Article

AI-Assisted Nutrition Estimation and Exercise Support for Integrated Lifestyle Management in Patients with Diabetes

Muhammad Jamil [1,2,*], Adnan Kavak [1,2,*], Hossein Fotouhi [3,*], Md. Rashed [4], Emre Gezer [5], Alpaslan Burak İnner [1,2] and Gautam Srivastava [6,7,8,]

[1] Department of Computer Engineering, Kocaeli University, 41001 İzmit, Türkiye; jamil138.amin@gmail.com (M.J.); akavak@kocaeli.edu.tr (A.K.); binner@kocaeli.edu.tr (A.B.İ.)
[2] Wireless Information and Intelligent Systems (WINS) Research Center, Kocaeli University, 41001 Kocaeli, Türkiye
[3] Department of Computer Science and Engineering, Mälardalen University (MDU), 72123 Västerås, Sweden; hossein.fotouhi@mdu.se (H.F.)
[4] Department of Information and Communication Engineering, Pabna University of Science and Technology, Pabna 6600, Bangladesh; rashedulislam.ice.pust@gmail.com (M.R.)
[5] Department of Endocrinology and Metabolism, Faculty of Medicine, Kocaeli University, Kocaeli, Türkiye; gezer$_e$$mre@hotmail.com(E.G.)$
[6] Sogang University, Seoul 04107, Republic of Korea; gsrivast@ieee.org (G.S.)
[7] Research Center for Interneural Computing, China Medical University, Taichung 40402, Taiwan
* Correspondence: jamil138.amin@gmail.com (M.J.); akavak@kocaeli.edu.tr (A.K.); hossein.fotouhi@mdu.se (H.F.)

Abstract

Background/Objectives: Multidimensional lifestyle interventions that combine healthy diet, physical activity, and psychosocial support are central to effective self-management of type 1 and type 2 diabetes. Although digital health tools can improve lifestyle monitoring, existing AI-based dietary assessment methods often struggle to estimate food quantity accurately because food size and scale are difficult to determine from images. This can lead to inconsistent estimates of food mass and energy content. **Methods:** In this study we introduce nutrition and exercise modules of our AI-assisted Diabetes Care (AIDCare) mHealth digital platform developed for diabetes patients. For effective diet management, we present the PC^{2}FoodNet model, which uses an EfficientNetV2-S backbone with multi-task regression and food-density priors to jointly estimate food volume, mass, and nutrients. For patient safety, we put nutrition-in-the-loop for making diet plans and rectifications of patients' daily key nutrient intake according to the requirements of the patients. The exercise module includes professionally designed exercises using Unreal Engine's MetaHuman plugin. The model was evaluated using 4-fold stratified cross-validation on a dataset of 22,070 images covering 40 Turkish food categories. **Results:** PC^{2}FoodNet achieved a mean Top-1 classification accuracy of 94.39% $\pm$ 0.47% and Top-5 accuracy of 98.90% $\pm$ 0.18% across the four folds. The corresponding Macro-F_1 and Weighted-F_1 scores were 93.14% $\pm$ 0.59% and 94.16% $\pm$ 0.36%, respectively. The physics-constrained regression framework achieved mean RMSE values of 3.01 mL for food volume, 30.17 g for mass, and 57.55 kcal for energy. The confidence-based referral mechanism identified uncertain cases for clinical review, resulting in a clinical referral rate of 23.81%. **Conclusions:** AIDCare combines physically consistent AI-based dietary assessment with personalized physical activity and continuous psychosocial support. By keeping nutrition involved when AI predictions are uncertain, the proposed framework reduces the risks associated with fully automated lifestyle recommendations. This approach provides a practical foundation for safer, more personalized, and evidence-based diabetes lifestyle management.

Received:
Accepted:
Published:

https://doi.org/10.3390/nu1010000

Keywords: AIDCare; diabetes self-management; dietary assessment; food image recognition; mHealth; nutrition estimation; physical activity

1. Introduction

Nutrition and physical activity are central pillars in the self-management of type 1 diabetes (T1D) and type 2 diabetes (T2D). Clinical guidelines strongly advocate individualized medical nutrition therapy, appropriate carbohydrate and energy monitoring, and regular aerobic and resistance exercise [1]. However, maintaining these behaviors in daily life remains challenging because dietary intake, physical activity, and psychosocial factors are continuously changing. Digital health technologies can reduce the burden of manual monitoring, but their effectiveness depends on sustained patient engagement, reliable measurements, and appropriate clinical support. Evidence from mobile health and diabetes technology studies indicates that digital interventions are more useful when automated monitoring is combined with individualized guidance and professional follow-up rather than functioning as isolated autonomous systems [2–6].

Dietary assessment is particularly challenging because a food image provides incomplete information about the physical quantity and nutritional composition of a meal. Although image-based food recognition has achieved substantial progress, identifying a food category does not directly determine its portion size, mass, or energy content. Monocular images contain no direct measurement of scale and can be affected by perspective, occlusion, food preparation, and visually similar dishes with different densities and nutritional properties [7–9]. Conventional approaches that first select a single food class and subsequently apply fixed nutritional values may therefore discard classification uncertainty before estimating physical quantities. This limitation is particularly important in diabetes management, where inaccurate dietary estimates can affect the interpretation of daily nutritional intake and potentially influence subsequent lifestyle decisions.

A further challenge is that reliable dietary assessment requires more than accurate point predictions. An AI system should also indicate when its predictions are uncertain so that potentially unreliable estimates are not presented to patients without appropriate review. Existing food-analysis approaches have explored visual recognition, portion estimation, geometric reconstruction, and multimodal reasoning, but these components are often treated as separate tasks and generally provide limited mechanisms for combining physical constraints with calibrated uncertainty and professional oversight [8,10–12]. This creates a need for an integrated approach in which uncertainty in food recognition can be propagated to downstream physical estimates and uncertain cases can be selectively escalated rather than being treated as equally reliable predictions.

To address these challenges, we present the Physics-Constrained, Confidence-Calibrated Food Analysis Network (PC^{2}FoodNet), a multi-task deep learning model developed as the dietary assessment engine of the AIDCare mHealth platform. PC^{2}FoodNet jointly performs food classification and physical quantity estimation while incorporating category-level density and energy-density priors into a differentiable inference pathway. Rather than relying exclusively on the most probable food class, the model uses the complete class-probability distribution to derive expected physical priors, allowing classification ambiguity to propagate into subsequent volume, mass, and energy estimates. The model further combines direct predictions with physics-based estimates and learns sample-specific uncertainty for continuous outputs.

PC^{2}FoodNet is integrated within AIDCare as one component of a broader lifestyle-management framework that combines nutrition and physical activity support. The Nutrition and Diet Module uses AI-assisted meal analysis to support dietary logging and individualized nutrition planning, while uncertain dietary predictions are referred to a Nutrition Professional (Dietitian) for verification. The Exercise Module provides structured physical activity planning and delivery under expert physiotherapist guidance, with movement-specific demonstrations developed using Unreal Engine's MetaHuman plugin. Psychotherapeutic support is incorporated to address behavioral adherence and motivation. Importantly, AIDCare does not automatically convert uncertain dietary estimates into compensatory exercise prescriptions; instead, nutrition and exercise remain coordinated but professionally governed domains.

The main contributions of this study are as follows:

1. We develop PC^{2}FoodNet, a multi-task food-analysis model that propagates the complete food-class probability distribution into expected density and energy-density priors for downstream physical estimation.
2. We introduce a physics-constrained inference pathway that links volume, mass, and energy estimation through bounded residual corrections and adaptive soft fusion.
3. We incorporate temperature scaling, conformal prediction sets, and heteroscedastic uncertainty estimation to identify ambiguous or unreliable predictions for selective professional review.
4. We evaluate the proposed approach on a 22,070-image Turkish food dataset covering 40 food categories using four-fold stratified cross-validation with explicit leakage-prevention procedures.
5. We demonstrate the integration of PC^{2}FoodNet within the AIDCare Nutrition and Diet Module and its coordination with an expert-guided Exercise Module under multidisciplinary professional oversight.

The remainder of this paper is organized as follows. Section 2 reviews related research in digital diabetes lifestyle support, exercise delivery, and image-based dietary assessment, and identifies the research gaps addressed by this study. Section 3 describes the study design, dataset, preprocessing and cross-validation protocol, PC^{2}FoodNet architecture, training procedure, physical reference targets, and evaluation metrics. Section 4 presents the integration of PC^{2}FoodNet within the AIDCare mHealth platform, including the Nutrition and Diet Module, Exercise Module, and multidisciplinary professional oversight. Section 5 reports the experimental results, including cross-validation performance, class-level evaluation, physical quantity estimation, and confidence-aware referral. Section 6 discusses the findings, practical implications, and limitations, while Section 7 concludes the study and outlines its overall contribution to AI-assisted diabetes lifestyle management.

2. Related Work

2.1. Digital Lifestyle Support in Diabetes

Diabetes mHealth systems increasingly combine physiological monitoring with food logs, activity records, coaching, and decision support. The American Diabetes Association (ADA) frames diabetes technology broadly, including software that supports lifestyle modification and data sharing, while stressing personalization [13]. The synthesis of the authors of [14] found that digital interventions for adults with T1D and T2D varied substantially in reach, uptake, delivery mode, and effectiveness. Reviews of mobile behavior-change components and prevention-oriented technologies likewise report modest average benefits accompanied by heterogeneity in engagement and outcomes [6,15]. Recent studies also indicate that digital health solutions support patient adherence and therapeutic outcomes when they incorporate human-in-the-loop oversight. Placing Nutrition Professionals (dietitians), directly in the clinical review loop provides psychological reinforcement prevents patient attrition, and ensures patient safety across dietary domains [16,17]. AIDCare therefore treats diet and exercise as measured, separable modules within a shared patient context, feeding structured AI outputs to domain specialist for review before patient delivery.

Wearable integration can enrich this shared context, but the evidence base remains small. A review of smartwatch-supported diabetes care identified only a few heterogeneous studies and called for stronger evaluation of glycemic control, exercise participation, medication adherence, and diet monitoring [18]. This supports using wearable and vital data as contextual inputs while avoiding assumptions that their availability alone improves clinical outcomes.

2.2. Digital Exercise, Immersive Delivery, and Professional Oversight

Remote exercise delivery is relevant when patients cannot regularly attend supervised sessions. The scoping review by the authors of [19] found that tele-exercise studies in T1D and T2D commonly combined aerobic and resistance exercise, but the evidence did not establish one universal delivery model. Meta-analytic evidence also indicates that exercise modality, total dose, behavioral techniques, and facilitator type should be considered when plans are individualized [2,3]. These findings support AIDCare's the use of an expert physiotherapist-curated exercise library rather than unrestricted automated exercise generation.

Immersive systems can provide repeatable demonstrations and remote engagement, yet evidence should be interpreted cautiously. The research work of [20,21] found generally favorable acceptability for immersive exercise and promising rehabilitation outcomes, alongside heterogeneous protocols and limited evidence of superiority over conventional care. A review of [22] reached a similar conclusion for Metaverse-aided rehabilitation, noting few eligible studies and unresolved questions about long-term effects, accessibility, safety, and integration with clinical workflows. The human-in-the-loop oversight therefore remains a foundational component of the proposed workflow.

2.3. Food Recognition and Nutrient Estimation

Large-scale food-recognition work has improved fine-grained category learning and transfer across food datasets [23]. Nutrient estimation is more difficult because visually similar foods can differ in preparation, density, and composition, while camera perspective obscures scale. Recent reviews found that reported error varies widely with the dataset, ground-truth definition, imaging protocol, and degree of nutrition-professional involvement [8,10–12,24]. Multimodal foundation models can recognize foods and estimate dietary variables without task-specific training, but controlled studies still report sensitivity to image conditions and prompting [25].

Recent portion-estimation systems illustrate several ways to recover scale or constrain quantity. The authors of [26] evaluated a population-specific food-photograph series for portion selection. While the authors of [27] estimated intake from differences between pre- and post-meal images. The authors of [28] modeled bowl geometry and visible fullness while Shao et al. [29] reconstructed three-dimensional food shape from a monocular image. CalorieMe used detection, segmentation, and a reference object [30], whereas newer attention-based pipelines infer ingredients or recipes before calculating energy [31]. These approaches improve particular parts of the estimation chain but do not remove the uncertainty created by occlusion, unknown density, hidden ingredients, and imperfect scale.

Recent application-oriented systems cover complementary parts of the problem. The authors of [32] classified food groups and discrete portion categories while Diet Engine proposed by [33] combined image recognition with personalized dietary guidance. Within AIDCare, in one of our previous studies we connected CNN classification and MLLM portion estimation to advisory insulin decision support for T1D [34] while in another study we used confidence-aware vision–language routing for culturally specific food recognition and nutritional monitoring [35]. Table 1 provides a comprehensive function-by-function comparison of these state-of-the-art approaches with PC^{2}FoodNet.

Table 1. Comparison of recent food recognition, portion estimation, and dietary assessment approaches with PC^{2}FoodNet.

Study / System	Year	Main Task & Input	Continuous Quantity	Physical Prior	Uncertainty / Referral	Multidisciplinary Oversight & Distinction
[23]	2023	Large-scale RGB food recognition	No	No	No	Category benchmark; lacks physical quantity estimation or clinical review.
[30]	2023	Detection, segmentation, reference card	Yes	No	User confirmation	Scale recovered via physical reference card; no joint density priors.
[29]	2023	3D food shape reconstruction	Yes	Geometry	No	3D mesh volume estimation; computationally heavy, no uncertainty routing.
[26]	2023	Photo-series portion selection	Discrete	No	No	Manual visual portion selection; relies heavily on user effort.
[28]	2024	Bowl geometry & fullness modeling	Yes	Geometry	No	Volume estimated for container-bound foods; uncalibrated for general dishes.
[25]	2024	Multimodal foundation model (VLM)	Yes	No	Prompt confidence	Zero-shot dietary assessment; output sensitive to prompt engineering.
[32]	2025	CNN food group & portion classification	Discrete	No	No	Portion modeled as discrete class rather than continuous physical variables.
[33]	2025	Detection, classification, advice	Limited	No	No	Mobile advisory workflow; lacks physics-constrained multi-task pathway.
[36]	2025	Multimodal food analysis	Yes	No	No	Joint nutritional advisory; lacks confidence calibration or gating.
[31]	2025	Attention-based ingredient energy	Yes	Recipe	No	Ingredient-level energy inference; no physical volume-weight fusion.
[34]	2026	CNN classification + MLLM portion	Yes	No	User check	Connects food logging to bolus advice; classification and portion remain decoupled.
[35]	2026	CNN + selective MLLM routing	Yes	No	Confidence routing	Selective Vision-Language Model (VLM) escalation; physical consistency is not trained end-to-end.
PC^{2}FoodNet (This study)	**2026**	**RGB recognition + linked volume, mass, energy regression**	**Yes**	**Yes**	**Conformal + Heteroscedastic**	**Evaluated on 22,070 images (40 classes); 100-g portion targets; integrated with Exercise Module under Nutritionist, Physiotherapist, and Psychotherapist oversight.**

2.4. Research Gaps

The literature supports the feasibility of automated dietary assessment but leaves several key technical and clinical gaps. Table 2 maps each gap to the corresponding design response in PC^{2}FoodNet and the AIDCare framework, alongside the empirical evidence established in this study.

3. Materials and Methods

3.1. Study Design and Role within AIDCare mHealth Digital Platform

We conducted the technical development and validation of PC^{2}FoodNet and the Exercise Module within the AIDCare mHealth digital platform. The proposed PC^{2}FoodNet processes monocular RGB food images to generate food classification probabilities, volume,

Table 2. Research gaps in automated dietary and lifestyle monitoring addressed by PC^{2}FoodNet and the AIDCare framework.

Research Gap	Clinical / Technical Impact	PC^{2}FoodNet / AIDCare Solution	Empirical Evidence & Boundary
Hard class-to-nutrient lookup	Single argmax class discards prediction ambiguity for visually similar dishes.	Expected density ($\bar{\rho}$) and energy density ($\bar{\kappa}$) probability-weighted across 40 classes.	End-to-end differentiable physical priors propagate classification uncertainty to quantity heads.
Independent quantity regressors	Unconstrained volume, mass, and energy predictions produce physical contradictions.	Hierarchical volume-weight-energy pathway, bounded residual corrections ($\delta = 0.20$), and soft fusion gates.	Achieves RMSEs of 2.72 mL for volume, 29.11 g for mass, and 55.84 kcal for energy on the Fold 4 validation set; four-fold mean RMSEs are 3.01 mL, 30.17 g, and 57.55 kcal, respectively.
Uncalibrated prediction confidence	High softmax probability can accompany severe portion or mass errors.	Temperature-scaled probabilities, conformal prediction sets, and heteroscedastic log-variance ($\hat{\sigma}_y^2$).	Identifies high-uncertainty predictions; selectively routes ambiguous cases to specialists.
Data leakage in cross-validation	Random image splits leak repeated dish specimens across training and validation sets.	4-fold stratified cross-validation ($K = 4$) with index-level splitting, fresh initializations, and fixed normalisation.	Strict data leakage prevention across 22,070 images (mean Val Acc@1 = 94.39% $\pm$ 0.47%).
Unsupervised decision escalation	Ambiguous AI outputs presented directly to patients risk inappropriate treatment.	Selective referral flags (23.81% referral rate) route high-uncertainty logs to clinical review queues.	Prevents autonomous AI error propagation by engaging clinical oversight.
Isolated mHealth lifestyle applications	Standalone food classifiers lack physical activity integration and clinical oversight.	Unified AIDCare ecosystem pairing Diet Module with Exercise Module (expert guides) under multidisciplinary oversight.	Connects Nutrition Professionals (dietitians), Physiotherapists, and Psychotherapists to a shared patient record.

mass, energy estimates, heteroscedastic uncertainty metrics, and a selective referral flag. Within AIDCare, these outputs form structured meal records feeding the Nutrition and Diet Module. Concurrently, the Exercise Module captures physical activity parameters, heart rate, active calories, and sleep metrics to support physiotherapist- and psychotherapist-curated exercise plans designed using the MetaHuman plugin.

The overarching design principle of AIDCare is to augment—rather than replace—clinical expertise. The system provides decision support for dietary logging and exercise planning, while predictions requiring further assessment are subject to expert review through a human-in-the-loop workflow.

3.2. Turkish Food Dataset and Domain Nutrition Priors

The experiments evaluate PC^{2}FoodNet on a Turkish food image dataset comprising $N = 22{,}070$ RGB images across $C = 40$ Turkish dish categories. The dataset was obtained from two of our previous studies [34,35], and the proposed model was applied to this dataset. We used per-class nutritional reference values per 100 g as a standard portion by specifying energy, calories$[c]$ (kcal), protein, protein$[c]$ (g), carbohydrate, carb$[c]$ (g), total fat, fat$[c]$ (g), and dietary fibre, fibre$[c]$ (g) for each dish class $c \in \{1, \dots, C\}$.

Prior to network training, two scalar physical domain priors are derived for each food category $c \in \{1, \dots, C\}$ directly from the nutritional database:

1. **Energy Density Prior** $\kappa_c \in \mathbb{R}_+$ (in kcal g^{-1}):

$$\kappa_c = \frac{\text{calories}[c]}{100} \tag{1}$$

2. **Macroscopic Density Prior** $\rho_c \in \mathbb{R}_+$ (in g mL^{-1}):

$$\rho_c = \frac{0.90 \cdot \text{fat}[c] + 1.30 \cdot (\text{protein}[c] + \text{carb}[c] + \text{fibre}[c])}{\text{protein}[c] + \text{carb}[c] + \text{fat}[c] + \text{fibre}[c] + \epsilon} \tag{2}$$

where constant coefficients $0.90\,\mathrm{g\,mL^{-1}}$ and $1.30\,\mathrm{g\,mL^{-1}}$ represent canonical component mass densities for lipids and non-lipid organic macronutrients, respectively, and $\epsilon = 10^{-8}$ prevents division by zero.

Because these priors are domain knowledge derived from nutritional databases rather than training images, they remain strictly fold-independent and introduce zero cross-validation data leakage.

3.3. Data Preprocessing and Leakage-Aware Stratified Cross-Validation

All images were resized to 224×224 pixels and normalized using fixed ImageNet statistics, with mean $\boldsymbol{\mu} = (0.485, 0.456, 0.406)$ and standard deviation $\boldsymbol{\sigma} = (0.229, 0.224, 0.225)$. During training, images underwent random cropping with a scale range of 0.6–1.0, horizontal flipping with probability 0.5, brightness, contrast, saturation, and hue variation of 0.3, 0.3, 0.3, and 0.05, respectively, and random rotation within $\pm 15^\circ$. Validation images were resized to 258 pixels and centrally cropped to 224×224 pixels without stochastic augmentation.

Four-fold stratified cross-validation ($K = 4$) was used, with an 80:20 training–validation split in each fold. For $N = 22{,}070$ images, each fold contained approximately 17,656 training and 4,414 validation images, corresponding to approximately 441 and 110 images per class, respectively. Data leakage was prevented by performing the partition before preprocessing, applying training and validation transformations independently, initializing the model independently for each fold, and preventing parameter or checkpoint sharing across folds. ImageNet normalization statistics and category-level nutritional and physical priors were fixed before cross-validation and remained unchanged across all folds. Validation images were therefore never used for model training or parameter estimation.

3.4. Network Architecture and Differentiable Inference Chain

PC^{2}FoodNet follows a sequential end-to-end workflow from food-image input to final nutritional outputs. The overall PC^{2}FoodNet architecture and inference workflow are illustrated in Figure 1. The process begins with image preprocessing, where the input image is resized, normalized, and prepared for model processing. The model then extracts visual features, identifies the food category, and estimates volume, mass, energy, and prediction uncertainty. Class probabilities are further used to incorporate physical food priors into the quantity estimates. The resulting predictions are then assessed for confidence; high-confidence predictions are accepted, while uncertain cases are flagged for professional review. The final outputs include food identity, estimated volume, mass, energy, uncertainty, and referral status.

3.5. Parameters of the Proposed Model

The model parameters θ are trained end-to-end using AdamW with gradient clipping (maximum norm $= 1.0$) and automatic mixed precision (AMP, FP16). The total multi-task loss objective $\mathcal{L}_{\text{total}}$ combines classification cross-entropy, Gaussian negative log-likelihood (NLL) regression, and a ramped physics consistency penalty:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{cls}}\mathcal{L}_{\text{CE}} + \lambda_{\text{phys}}\alpha(t)\mathcal{L}_{\text{phys}} + \sum_{y\in\{v,w,e\}} \lambda_y \mathcal{L}_{\text{NLL}}(\hat{y}, \hat{s}_y, y^*) \tag{3}$$

where components are defined as follows:

1. **Classification Loss** $\mathcal{L}_{\text{CE}}$ with label smoothing $\varepsilon = 0.1$:

$$\mathcal{L}_{\text{CE}} = -\sum_{c=1}^{C} \tilde{p}_c \log \hat{p}_c, \quad \text{where } \tilde{p}_c = (1-\varepsilon)\,\mathbf{1}[c = y^*] + \frac{\varepsilon}{C} \tag{4}$$

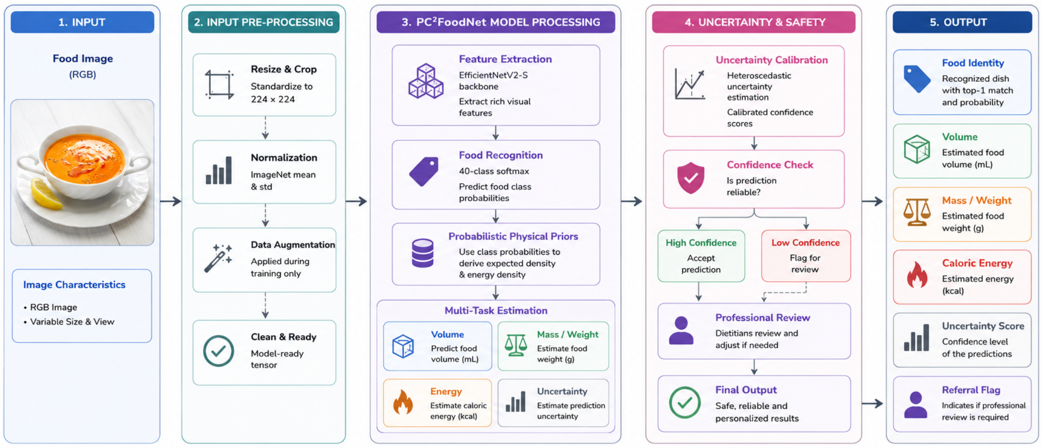

Figure 1. PC²FoodNet workflow from food-image input and preprocessing through model-based food recognition, physical quantity estimation, uncertainty assessment, professional review, and final nutritional outputs.

where $y^* \in \{1, \dots, C\}$ is the true target class index and $\mathbf{1}[\cdot]$ is the indicator function.

2. **Gaussian Negative Log-Likelihood Loss** $\mathcal{L}_{\text{NLL}}$ for continuous regression target y^*:

$$\mathcal{L}_{\text{NLL}}(\hat{y}, \hat{s}_y, y^*) = \frac{1}{2}\left(\hat{s}_y + \frac{(\hat{y} - y^*)^2}{\exp(\hat{s}_y) + \epsilon}\right) \tag{5}$$

where $\hat{y}$ is predicted mean, $\hat{s}_y$ is predicted log-variance, and $\epsilon = 10^{-8}$.

3. **Physics Consistency Loss** $\mathcal{L}_{\text{phys}}$ (Smooth-L1 penalty):

$$\mathcal{L}_{\text{phys}} = \text{SmoothL1}\Big(\text{clip}(\hat{w}, 0, 5000), \text{clip}(\hat{w}_{\text{phys}}, 0, 5000)\Big) \tag{6}$$

where $\hat{w}_{\text{phys}}$ is detached from gradient computation so only direct heads and gates receive consistency gradients.

4. **Physics Loss Ramp** $\alpha(t)$:

$$\alpha(t) = \min\left(1, \frac{t}{T_{\text{phys}}}\right), \quad T_{\text{phys}} = 10 \tag{7}$$

where t is the current epoch index, preventing untrained physics heads from destabilizing early gradient updates.

Table 3 lists loss term weight coefficients, and Table 4 details complete training hyperparameters.

Table 3. Loss term weights used in all experiments.

Term	Symbol	Value
Classification	λ_{cls}	1.0
Weight NLL	λ_{wgt}	1.0
Energy NLL	λ_{nrg}	1.0
Volume NLL	λ_{vol}	0.5
Physics consistency	λ_{phys}	0.2

Table 4. Hyperparameters used in all 4-fold cross-validation runs.

Hyperparameter	Value
Epochs per fold	50
Batch size	32
Learning rate	3×10^{-4}
Weight decay	1×10^{-4}
LR warmup	Linear, 5 epochs (0.1 → 1.0)
LR decay	Cosine annealing, $\eta_{\min} = 10^{-6}$
Gradient clip	Max norm = 1.0
Mixed precision	AMP (FP16)
Random seed	42
GPU	NVIDIA RTX 5000 Ada Generation (64 GB VRAM)
CUDA version	12.4

3.6. Reference Portion Targets for Physical Quantity Regression

In the Turkish Food Dataset, nutritional reference data and derived physical targets are anchored to a standard 100 g portion serving baseline ($W_{\text{ref}} = 100\,\text{g}$) per dish category. The corresponding reference volume $V_{\text{ref}}^{(i)}$ (mL) and reference energy $E_{\text{ref}}^{(i)}$ (kcal) for each sample i of target food class y_i are computed using the category-level macroscopic density prior $\rho^{(y_i)}$ and energy density prior $\kappa^{(y_i)}$, which are derived from the class-level nutritional reference database and fixed before cross-validation:

$$V_{\text{ref}}^{(i)} = \frac{W_{\text{ref}}}{\rho^{(y_i)}}, \qquad E_{\text{ref}}^{(i)} = W_{\text{ref}} \cdot \kappa^{(y_i)} \tag{8}$$

where $W_{\text{ref}} = 100\,\text{g}$. These class-level reference targets activate the Gaussian NLL loss components during multi-task optimization and provide the reference values for evaluating Root Mean Square Error (RMSE) for volume, weight, and energy estimation.

3.7. Evaluation Metrics

Classification performance is evaluated using Top-1 Accuracy (Acc@1), Top-5 Accuracy (Acc@5), Macro Precision, Macro Recall, Macro-F_1, and Weighted-F_1:

$$\text{Macro-}F_1 = \frac{1}{C}\sum_{c=1}^{C} F_1^{(c)}, \qquad \text{Weighted-}F_1 = \frac{\sum_{c=1}^{C} n_c \cdot F_1^{(c)}}{\sum_{c=1}^{C} n_c} \tag{9}$$

where n_c is the validation sample count for class c. Cross-fold aggregate metrics are reported as mean ± standard deviation across all $K = 4$ folds. Physical quantity regression performance is evaluated using Root Mean Square Error (RMSE).

4. AIDCare Lifestyle mHealth Platform

AIDCare integrates nutrition, physical activity, and psychological support through a shared longitudinal patient record. The platform separates dietary assessment from exercise planning and does not apply automated calorie-based compensation between the two domains.

In the Nutrition and Diet Module (Figure 2), patients capture meal images, review AI-estimated volume, weight, and caloric content, and submit confirmed meal records. PC^{2}FoodNet provides confidence-calibrated dietary estimates within this workflow. When classification ambiguity or high regression uncertainty is detected, the meal record is routed to a Nutrition Professional (Dietitian) for verification and adjustment before inclusion in the patient's longitudinal dietary record. The Dietitian also develops individualized dietary plans according to the patient's nutritional requirements.

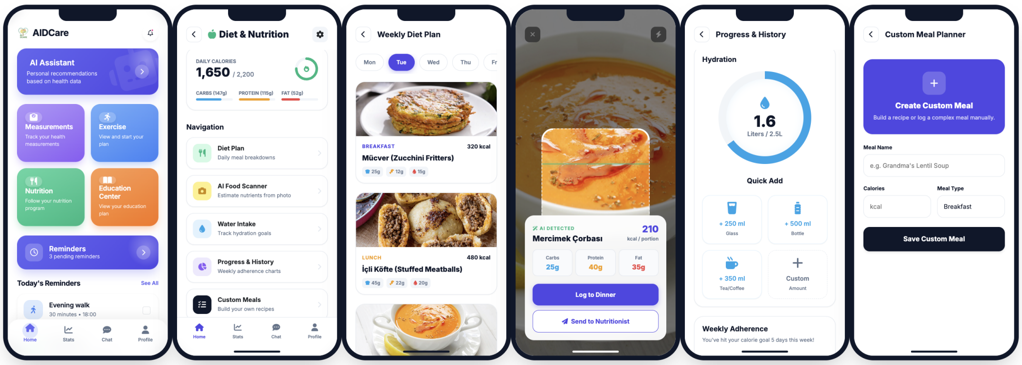

Figure 2. AIDCare Nutrition and Diet Module showing patient's diet plan, AI-based food analysis with nutritional estimation, progress and history tracking, and user-confirmed dietary recording.

The Exercise Module provides an integrated workflow for exercise planning, scheduling, monitoring, and guided delivery (Figure 3). Patients can access exercise information from the shared AIDCare home screen, review activity and vital information, schedule planned activities, and monitor their exercise history. Exercise plans are designed under the guidance of expert physiotherapist, who consider the patient's physical capabilities and relevant physiological context when determining appropriate exercise type, intensity, and duration. The selected exercises are delivered through movement-specific 3D demonstrations developed using Unreal Engine's MetaHuman plugin.

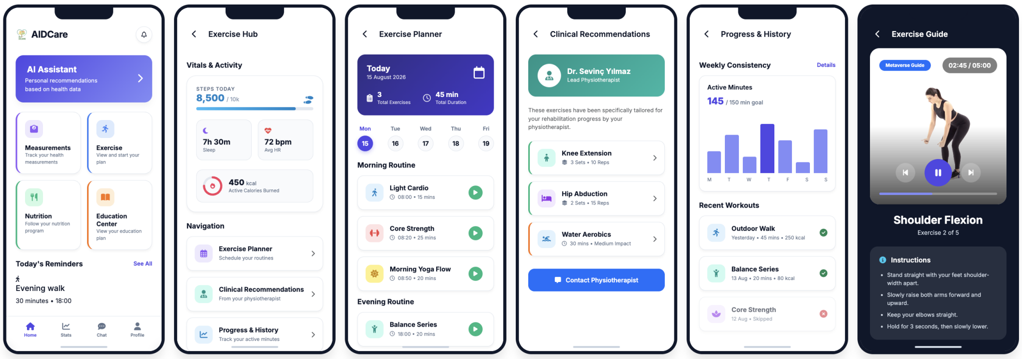

Figure 3. AIDCare Exercise Module showing exercise access, activity and vital information, scheduling, clinician recommendations, adherence tracking, and MetaHuman-based exercise demonstrations.

AIDCare further incorporates multidisciplinary professional oversight through a dedicated professional workflow (Figure 4). By means of a clinican web panel in the AIDCare platform , dietitians can review dietary intake and nutritional targets, physiotherapists can oversee exercise planning and physical safety. Psychotherapists can also provide support related to behavioral adherence, motivation, and patient feedback. This clinician-in-the-loop design maintains professional accountability within each domain while allowing nutrition, exercise, and psychological information to be coordinated through the shared AIDCare platform.

5. Results

5.1. 4-Fold Stratified Cross-Validation Performance

We evaluated PC^{2}FoodNet using a 4-fold stratified cross-validation protocol ($K = 4$) on the 22,070-image Turkish Food Dataset across 40 food categories. In each fold, one stratified partition was used for validation while the remaining partitions were used for training, ensuring that each sample contributed to validation once across the complete cross-validation procedure. The same multi-task learning framework, including the physics-

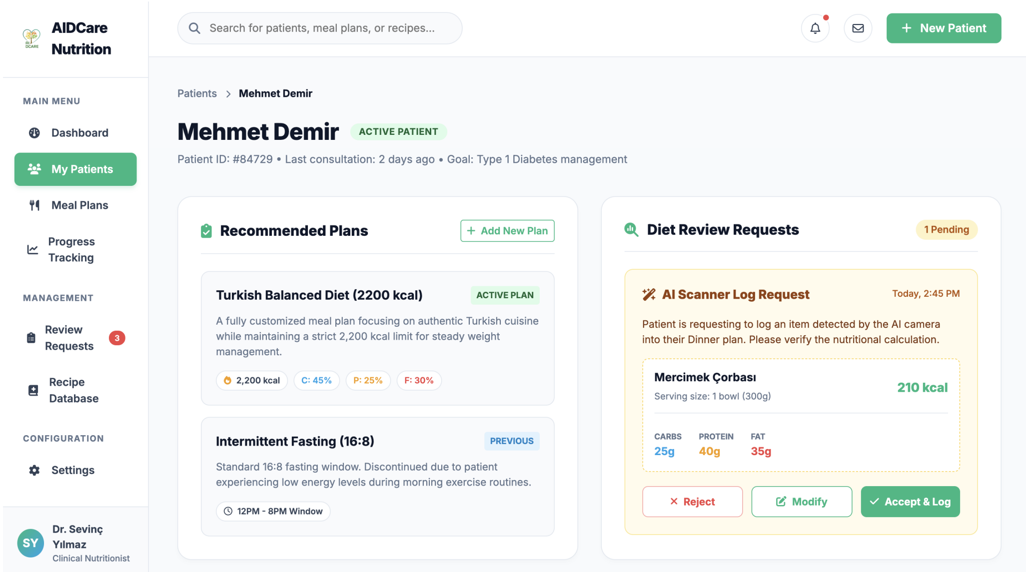

Figure 4. AIDCare Professional Dashboard showing diet review requested by the nutritionst.

constrained physical quantity estimation components, was applied consistently across all four folds. Table 5 summarizes the validation performance at the selected best epoch for each fold.

Table 5. Final 4-fold stratified cross-validation performance of PC^{2}FoodNet on the 40-class Turkish Food Dataset (22,070 images).

Fold	Best Epoch	Val Acc@1 (%)	Val Acc@5 (%)	Macro F_1 (%)	Weighted F_1 (%)	RMSE$_{\text{Vol}}$ (mL)	RMSE$_{\text{Wgt}}$ (g)	RMSE$_{\text{Nrg}}$ (kcal)
1	40	93.74	98.71	92.31	93.68	3.12	30.84	58.21
2	44	94.36	98.91	93.12	94.10	2.86	29.47	56.73
3	47	94.81	99.14	93.64	94.52	3.35	31.26	59.42
4	46	94.65	98.84	93.48	94.35	2.72	29.11	55.84
Mean	–	**94.39**	**98.90**	**93.14**	**94.16**	**3.01**	**30.17**	**57.55**
Std	–	**0.47**	**0.18**	**0.59**	**0.36**	**0.28**	**1.04**	**1.58**

Across the four folds, PC^{2}FoodNet demonstrated stable classification performance, with a mean Top-1 validation accuracy of **94.39%** $\pm$ **0.47%** and Top-5 accuracy of **98.90%** $\pm$ **0.18%**. The corresponding Macro-F_1 and Weighted-F_1 scores were **93.14%** $\pm$ **0.59%** and **94.16%** $\pm$ **0.36%**, respectively. For physical quantity estimation, the mean RMSE values were **3.01** $\pm$ **0.28 mL** for volume, **30.17** $\pm$ **1.04 g** for weight, and **57.55** $\pm$ **1.58 kcal** for energy. The relatively small fold-to-fold variation across the reported metrics indicates stable performance across the different cross-validation folds. Figure **??** presents the training and validation trajectories across the four folds.

5.2. Continuous Physical Quantity Regression

The physical quantity estimation component was evaluated jointly with the food classification task across all four cross-validation folds. PC^{2}FoodNet achieved low estimation errors for volume, weight, and energy across the evaluation folds. The best fold-level results were 2.72 mL for volume, 29.11 g for weight, and 55.84 kcal for energy. Across the four folds, the corresponding mean RMSE values were 3.01 mL for volume, 30.17 g for weight, and 57.55 kcal for energy, with standard deviations of 0.28 mL, 1.04 g, and 1.58 kcal, respectively.

$$\mathrm{RMSE}_{\mathrm{Vol}} = 3.01 \pm 0.28\,\mathrm{mL}, \quad \mathrm{RMSE}_{\mathrm{Wgt}} = 30.17 \pm 1.04\,\mathrm{g}, \quad \mathrm{RMSE}_{\mathrm{Nrg}} = 57.55 \pm 1.58\,\mathrm{kcal}. \tag{10}$$

The relatively small variation between folds indicates stable physical quantity estimation under different training–validation partitions. Figure **??** summarizes the physical quantity estimation performance across the four folds.

5.3. 40-Class Categorical Evaluation

Table 6 presents the per-class precision, recall, F_1-score, and physical quantity estimation errors for the 40 food categories evaluated on the Fold 4 validation set. The per-class analysis provides a detailed view of category-level recognition performance and complements the aggregate cross-validation results by identifying both well-recognized and comparatively challenging food categories.

Table 6. Per-class classification metrics and physical quantity estimation errors on the Fold 4 validation set.

Food Category	**Precision (%)**	**Recall (%)**	F_1 **(%)**	$\mathbf{RMSE_{Vol}}$ (mL)	$\mathbf{RMSE_{Wgt}}$ (g)	$\mathbf{RMSE_{Nrg}}$ (kcal)
Adana Kebap	93.7	97.4	95.5	5.0	40.2	30.5
Aşure	91.4	99.7	95.4	1.6	41.2	7.4
Baklava	100.0	99.5	99.8	1.8	0.6	17.8
Beyaz Ekmek	97.6	96.4	97.0	2.1	40.6	60.1
Bulgur Pilavı	92.3	98.5	95.3	2.3	40.5	31.5
Börek	98.1	98.3	98.2	1.5	2.1	20.9
Esmer Ekmek	96.2	91.1	93.6	2.8	40.0	161.6
Ezogelin Çorbası	90.0	84.4	87.1	2.7	39.2	34.0
Gözleme	98.4	100.0	99.2	1.8	0.5	10.1
Hamburger	97.8	99.2	98.5	2.1	40.9	11.9
Haşlanmış Yumurta	98.8	98.8	98.8	4.5	39.6	12.5
Kadayıf	98.3	77.4	86.6	2.1	0.5	19.9
Karnıyarık	98.8	99.2	99.0	1.9	41.3	8.3
Kumpir	98.7	96.2	97.5	4.4	40.1	22.4
Kuru Fasulye	97.2	95.8	96.5	1.9	40.8	45.4
Köfte	94.8	96.1	95.4	3.5	41.0	41.1
Künefe	90.8	86.8	88.7	3.9	8.7	54.4
Kısır	97.7	97.7	97.7	1.8	41.0	8.7
Lahmacun	98.6	98.6	98.6	2.2	40.7	25.7
Mantı	100.0	100.0	100.0	1.6	41.8	6.3
Menemen	94.7	85.7	90.0	6.6	39.1	16.3
Mercimek Çorbası	85.7	86.7	86.2	2.1	40.0	15.2
Mısır Ekmeği	46.7	75.0	57.5	2.7	4.4	107.8
Nohut	98.7	98.7	98.7	2.1	40.8	20.6
Pastane Poğaçası	100.0	98.4	99.2	2.2	0.5	13.6
Patates Kızartması	97.7	98.8	98.3	2.5	40.5	20.6
Pide	98.6	98.0	98.3	2.2	40.8	48.0
Pirinç Pilavı	98.8	96.4	97.6	2.1	40.6	50.5
Revani	100.0	97.8	98.9	2.1	0.5	11.2
Sütlaç	100.0	99.3	99.7	1.5	41.6	9.5
Tarhana Çorbası	72.8	81.2	76.8	8.6	38.1	25.0
Tavuk	93.7	98.7	96.1	4.1	40.0	58.0
Tavuk Döner	98.6	94.7	96.6	4.3	39.4	99.2
Tulumba Tatlısı	83.1	100.0	90.8	1.8	0.5	4.8
Yaprak Sarma	100.0	100.0	100.0	1.7	41.3	3.0
Yulaf Ekmeği	93.5	96.0	94.7	3.4	40.2	43.8
Çiğ Köfte	100.0	99.4	99.7	2.0	40.8	6.2
İskender	100.0	9.0	16.5	2.5	7.8	209.8
Şakşuka	91.0	94.7	92.8	5.9	40.0	19.4
Şekerpare	97.6	100.0	98.8	1.6	0.5	11.6

The per-class results indicate generally strong recognition performance across most food categories, while also revealing several categories that remain more challenging. High F_1-scores were observed for Mantı (100.0%), Yaprak Sarma (100.0%), Baklava (99.8%), Sütlaç (99.7%), and Çiğ Köfte (99.7%). In contrast, lower scores were observed for İskender

(F_1 = 16.5%), Mısır Ekmeği (F_1 = 57.5%), and Tarhana Çorbası (F_1 = 76.8%). These variations may reflect differences in visual appearance, intra-class variability, and similarity between certain food categories. The per-class analysis therefore provides a complementary view of model behavior beyond the aggregate cross-validation metrics and highlights categories that may benefit from additional training data or further model refinement.

5.4. Confidence-Aware Referral & Multidisciplinary Escalation

The selective confidence-calibrated policy was evaluated on a predefined subset of 420 validation images. The policy flagged 100 cases (23.81%) for optional multidisciplinary reassessment. Among the flagged cases, 39 cases had conformal prediction sets containing 2–4 candidate food classes, indicating classification ambiguity, while 61 cases had singleton predictions but exceeded the validation-derived threshold for predicted regression log-variance $\hat{s}_y$.

These cases were therefore identified for additional professional review before their outputs were treated as sufficiently reliable for downstream use. The confidence-aware referral mechanism provides a human-in-the-loop safeguard by directing uncertain predictions toward professional assessment rather than treating all model outputs as equally reliable. Within the AIDCare framework, such cases can be reviewed by the appropriate professionals according to the nature of the identified uncertainty.

6. Discussion

PC^{2}FoodNet achieves strong food-classification performance while estimating three linked physical quantities from a single RGB image. The central contribution is the integration of class-level macronutrient priors into a differentiable inference chain. By propagating the full class probability distribution into the physical prior calculation, constraining the dependency order of volume, weight, and energy, and incorporating learnable soft fusion gates, the network addresses key limitations of conventional food-computing systems [7–9].

On the 40-class Turkish Food Dataset (22,070 images), 4-fold stratified cross-validation demonstrated consistent classification performance across folds, with a mean Top-1 accuracy of 94.39% ± 0.47%, Top-5 accuracy of 98.90% ± 0.18%, Macro-F_1 of 93.14% ± 0.59%, and Weighted-F_1 of 94.16% ± 0.36%. For physical quantity estimation, the corresponding mean RMSE values were 3.01 ± 0.28 mL for volume, 30.17 ± 1.04 g for weight, and 57.55 ± 1.58 kcal for energy. The best individual fold (Fold 4) achieved RMSE values of 2.72 mL, 29.11 g, and 55.84 kcal for volume, weight, and energy, respectively. The relatively small fold-to-fold variation across the classification and regression metrics indicates stable model behavior under different training–validation partitions.

In the AIDCare platform, integrating the Nutrition and Diet Module (PC^{2}FoodNet) and the Exercise Module under a Human-in-the-Loop Architecture provides coordinated support across lifestyle domains. Standalone digital mHealth applications can suffer from patient attrition or clinical misalignment when automated algorithms provide unmonitored recommendations. By embedding Nutrition Professionals (dietitians), Physiotherapists, and Psychotherapists directly into the decision workflow, AIDCare supports personalized and professionally governed care. Dietitians can verify dietary and macronutrient estimates; physiotherapists guide exercise planning and ensure movement safety; and psychotherapists address behavioral barriers to exercise adherence and dietary consistency. By keeping diet and exercise records uncoupled from automated caloric compensation rules, the system avoids feedback loops in which uncertain calorie predictions automatically trigger inappropriate exercise prescriptions [16,17].

6.1. Limitations and Future Work

This study has several technical and data-related limitations. First, while the Turkish Food Dataset provides 22,070 images across 40 classes, continuous physical quantity regression was evaluated against standard 100 g reference portion targets ($W_{\text{ref}} = 100\,\text{g}$) defined at the dish-category level. These reference targets are derived from category-level nutritional and physical priors rather than direct image-specific measurements of portion size. Second, the dataset comprises single, isolated dishes rather than complex mixed meals, hidden ingredients, or heavy plate occlusions. Third, monocular RGB images lack depth information; volume prediction therefore relies on learned visual priors rather than explicit 3D geometric reconstruction. Fourth, clinical outcomes, user adherence, and long-term glycemic effects were not evaluated in this technical study. Consequently, the reported model performance should not be interpreted as evidence of clinical effectiveness or improved diabetes outcomes.

Future work will expand evaluation to multi-item mixed meals, incorporate depth sensing or reference objects for direct scale recovery, and conduct prospective clinical studies within the AIDCare framework to evaluate patient engagement, multidisciplinary professional agreement, and dietary/exercise logging fidelity.

7. Conclusions

This study presented PC^{2}FoodNet, an AI-based food analysis model developed as part of the AIDCare mHealth platform for diabetes lifestyle management. The proposed model combines food recognition with volume, mass, and energy estimation while using physical food priors and uncertainty estimation to improve the reliability of dietary assessment. Instead of treating every AI prediction as equally reliable, the system identifies uncertain cases and refers them for professional review.

PC^{2}FoodNet was evaluated on 22,070 images covering 40 Turkish food categories using 4-fold stratified cross-validation. The model achieved a mean Top-1 accuracy of 94.39% $\pm$ 0.47% and Top-5 accuracy of 98.90% $\pm$ 0.18%. The corresponding Macro-F_1 and Weighted-F_1 scores were 93.14% $\pm$ 0.59% and 94.16% $\pm$ 0.36%, respectively. For physical quantity estimation, the model achieved mean RMSE values of 3.01 $\pm$ 0.28 mL for volume, 30.17 $\pm$ 1.04 g for mass, and 57.55 $\pm$ 1.58 kcal for energy across the four folds. The best individual fold achieved RMSE values of 2.72 mL, 29.11 g, and 55.84 kcal for volume, mass, and energy, respectively. These results demonstrate that the proposed approach can provide consistent food recognition together with physical quantity estimates from food images.

Within AIDCare, the Nutrition and Diet Module is coordinated with an expert-guided Exercise Module and multidisciplinary professional support. Nutrition Professionals review uncertain dietary estimates, Physiotherapists guide exercise planning, and Psychotherapists support behavioral adherence. This human-in-the-loop design keeps AI in a decision-support role rather than replacing professional care. The proposed framework provides a practical foundation for safer and more personalized digital lifestyle support for people with diabetes.

Author Contributions: Conceptualization, M.J., A.K., M.R., E.G. and A.B.İ.; methodology, M.J., M.R., E.G. and A.B.İ.; software, M.J.; validation, M.J. and A.K.; formal analysis, M.J.; investigation, M.J. and A.K.; resources, A.K., G.S. and H.F.; data curation, M.J.; writing—original draft preparation, M.J., M.R., E.G. and A.B.İ.; writing—review and editing, A.K., H.F. and G.S.; visualization, M.J.; supervision, A.K., G.S. and H.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Excellence in Production Research Framework through XPRES (Excellence in Production Research), the R2Microgrid project under the RESILIENT Competence Center financed by the Swedish Energy Agency and co-financed by Mälardalen University and industrial partners, and the Turkish Health Institutes (TÜSEB) under Grant No. 33987.

Institutional Review Board Statement: Ethical review and approval were not required because this study reports technical model development using food images and reference physical measurements and did not involve a prospective clinical intervention, identifiable patient data, or patient outcomes.

Informed Consent Statement: Not applicable. The study did not involve human participants or identifiable personal or health information.

Data Availability Statement: The derived metrics, out-of-fold predictions, audit files, and figures supporting this study are included in the project materials. The underlying food images and complete training implementation are available from the corresponding author subject to ownership and data-sharing restrictions.

Use of Artificial Intelligence: AI tools, including Gemini Pro 3.1 and Grammarly, were used only to improve grammar, spelling, and language clarity. No part of the research design, implementation, experimental analysis, results generation, or scientific interpretation was generated by these tools. All technical and intellectual contributions in this manuscript were made by the authors.

Conflicts of Interest: The authors declare no conflicts of interest related to this work.

References

1. American Diabetes Association Professional Practice Committee. 5. Facilitating Positive Health Behaviors and Well-Being to Improve Health Outcomes: Standards of Care in Diabetes—2025. *Diabetes Care* **2025**, *48*, S86–S127. https://doi.org/10.2337/dc25-S005.
2. Liang, Z.; Zhang, M.; Wang, C.; Hao, F.; Yu, Y.; Tian, S.; Yuan, Y. The Best Exercise Modality and Dose to Reduce Glycosylated Hemoglobin in Patients with Type 2 Diabetes: A Systematic Review with Pairwise, Network, and Dose–Response Meta-Analyses. *Sports Medicine* **2024**, *54*, 2557–2570. https://doi.org/10.1007/s40279-024-02057-6.
3. Zhao, X.; Forbes, A.; Abu Ghazaleh, H.; He, Q.; Huang, J.; Asaad, M.; Cheng, L.; Duaso, M. Interventions and Behaviour Change Techniques for Improving Physical Activity Level in Working-Age People (18–60 Years) with Type 2 Diabetes: A Systematic Review and Network Meta-Analysis. *International Journal of Nursing Studies* **2024**, *160*, 104884. https://doi.org/10.1016/j.ijnurstu.2024.104884.
4. Xue, H.; Zhang, L.; Shi, Y.; Zhang, H.; Zhang, C.; Liu, Y.; Tan, W.; Liu, Y. The Effectiveness of Digital Health Intervention on Glycemic Control and Physical Activity in Patients with Type 2 Diabetes: A Systematic Review and Meta-Analysis. *Frontiers in Digital Health* **2025**, *7*, 1630588. https://doi.org/10.3389/fdgth.2025.1630588.
5. Zhang, D.; Huang, J.; Zhang, Y.; Wei, Z.; Long, T.; Guo, X.; Li, M. Effects of Continuous Glucose Monitoring on Dietary Behavior and Physical Activity: A Systematic Review and Meta-Analysis. *Diabetes Research and Clinical Practice* **2025**, *229*, 112907. https://doi.org/10.1016/j.diabres.2025.112907.
6. Tarricone, R.; Petracca, F.; Svae, L.; Cucciniello, M.; Ciani, O. Which Behaviour Change Techniques Work Best for Diabetes Self-Management Mobile Apps? Results from a Systematic Review and Meta-Analysis of Randomised Controlled Trials. *eBioMedicine* **2024**, *103*, 105091. https://doi.org/10.1016/j.ebiom.2024.105091.
7. Zheng, J.; Wang, J.; Shen, J.; An, R. Artificial Intelligence Applications to Measure Food and Nutrient Intakes: Scoping Review. *Journal of Medical Internet Research* **2024**, *26*, e54557. https://doi.org/10.2196/54557.
8. Shonkoff, E.; Cara, K.C.; Pei, X.A.; Chung, M.; Kamath, S.; Panetta, K.; Hennessy, E. AI-Based Digital Image Dietary Assessment Methods Compared to Humans and Ground Truth: A Systematic Review. *Annals of Medicine* **2023**, *55*, 2273497. https://doi.org/10.1080/07853890.2023.2273497.
9. Phalle, A.; Gokhale, D. Navigating Next-Gen Nutrition Care Using Artificial Intelligence-Assisted Dietary Assessment Tools—A Scoping Review of Potential Applications. *Frontiers in Nutrition* **2025**, *12*, 1518466. https://doi.org/10.3389/fnut.2025.1518466.
10. Kaushal, S.; Tammineni, D.K.; Rana, P.; Sharma, M.; Sridhar, K.; Chen, H.H. Computer Vision and Deep Learning-Based Approaches for Detection of Food Nutrients/Nutrition: New Insights and Advances. *Trends in Food Science & Technology* **2024**, *146*, 104408. https://doi.org/10.1016/j.tifs.2024.104408.
11. Konstantakopoulos, F.S.; Georga, E.I.; Fotiadis, D.I. A Review of Image-Based Food Recognition and Volume Estimation Artificial Intelligence Systems. *IEEE Reviews in Biomedical Engineering* **2024**, *17*, 136–152. https://doi.org/10.1109/RBME.2023.3283149.
12. Liu, D.; Zuo, E.; Wang, D.; He, L.; Dong, L.; Lu, X. Deep Learning in Food Image Recognition: A Comprehensive Review. *Applied Sciences* **2025**, *15*, 7626. https://doi.org/10.3390/app15147626.

13. American Diabetes Association Professional Practice Committee. 7. Diabetes Technology: Standards of Care in Diabetes—2025. *Diabetes Care* **2025**, *48*, S146–S166. https://doi.org/10.2337/dc25-S007.
14. Moschonis, G.; Siopis, G.; Jung, J.; Eweka, E.; Willems, R.; Kwasnicka, D.; Asare, B.Y.A.; Kodithuwakku, V.; Verhaeghe, N.; Vedanthan, R.; et al. Effectiveness, Reach, Uptake, and Feasibility of Digital Health Interventions for Adults with Type 2 Diabetes: A Systematic Review and Meta-Analysis of Randomised Controlled Trials. *The Lancet Digital Health* **2023**, *5*, e125–e143. https://doi.org/10.1016/S2589-7500(22)00233-3.
15. Nguyen, V.; Ara, P.; Simmons, D.; Osuagwu, U.L. The Role of Digital Health Technology Interventions in the Prevention of Type 2 Diabetes Mellitus: A Systematic Review. *Clinical Medicine Insights: Endocrinology and Diabetes* **2024**, *17*, 11795514241246419. https://doi.org/10.1177/11795514241246419.
16. Ouanes, K.; Farhah, N. Effectiveness of Artificial Intelligence (AI) in Clinical Decision Support Systems and Care Delivery. *Journal of Medical Systems* **2024**, *48*, 74. https://doi.org/10.1007/s10916-024-02098-4.
17. Wilhelm, C.; Steckelberg, A.; Rebitschek, F.G. Benefits and Harms Associated with the Use of AI-Related Algorithmic Decision-Making Systems by Healthcare Professionals: A Systematic Review. *The Lancet Regional Health–Europe* **2025**, *48*, 101145. https://doi.org/10.1016/j.lanepe.2024.101145.
18. Diez Alvarez, S.; Fellas, A.; Wynne, K.; Santos, D.; Sculley, D.; Acharya, S.; Navathe, P.; Gironès, X.; Coda, A. The Role of Smartwatch Technology in the Provision of Care for Type 1 or 2 Diabetes Mellitus or Gestational Diabetes: Systematic Review. *JMIR mHealth and uHealth* **2024**, *12*, e54826. https://doi.org/10.2196/54826.
19. Albalawi, H.F.A. The Role of Tele-Exercise for People with Type 2 Diabetes: A Scoping Review. *Healthcare* **2024**, *12*, 917. https://doi.org/10.3390/healthcare12090917.
20. Doré, B.; Gaudreault, A.; Everard, G.; Ayena, J.C.; Abboud, A.; Robitaille, N.; Batcho, C.S. Acceptability, Feasibility, and Effectiveness of Immersive Virtual Technologies to Promote Exercise in Older Adults: A Systematic Review and Meta-Analysis. *Sensors* **2023**, *23*, 2506. https://doi.org/10.3390/s23052506.
21. Bilika, P.; Karampatsou, N.; Stavrakakis, G.; Paliouras, A.; Theodorakis, Y.; Strimpakos, N.; Kapreli, E. Virtual Reality-Based Exercise Therapy for Patients with Chronic Musculoskeletal Pain: A Scoping Review. *Healthcare* **2023**, *11*, 2412. https://doi.org/10.3390/healthcare11172412.
22. Vecchio, M.; Chiaramonte, R.; Buccheri, E.; Tomasello, S.; Leonforte, P.; Rescifina, A.; Ammendolia, A.; Longo, U.G.; de Sire, A. Metaverse-Aided Rehabilitation: A Perspective Review of Successes and Pitfalls. *Journal of Clinical Medicine* **2025**, *14*, 491. https://doi.org/10.3390/jcm14020491.
23. Min, W.; Wang, Z.; Liu, Y.; Luo, M.; Kang, L.; Wei, X.; Wei, X.; Jiang, S. Large Scale Visual Food Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2023**, *45*, 9932–9949. https://doi.org/10.1109/TPAMI.2023.3237871.
24. Chotwanvirat, P.; Prachansuwan, A.; Sridonpai, P.; Kriengsinyos, W. Advancements in Using AI for Dietary Assessment Based on Food Images: Scoping Review. *Journal of Medical Internet Research* **2024**, *26*, e51432. https://doi.org/10.2196/51432.
25. Lo, F.P.W.; Qiu, J.; Wang, Z.; Chen, J.; Xiao, B.; Yuan, W.; Giannarou, S.; Frost, G.; Lo, B. Dietary Assessment with Multimodal ChatGPT: A Systematic Analysis. *IEEE Journal of Biomedical and Health Informatics* **2024**, *28*, 7577–7587. https://doi.org/10.1109/JBHI.2024.3417280.
26. Sharma, V.; Chadha, R. Development and Evaluation of Food Photograph Series Software for Portion Size Estimation among Urban North Indian Adults. *Mediterranean Journal of Nutrition and Metabolism* **2023**, *16*, 293–312. https://doi.org/10.3233/MNM-230039.
27. Kim, J.h.; Lee, D.s.; Kwon, S.k. Food Classification and Meal Intake Amount Estimation through Deep Learning. *Applied Sciences* **2023**, *13*, 5742. https://doi.org/10.3390/app13095742.
28. Jia, W.; Li, B.; Xu, Q.; Chen, G.; Mao, Z.H.; McCrory, M.A.; Baranowski, T.; Burke, L.E.; Lo, B.; Anderson, A.K.; et al. Image-Based Volume Estimation for Food in a Bowl. *Journal of Food Engineering* **2024**, *372*, 111943. https://doi.org/10.1016/j.jfoodeng.2024.111943.
29. Shao, Z.; Vinod, G.; He, J.; Zhu, F. An End-to-End Food Portion Estimation Framework Based on Shape Reconstruction from Monocular Image. In Proceedings of the 2023 IEEE International Conference on Multimedia and Expo (ICME), 2023, pp. 942–947. https://doi.org/10.1109/ICME55011.2023.00166.
30. Magid, B.; Ibrahim, M.; Kawashti, Y.A.; Mohamed, M.; Sabry, M.; Hindy, H.; Khaled, M.; Mohamed, W. CalorieMe: An Image-Based Calorie Estimator System. In Proceedings of the 2023 Eleventh International Conference on Intelligent Computing and Information Systems (ICICIS), 2023, pp. 555–560. https://doi.org/10.1109/ICICIS58388.2023.10391113.
31. Gautam, S.; Patnaik, S.S.; Kush, R.; Pantola, D.; Mishra, V.K. Enhancing Food Analysis with Attention-Based Deep Learning: Ingredients, Recipes, and Calorie Estimation. *Procedia Computer Science* **2025**, *260*, 692–700. https://doi.org/10.1016/j.procs.2025.03.248.
32. Nogay, H.S.; Nogay, N.H.; Adeli, H. Image-Based Food Groups and Portion Prediction by Using Deep Learning. *Journal of Food Science* **2025**, *90*, e70116. https://doi.org/10.1111/1750-3841.70116.

33. Saad, A.M.; Rahi, M.R.H.; Islam, M.M.; Rabbani, G. Diet Engine: A Real-Time Food Nutrition Assistant System for Personalized Dietary Guidance. *Food Chemistry Advances* **2025**, *7*, 100978. https://doi.org/10.1016/j.focha.2025.100978.
34. Velombe, J.C.; Bayraktar, S.; Kavak, A.; Jamil, M.; İnner, A.B.; Srivastava, G.; Fotouhi, H. A Hybrid CNN–MLLM Architecture for Image-Based Nutrition Estimation and Advisory Insulin Decision Support in Type 1 Diabetes. *Nutrients* **2026**, *18*, 2205. https://doi.org/10.3390/nu18132205.
35. Göz, F.; Jamil, M.; Kavak, A.; Bayraktar, S.; Doğru, A.C.; Srivastava, G.; Fotouhi, H. A Confidence-Aware Hybrid Vision–Language Framework for Food Recognition and Nutritional Monitoring. *Nutrients* **2026**, *18*, 2449. https://doi.org/10.3390/nu18152449.
36. Cheng, S.T.; Lyu, Y.J.; Teng, C. Image-Based Nutritional Advisory System: Employing Multimodal Deep Learning for Food Classification and Nutritional Analysis. *Applied Sciences* **2025**, *15*, 4911. https://doi.org/10.3390/app15094911.