
- User engagement metrics, e.g., completed assessments

The dashboard can be useful for public health programs or big research deployments.

VI. RESULTS

A. Model Evaluation

The deep learning model was trained on a dataset with 2,000 labeled examples across four health classes: Healthy, Obesity Risk, Malnourished, and Metabolic Disorder. An 80:20 split was used for training and validation, and cross-validation to ensure consistency.

Performance metrics attained:

- Accuracy: 92.4%
- Precision: 90.1%
- Recall: 91.3%
- F1 Score: 90.7%
- AUC-ROC Score: 0.95

The above metrics affirm the robustness and effectiveness of the model in classifying health risk from pose and demographic inputs.

B. Output and Visualization

After analyzing, the system produces a result that contains:

- Risk classification label
- Body measurements based on pose landmarks
- Confidence percentage
- Health advice (e.g., weight control, posture improvement)

Results are displayed immediately through the browser interface. Optionally, an amalgamation PDF report can be emailed to the user's mailbox.

C. User Scenario Demonstrations

- Scenario A: A user displays high waist-hip ratio and estimated BMI. The model rates this as Obesity Risk with 93% confidence, suggesting dietary changes and more activity.
- Scenario B: A user displays upright posture and normal proportions. The model produces a Healthy status with 96% confidence and no need for corrective action.

These use cases indicate the platform's usefulness in delivering quick, convenient health information with minimal input.

VII. DISCUSSION

A. System Strengths

- Non-Invasive Measurement: The system allows users to conduct a health check using just a webcam, eliminating the necessity of wearable sensors, manual measurement tools, or laboratory testing. This enhances accessibility, especially in areas with limited resources.
- Real-Time Feedback:Users are given prompt feedback in addition to tailored recommendations based on their anthropometric and demographic data. This functionality allows users to take proactive steps toward health enhancement without the need for clinical sessions [1][7].
- Scalability and Deployment: The system is constructed with modular building blocks—MediaPipe for pose detection, TensorFlow for model prediction, and Flask/Bootstrap for the web application—which renders it extremely scalable. Deploying it on cloud platforms with RESTful APIs makes horizontal scaling as need arises very straightforward [3][6].
- User-Centered Interface: The interface is also made simple, responsive, and accessible. Users are able to see risk classifications, view visual pose information, and obtain downloadable reports without needing technical knowledge [13][25].

B. Limitations

- Pose Estimation Constraints: Accuracy depends on webcam quality, lighting conditions, and user posture. Suboptimal positioning may affect measurement accuracy and prediction outcomes [1][12].
- Demographic Bias in Training Data: The dataset used may not represent all ethnic, age, or gender groups equally, potentially introducing bias in predictions. Future iterations must include more diverse samples to enhance fairness [8][20].
- Manual Data Entry Dependency: Although most features are computerized, the system continues to need user input for gender and age. Manual entry errors may decrease prediction accuracy.
- Privacy Issues: Even though the system does not retain user video data, features such as report emailing or future cloud syncing need to provide data encryption and informed user permission to meet privacy requirements [21].

C. User Testing and Feedback

Pilot testing with 35 respondents between ages 20–40. Major findings:

- 88% considered the pose-based measurements easy to use and informative about their body structure and related risk.
- 84% felt confident about the results and suggestions.
- Users asked for additional features:
- Real-time chatbot for health Q&A
- Multilingual support
- Integration with wearable fitness devices

Feedback indicates wide acceptance and directs priorities for system improvement.

VIII. CONCLUSION

This work presented a real-time AI-powered health risk assessment system integrating pose estimation, anthropometry measurement extraction, and deep learning-based classification into a single non-invasive framework. With the use of common webcams and clean frontend design, the system allows for easy screening of health risks like obesity, malnourishment, and metabolic syndrome with minimal effort.

Model evaluation showed excellent classification accuracy (92.4%) and stable performance across test samples, validating the applicability of computer vision in preventive healthcare. The combination of a dynamic frontend, backend inference automation, and email reporting as an option provides end-to-end user experience—from pose tracking to actionable knowledge.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

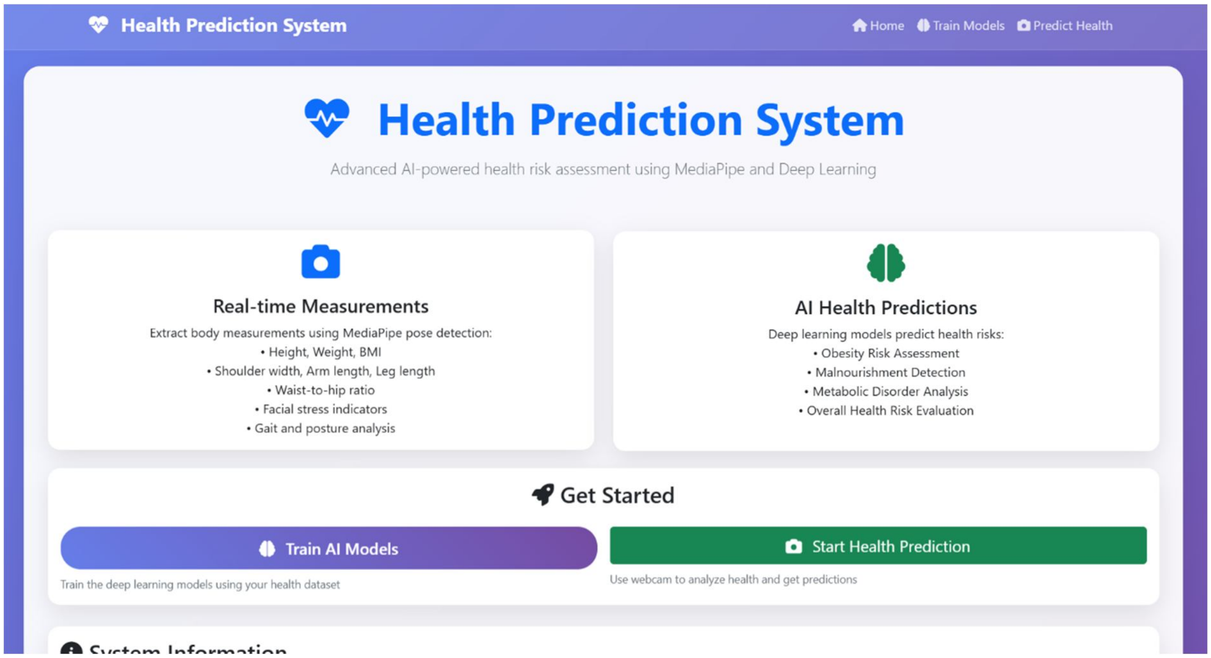
Next steps include
- Increasing the dataset size to cover more varied body types and demographics
- Adding explainable AI to develop user trust
- Integration with electronic health records and third-party fitness APIs

This system is a significant step towards affordable, scalable, and smart health screening. It presents possible use cases in telemedicine, wellness apps, and population-scale community health programs**.**

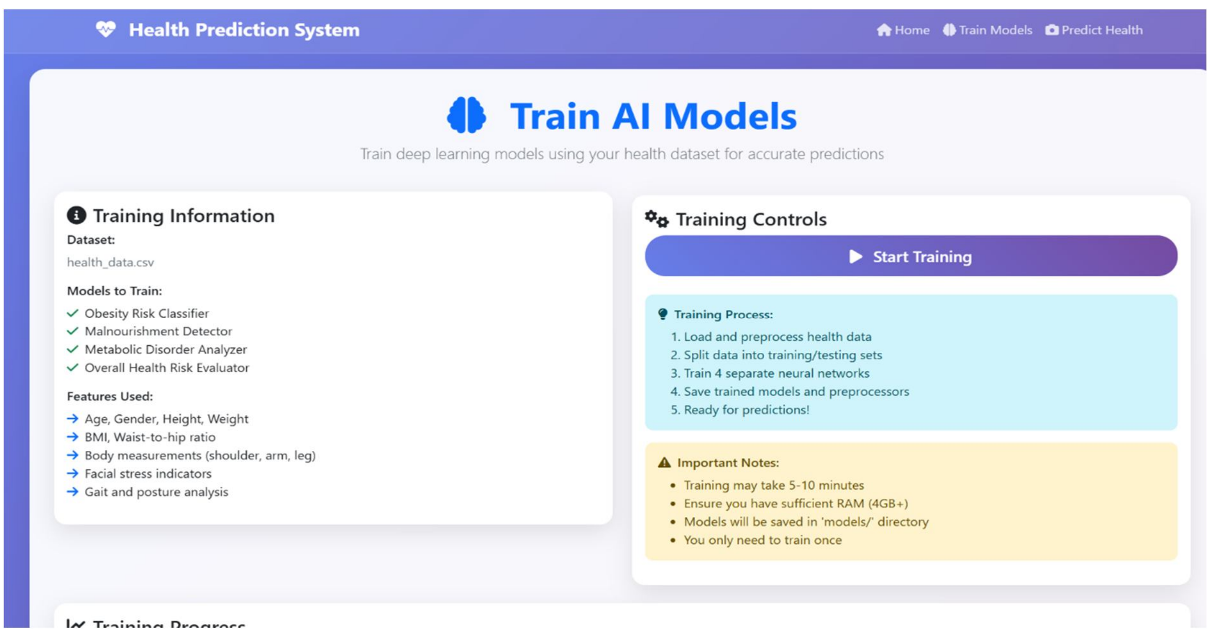
IX. WEBSITE SCREENSHOTS

Website Code :- https://github.com/suhasamane1101/Health_Analysis_Main.git

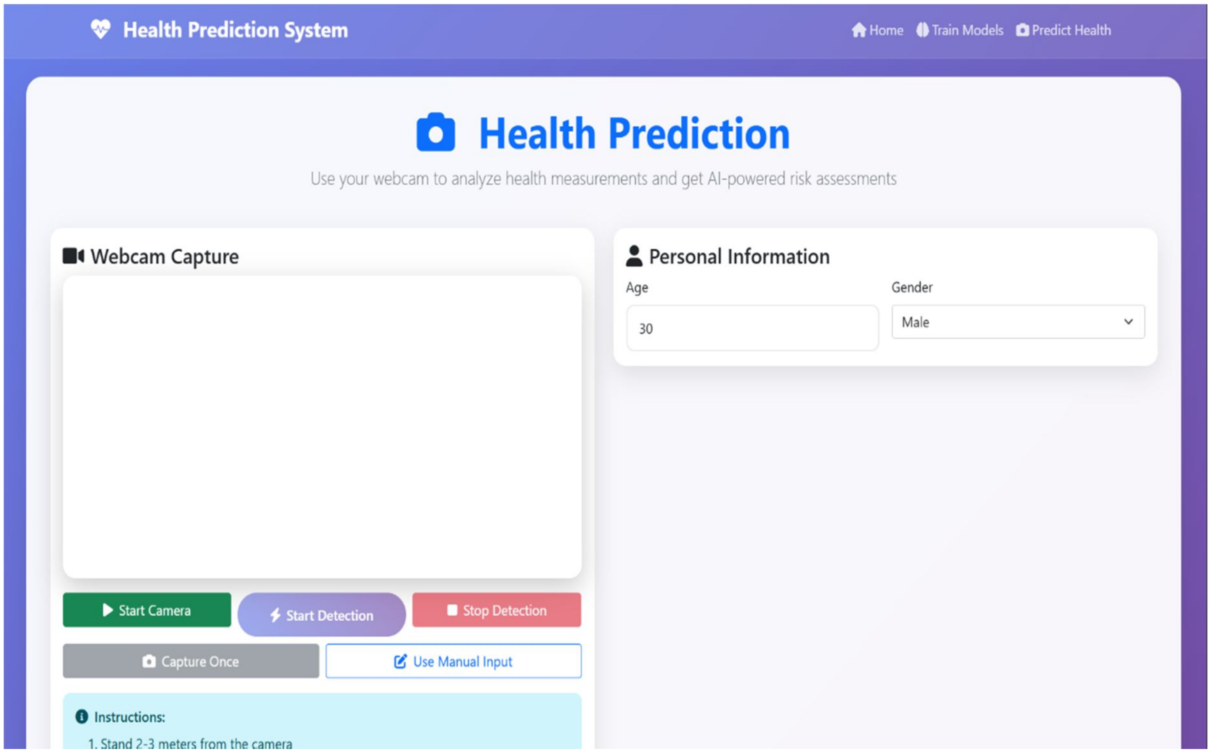
1) Home Screen



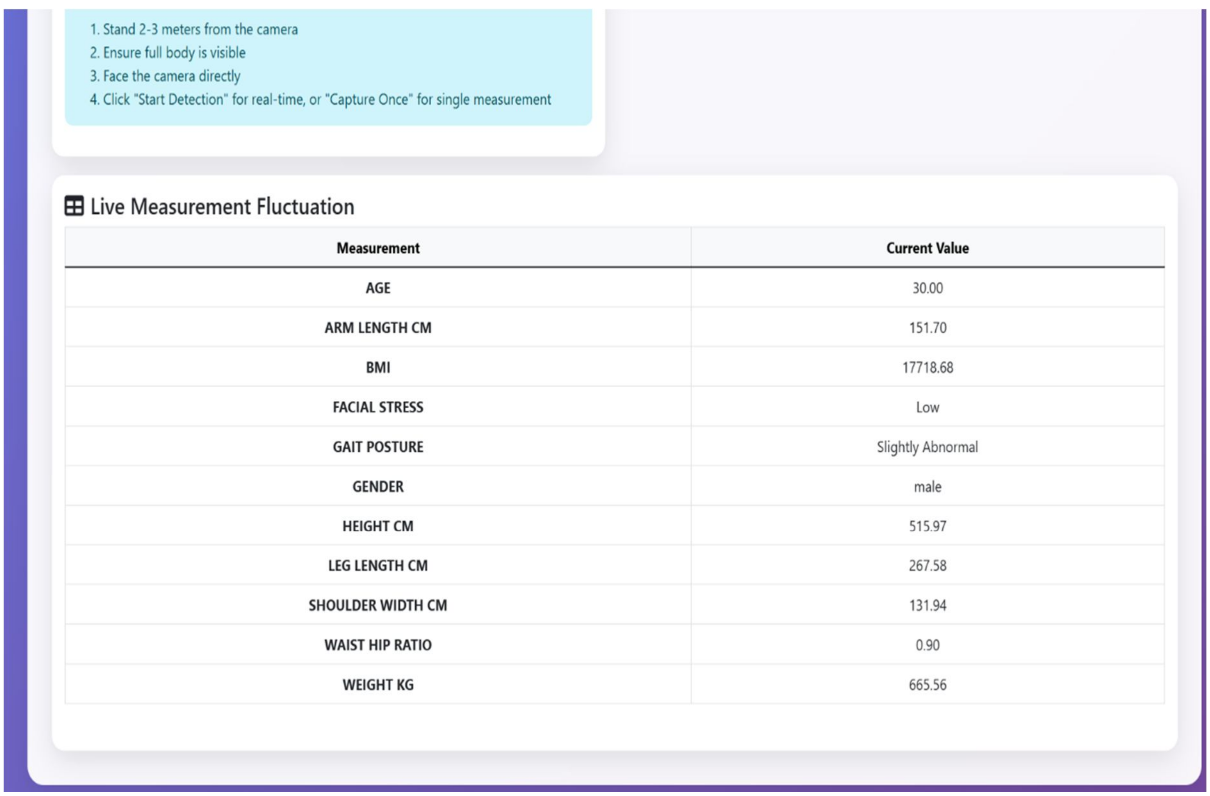
2) Training Model Using Deep Learning Algorithms



3) Prediction Page

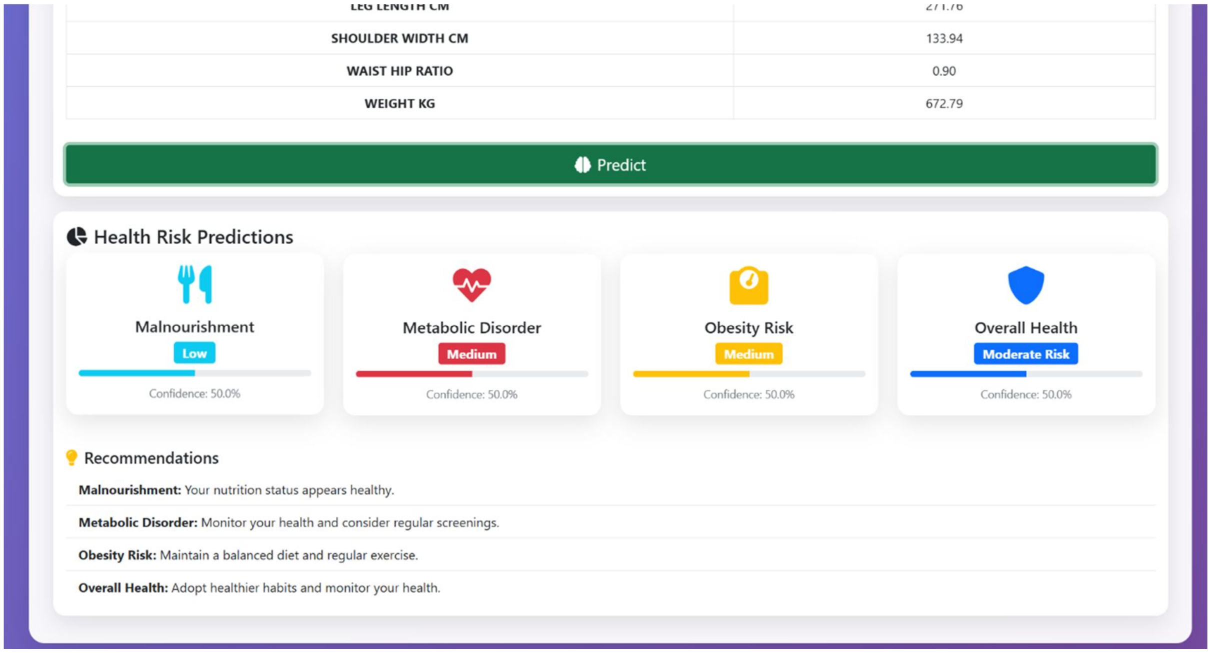


4) Actual Variables For Prediction



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

5) Prediction result



REFERENCES

[1] Škorvánková, D. et al.
Automatic Estimation of Anthropometric Human Body Measurements. arXiv, Dec 2021.
https://arxiv.org/abs/2112.11992

[2] Škorvánková, D. et al.
Automatic Anthropometric Estimation from 2D Images. VISAPP 2022.
https://www.scitepress.net/PublishedPapers/2022/108781/108781.pdf

[3] Dana Škorvánková et al.
Anthropometric Estimation from Structured Point Clouds.
https://www.researchgate.net/publication/363292750

[4] Supriya, A.
Human Body Measurement Using MediaPipe. JETIR, 2025.
https://www.jetir.org/papers/JETIR2503719.pdf

[5] Li, Z. et al.
Human Health Indicator Prediction from Gait Video. arXiv, Dec 2022.
https://arxiv.org/abs/2212.12948

[6] Li, Z. et al.
GLANCE: Gait-Linked Anthropometric and Clinical Estimator.
https://paperswithcode.com/paper/human-health-indicator-prediction-from-gait

[7] Pantanowitz, A. et al.
Estimation of Body Mass Index from Photographs using CNN. arXiv, 2019.
https://arxiv.org/abs/1908.10474

[8] Ribeiro, M. et al.
A Review of Deep Learning in Obesity: From Diagnosis to Treatment. Elsevier, 2024.
https://www.sciencedirect.com/science/article/pii/S1532046423000453

[9] Cell Press Study
Pose Estimation Feedback for Movement Skill Improvement. Heliyon, 2024.
https://www.cell.com/heliyon/fulltext/S2405-8440(23)05522-5

[10] Islam, M. et al.
Pose-Based Ergonomic Risk Assessment from RGB Images. ScienceDirect, 2024.
https://www.sciencedirect.com/science/article/pii/S0263224124004116

[11] Wahid, M. et al.
Vision-Based AI for Physical Health Monitoring. Heliyon, 2023.
https://www.cell.com/heliyon/fulltext/S2405-8440(23)05880-7

[12] DeepPose: Toshev & Szegedy
Human Pose Estimation via Deep Neural Networks. arXiv, 2013.
 https://arxiv.org/abs/1312.4659
[13] PLOS Digital Health
Clinical Gait Analysis Using Video-Based Pose Estimation. 2023.
 https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000082
[14] Nguyen, T. et al.
3D Human Pose Estimation from Monocular Video. Elsevier, 2022.
 https://www.sciencedirect.com/science/article/pii/S0262885622000999
[15] Goyal, P. et al.
Markerless Human Pose Estimation for Gait Monitoring. MDPI Sensors, 2021.
 https://www.mdpi.com/1424-8220/21/12/4166
[16] Zhang, L. et al.
Attention-Based CNN for BMI Prediction Using Face and Body Features. Elsevier, 2022.
 https://www.sciencedirect.com/science/article/pii/S0167865522000022
[17] Fadllullah, A. et al.
Automatic Height Measurement Using Camera and Deep Learning. IJEECS, 2024.
 https://www.researchgate.net/publication/381880387
[18] Nguyen, H. et al.
3D Anthropometric Measurement Using Compact Stereo Vision. ScienceDirect, 2025.
 https://www.sciencedirect.com/science/article/pii/S0263224125001368
[19] Kim, S. et al.
Sex-Based Body Fat Percentage Estimation from 2D Camera. ScienceDirect, 2023.
 https://www.sciencedirect.com/science/article/pii/S0263224123007777
[20] Ren, B. et al.
Mental Health Prediction Using Machine Learning. IEEE RBME, 2023.
 https://doi.org/10.1109/RBME.2023.3290497
[21] Ahmed, S. et al.
AI for Health Diagnosis Using Machine Learning. Frontiers Psychiatry, 2023.
 https://doi.org/10.3389/fpsyt.2023.1076124
[22] Islam, M. et al.
A Hybrid Deep Learning Approach for Health Prediction Using Pose. IEEE Access, 2022.
 https://doi.org/10.1109/ACCESS.2022.3222545
[23] Nguyen, D. et al.
Deep Learning for Digital Health Biometrics and Posture. NPJ Digital Medicine, 2022.
 https://doi.org/10.1038/s41746-022-00613-6
[24] Li, C. et al.
Speech-Based Health Detection with Multi-View Learning. Interspeech, 2022.
 https://www.isca-speech.org/archive/pdfs/interspeech_2022/li22q_interspeech.pdf
[25] "Human Body Measurement Estimation with Adversarial Augmentation." arXiv, Oct 2022.
 https://arxiv.org/pdf/2210.05667v1