**"AI Coach" Project Proposal: A Strategic and Technical Framework**

**Prepared by:** The Graduate Project Team:

* Abdulsalam ElSayed
* Amr ElDosoky
* Seif Hassan

**Date:** October 4, 2025

**Table of Contents**

**1.0 Executive Summary & Strategic Recommendation**

1.1 Project Vision

1.2 Core Problem

1.3 Strategic Recommendation

**2.0 Technical Deep Dive & Proposed Architecture**

2.1 Input Modality & Data Processing

2.2 Ground Truth & Data Acquisition Strategy

2.3 Proposed Multimodal Architecture

**3.0 Market & Competitive Landscape Analysis**

3.1 Competitor Analysis

3.2 Unique Selling Proposition (USP)

**4.0 Actionable Project Roadmap**

4.1 Phase 1: Foundation & Data Acquisition (Months 1-2)

4.2 Phase 2: Core Model Development (Months 3-5)

4.3 Phase 3: Prototyping & Validation (Months 6-7)

4.4 Phase 4: User Testing & Final Report (Month 8+)

**5.0 Ethical Considerations**

5.1 Data Privacy and Security

5.2 Algorithmic and Dataset Bias

5.3 Social Impact and Body Image

**1.0 Executive Summary & Strategic Recommendation**

**1.1 Project Vision**

The "AI Coach" project aims to develop a prototype mobile application that provides personalized, accurate, and interpretable body composition analysis. By leveraging multimodal AI, the application will move beyond the limitations of commercial Bioelectrical Impedance Analysis (BIA) devices, offering users and their trainers a more trustworthy and empowering tool for tracking fitness and wellness goals.

**1.2 Core Problem**

Initial market research and surveys with fitness professionals confirm two primary pain points with incumbent technologies like InBody: a fundamental distrust in the accuracy of raw outputs, which are easily skewed by transient factors (e.g., hydration), and a poor user experience, with reports that are complex and difficult for clients to interpret. The "AI Coach" directly addresses this market gap by aiming for superior accuracy and delivering results through a "coach-eye simulation" that explains the

*why* behind the numbers.

**1.3 Strategic Recommendation**

This document outlines a strategic and technical framework for the project, focusing on an **image-centric, hardware-agnostic approach**. This path is more innovative, aligns with state-of-the-art research, and positions the project at the forefront of computer vision in fitness. The core of this strategy involves pre-training models on a large-scale synthetic dataset and leveraging Explainable AI (XAI) to create the project's key market differentiator: trust through transparency.

**2.0 Technical Deep Dive & Proposed Architecture**

This section details the recommended technical pathway for developing the AI Coach prototype, focusing on an image-centric multimodal prediction model.

**2.1 Input Modality & Data Processing**

The most feasible and user-friendly approach is to build the model around **2D photos** captured by a standard smartphone.

* **User Input:** The application's front-end will capture front and lateral 2D photos. This method significantly lowers the barrier to entry for users, as it requires no specialized hardware.
* **Preprocessing Pipeline:** The captured photos will be processed into binary **silhouettes**. This critical step standardizes the input and makes the model robust to real-world variations in lighting, background, and clothing texture, a methodology validated by the success of the BodyM model. This ensures the model focuses purely on the geometric shape of the body.

**2.2 Ground Truth & Data Acquisition Strategy**

The primary technical challenge is the scarcity of large-scale datasets pairing body images with "gold standard" body composition labels (e.g., from DEXA scans). To overcome this, a **synthetic-first** training strategy is essential.

1. **Pre-training with Synthetic Data:**
   * **Generation:** A parametric 3D body model, such as **SMPL** (Skinned Multi-Person Linear model), will be used to programmatically generate hundreds of thousands of unique 3D human bodies by randomizing shape and pose parameters. This approach is validated by the "Adversarial Body Sim" paper, which used it to create challenging and diverse training examples.
   * **Labeling:** For each synthetic body, all ground-truth data—including the 3D mesh, 14+ anthropometric measurements, and 2D silhouettes (via a **differentiable renderer**)—are automatically and perfectly known. This allows for the creation of a massive, perfectly-labeled dataset for pre-training the core computer vision components of the model.
2. **Sourcing Real-World Data (for Fine-Tuning & Validation):**
   * **Public Datasets:** The project should pursue academic access to public anthropometric datasets to bridge the gap between synthetic and real-world data. The **CAESAR** dataset, which contains 3D scans and measurements for over 4,000 individuals, is the most promising option and is available for academic use. Other potential datasets include

**DFAUST** (4D scans) and **OSSO** (3D models paired with DXA scans), though access may be more restrictive.

* + **Research Collaboration:** A targeted outreach plan should be initiated to collaborate with university kinesiology departments, sports medicine clinics, or high-end training facilities that operate a DEXA scanner. By offering a free research tool in exchange for anonymized, consent-based data (images + DEXA results), the project can acquire a small but invaluable high-quality dataset for fine-tuning the model and, most importantly, for **rigorously validating its real-world accuracy**.

**2.3 Proposed Multimodal Architecture**

A modular, multi-step architecture is proposed to systematically address the problem from input to explainable output.

| Step | Component | Description | Training Data |
| --- | --- | --- | --- |
| **1. Input** | Mobile App UI | Capture front/side photos, weight, height, age, gender, and lifestyle questionnaire data (e.g., sleep, activity level). | N/A |
| **2. Image Processing** | Silhouette Extraction | A segmentation network converts user photos into clean, binary silhouette images for standardized input. | Public segmentation datasets. |
| **3. Measurement Extraction** | CNN Model (BMnet-inspired) | Takes the pair of silhouettes as input and regresses the 14 key anthropometric measurements. | **Pre-trained** on the large synthetic SMPL dataset. **Fine-tuned** on real scan data (e.g., CAESAR). |
| **4. Multimodal Fusion** | Fusion Model (MLP or Gradient Boosting) | Fuses all tabular data: the 14 predicted measurements from Step 3, plus user-provided height, weight, age, gender, and lifestyle data. | Trained on the synthetic dataset, using anthropometric formulas (e.g., RFM) to generate "pseudo ground-truth" body fat % labels. |
| **5. Output Prediction** | Regression Head | The final layer of the Fusion Model, which outputs the predicted body composition values (e.g., Body Fat %, Lean Mass %). | (Same as Step 4) |
| **6. Explanation Generation** | Explainable AI (XAI) Module | Applies a method like **SHAP** or **LIME** to the trained Fusion Model to determine the contribution of each input feature to the final prediction. | Uses the trained model from Step 4. |
| **7. User-Facing Output** | Mobile App UI | Displays the predicted numbers and a coach-style textual explanation generated from the XAI output (e.g., "This estimate was primarily influenced by your waist-to-hip ratio."). | N/A |

**3.0 Market & Competitive Landscape Analysis**

**3.1 Competitor Analysis**

The market consists of three primary segments:

1. **B2B Hardware Scanners:** Incumbent technology targeting gyms and clinics.
   * **Companies:** **Fit3D**, **Styku**, **TC2**.
   * **Business Model:** Expensive hardware (~$9,000-$13,000+) with a monthly software subscription.
   * **Key Threat:** **Fit3D's SNAP** product is a hardware-free, app-based solution, making them a direct and formidable competitor leveraging a massive dataset of over 6 million scans.
2. **B2C App-Based Solutions:** Direct competitors for individual users.
   * **Companies:** **Spren**, **Zing Coach**, and **Fytted** offer body composition or measurement tracking from photos.

**NuraLogix (Anura)** uses video selfies for vitals but is less focused on detailed body composition.

* + **Business Model:** Freemium or subscription-based mobile apps.

1. **Related Technologies:** Market indicators validating the underlying technology.
   * **Virtual Try-On (VTO):** Companies like **3DLOOK** and platforms like Snapchat are maturing the technology for 3D human reconstruction from photos for fashion.

**3.2 Unique Selling Proposition (USP)**

The "AI Coach" project has three powerful differentiators:

1. **Accessibility & Low Cost:** As a smartphone-only solution, the app eliminates the need for specialized hardware, making it accessible to freelance trainers, small studios, and individuals who cannot afford a $10,000 scanner.
2. **Holistic, Multimodal Analysis:** The model's ability to fuse visual data with lifestyle inputs (sleep, activity) and historical trends allows for a more personalized and context-aware assessment that improves over time for each user.
3. **Trust Through Explainability:** This is the "killer feature." The core user problem is not just inaccuracy but a lack of understanding. By using XAI to provide "coach-eye" explanations, the app builds trust and provides actionable feedback, moving beyond being a black-box calculator to become an analysis tool that explains its reasoning.

**4.0 Actionable Project Roadmap**

**4.1 Phase 1: Foundation & Data Acquisition (Months)**

* **Data Scoping:** Begin the application process for academic access to the **CAESAR** dataset.
* **Outreach:** Draft a proposal for collaboration and contact local university labs and clinics with DEXA scanners.
* **Technical Setup:** Implement the SMPL model to begin building the synthetic data generation pipeline.

**4.2 Phase 2: Core Model Development (Months)**

* **Synthetic Data Generation:** Generate a large-scale (100k+ samples) synthetic dataset of SMPL bodies with corresponding silhouettes and measurements.
* **Model Training:**
  1. Train the image-to-measurement CNN (Step 3) on the synthetic data.
  2. Train the multimodal fusion model (Step 4) on the synthetic data using formula-derived BFP labels.

**4.3 Phase 3: Prototyping & Validation (Months)**

* **App Prototype:** Develop a minimal viable mobile app to capture inputs and display predictions.
* **Fine-Tuning:** If a real-world dataset is acquired, fine-tune the end-to-end model.
* **XAI Implementation:** Integrate a library like SHAP to generate feature importance scores and convert them into human-readable text.

**4.4 Phase 4: User Testing & Final Report (Month)**

* **Pilot Test:** Deploy the prototype to a small group of target users (e.g., coaches from the initial survey).
* **Gather Feedback:** Collect qualitative and quantitative feedback on accuracy, usability, and the value of the XAI-driven explanations.
* **Finalize:** Document results and write the final graduation project report.

**5.0 Ethical Considerations**

**5.1 Data Privacy and Security**

3D body scan data and biometric measurements are sensitive personal data under regulations like GDPR. The project must ensure all data handling practices are secure, user data is anonymized, and explicit user consent is obtained via a clear privacy policy.

**5.2 Algorithmic and Dataset Bias**

The model's accuracy will reflect its training data. If the datasets lack diversity in body types, ethnicities, and ages, the model may perform poorly for under-represented groups, perpetuating harmful biases. A conscious effort must be made to create a balanced and diverse dataset.

**5.3 Social Impact and Body Image**

An application that quantifies the human body can significantly impact users' mental health and body image. The user experience must be designed responsibly:

* Frame results as **estimations** with a stated margin of error.
* Prioritize displaying **healthy ranges** over singular "ideal" numbers.
* Accompany results with positive, health-focused educational content.