ISSN: 3139-0862

NUTRICOACH PRO: AN AGENTIC REAL-TIME RETRIEVAL-AUGMENTED GENERATION AND MULTIMODAL INTELLIGENCE SYSTEM FOR PRECISION CLINICAL NUTRITION AND METABOLIC HEALTH TRACKING

Kubera B, Sharath S R, Anoushka, Lavanya S U, Maseeha Banu

Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), Malnad College of Engineering, Hassan, Karnataka, India

Corresponding author: Kubera B | kkubera307@gmail.com

Abstract—

Metabolic syndrome is a growing global cardiometabolic burden and is increasingly accompanied by demand for continuous, personalized, and evidence-grounded nutrition support. A 2025 systematic review and modelling analysis estimated that 1.54 billion adults were living with metabolic syndrome in 2023, with global prevalence estimated at 28.4% across women and men. Conventional nutrition trackers provide structured quantitative logging, but manual food search, portion selection, and fragmented workflows can increase interaction burden. Large language models (LLMs) provide a natural conversational interface, yet unrestricted generation can produce unsupported nutritional or clinical statements. This paper presents NutriCoach Pro, an agentic nutrition-intelligence architecture that separates deterministic computation, verified food data, evidence retrieval, multimodal perception, and language generation. The system combines a Mifflin-St Jeor energy-estimation engine, activity-adjusted energy expenditure, configurable macronutrient targets, hydration tracking, a bounded Health Compliance Score, deterministic multi-item food parsing, multimodal meal recognition, a seven-domain clinical knowledge base, local retrieval-augmented generation (RAG), optional cloud-model adapters, schema-controlled database actions, and WebSocket-based real-time synchronization. A prototype evaluation plan is defined for NLP extraction, calorie-estimation error, unsupported clinical claims, latency, multimodal recognition, and user task performance. Project-reported prototype figures supplied for this manuscript are F1=96.4%, mean absolute calorie error=12.3 kcal, local-RAG latency=42.1 ms, WebSocket synchronization latency=28.4 ms, and unsupported-claim rate below 1.2%; these values are explicitly treated as provisional until reproduced with a documented held-out test protocol. The contribution is an explainable, modular, privacy-oriented architecture that places hard numerical and safety boundaries around generative Al rather than treating an LLM as the source of truth or as an autonomous Clinical decision maker.

Keywords— retrieval-augmented generation, multimodal AI, nutrition tracking, metabolic health, agentic AI, large language models, food recognition, mHealth, clinical safety, personalized nutrition

INTRODUCTION

Metabolic syndrome is a cluster of cardiometabolic abnormalities associated with increased risk_ of cardiovascular disease and type-2 diabetes. A systematic review and Bayesian modelling analysis published in Nature Communications estimated global metabolic-syndrome prevalence at 28.4% among women and men in 2023, compared with 11.9% in 2000, and estimated 1.54 billion adults affected in 2023. The same analysis estimated 31.0% prevalence among women and 25.7% among men in 2023, with substantial regional variation and particularly large burdens in East and South Asia [1]. These estimates establish the scale of the public-health problem but do not imply that consumer software can diagnose or treat metabolic disease.

Digital nutrition systems can support self-monitoring, education, and behavior change. Reviews of mobile nutrition applications report potential benefits, but also identify heterogeneity in functionality, adherence, nutrient coding, and_ validation. Dietary-record apps can underestimate intake relative to reference methods, showing why transparent data provenance and validation are important.

NutriCcoach Pro addresses this problem as a systems-engineering contribution. The architecture treats numerical nutrition values as structured data, evidence as retrieved knowledge, and natural-language generation as a presentation layer. The central design hypothesis is that separating these responsibilities can improve traceability and reduce unsupported clinical generation.

The principal contributions are: (1) a deterministic metabolic computation engine; (2) a multi-item dietary parser with verified-food matching; (3) a seven-domain evidence repository and local RAG pipeline; (4) a multimodal food-ingestion pathway; (5) a schema-controlled agentic action controller; (6) a bounded adherence score; and (7) an experimental framework for measuring extraction, nutrition-estimation, safety, latency, and usability performance.

Fig. 1. Four-tier NutriCoach Pro system architecture

CLIENT PRESENTATION TIER

PWA * Moblle/Desktop Dashboard * Food Camera * voice/Text * SVG Health Rings

GATEWAY & INGESTION

Flask REST « Flask-SocketlO + Auth » NLP Tokenizer + Input Normalization + Intent Classifier

KNOWLEDGE & DATA

SQLite/Transactional DB + 200+ verified foods » User/Meal/Water/Workout Logs » 7-domain evidence corpus « Vector index

MULTI-LLM AGENTIC ENGINE

Deterministic Metabolic Engine + Local RAG + Vision Adapter + Gemini/OpenAl/Groq/OpenRouter Adapters + Action Controller

Fig. 3. Four-tier NutriCoach Pro system architecture.

TABLE II — GLOBAL METABOLIC SYNDROME BURDEN,

2000- 2023 [1]

2023

. 2000 2023 .

ISSN: 3139-0862

Population estimated
prevalence prevalence
cases
Women 14.7% 31.0% 846 million
Men 9.0% 25.7% 692 million
Combined 11.9% 28.4% 1.54 billion
Global metabolic syndrome prevalence, 2000 vs 2023
mm 2000 MM 2023
30 4
&
2 204
o
$
oO
& 107
(0) A
Women Men Combined

Fig. 2. Modelled global metabolic syndrome prevalence in 2000 and 2023 [1].

RELATED WORK

Mobile nutrition applications have evolved from manual food selection toward barcode, voice, photo, and coaching features. Systematic reviews show that smartphone dietary tools commonly combine photo-based assessment, portion estimation, free-text input, food databases, and classification. Other reviews indicate that app-based interventions can improve' selected nutrition and weight-related outcomes, although evidence quality and intervention designs vary.

Image-assisted dietary assessment is an active research area. Asystematic review of image-based food-recognition systems found that the pipeline typically includes image acquisition, preprocessing, segmentation, classification, volume estimation, and nutrient lookup. More recent reviews report rapid adoption of deep learning while emphasizing portion-size estimation and validation as persistent challenges.

LLM-based healthcare systems create a second research problem. A systematic review of 519 health-care LLM evaluation studies found that only 5% used _ real patient-care data and that fairness, bias, toxicity, calibration, uncertainty, and deployment considerations were much less frequently evaluated than accuracy [16]. A 2026 systematic review of _ hallucination-mitigation strategies identified retrieval-augmented generation, knowledge graphs, self-reflection, specialized evaluation, human-in-the-loop methods, training techniques, and red teaming as major approaches [17]. RAG therefore provides an explicit mechanism for retrieving non-parametric evidence before generation, making it suitable for an evidence-aware nutrition assistant.

NutriCoach Pro differs from a conventional chatbot because the LLM is not permitted to invent calories, directly modify arbitrary database records, or act as an autonomous clinical authority. Instead, deterministic functions, evidence retrieval, and typed tools form explicit boundaries around generation.

TABLE | — POSITIONING OF NUTRICOACH PRO
an Conventional Vanilla NutriCoach
Capability
tracker LLM Pro
Structured Not
Yes Yes
nutrition database guaranteed
Natural-language a ;
; Limited/varies Yes Yes
meal logging
Deterministic Not
; ; Yes Yes
numerical engine guaranteed
. . . Not
Evidence retrieval Varies Yes
guaranteed
Multimodal food Increasingly
. ; Yes Yes
input available
Unsafe
Schema-controlled . ;
. Limited without Yes
actions
controls
Core
Local RAG option Product-dependent | Rare ;
design
Clinical safety ; Requires Explicit
; Varies
routing controls layer

PROBLEM STATEMENT AND DESIGN REQUIREMENTS

The system must transform low-structure user input into reliable structured nutrition records while preserving a distinction between measured data, estimated values, retrieved evidence, and generated explanations. The requirements were therefore defined around correctness, traceability, safety, latency, privacy, and usability.

. R1: Accept text, voice-transcribed text, and meal photographs.
* R2: Parse multiple foods, quantities, and units from one interaction.
* R83: Resolve nutrition values from a verified database rather than free-form model memory.
+ R4: Estimate energy expenditure using an explicit mathematical model.
+ R5: Retrieve evidence for health-related explanations.
+ R6: Restrict database-changing actions to typed and authorized tools.
* R7: Synchronize dashboard state in real time.
* R8: Provide uncertainty and safety messages where inputs are incomplete.
* R9: Support local/offline processing for sensitive workflows where feasible.

ISSN: 3139-0862

* R10: Produce reproducible logs of model, source, latency, and action provenance.

MATHEMATICAL METABOLIC ENGINE

The metabolic engine converts user attributes and logged activity into transparent estimates. The system intentionally labels these quantities as estimates rather than direct physiological measurements.

Resting energy expenditure is estimated with the Mifflin-St Jeor equation. For males, BMR = 10W + 6.25H - 5A + 5; for females, BMR = 10W + 6.25H - 5A - 161, where W is body mass in kilograms, H is height in centimetres, and Ais age in years.

Total daily energy expenditure is represented as TDEE = BMRx a, where ais an activity multiplier. A prototype may expose a range such as 1.20- 1.90, but the selected multiplier remains an estimate and should not be interpreted as measured expenditure.

For a configurable energy goal, E_target = TDEE - D for a deficit, E_target = TDEE for maintenance, or E_target = TDEE + S fora surplus. The application should not enforce universal calorie floors because _ individual energy requirements and clinical contexts differ.

Protein is represented as P_target = pW, where p is a configurable g/kg/day parameter. The proposed 1.8 g/kg/day value is treated as an optional high-protein configuration, not a universal clinical requirement. Fat energy is E_F = 9F and protein energy is E_P = 4P. Residual carbohydrate allocation is C = [E_target- (4P + 9F)]/4.

Daily net energy is represented as E_net = 2E_consumed - YE_burned. Because both intake and expenditure can contain estimation error, the dashboard reports these as tracked estimates rather than exact physiological balance.

TABLE Ill — WORKED NUMERICAL EXAMPLE (ILLUSTRATIVE, NOT A CLINICAL PRESCRIPTION)

Input Value
Body mass W 70kg
Height H 175cm
AgeA 25 years
Sex Male
Activity factor a 1.55
Estimated BMR 1,673.75 kcal/day
Estimated TDEE 2,594.31 kcal/day
Illustrative 10% deficit target 2,334.88 kcal/day

Fig. 3. Deterministic metabolic and macronutrient calculation engine

User profile
W+H+A+sex
Mifflin-St Jeor
BMR
Activity factor
TDEE = BMR x a pene
Bo
mainte
surplus
ZL
Protein target Fat allocation Residual carbs
P= pw 2 EF =9F [| c= rl

Fig. 4. Deterministic metabolic and macronutrient calculation engine.

HEALTH COMPLIANCE SCORE

NutriCoach Pro introduces a Health Compliance Score (HCS) for longitudinal visualization of adherence. It is a software metric and is not presented as a validated clinical outcome.

Let S_cal, S_pro, S_water, S_activity, and S_quality be component scores in [0,100], with non-negative weights that sum to one. HCS = w_cS_cal + w_pS_pro + w_wS_water + w_aS_activity + w_qS_quality. Because every component is bounded and the weights sum to one, 0<HCS < 100.

A continuous calorie-adherence component may be defined as S_cal = 100 max(0, 1 - |E_consumed - E_target|/t_c). Protein adequacy may use S_pro = 100 min(1, P_consumed/P_target), while hydration may use S_water = 100 min(1, W_consumed/W_target). These functions are configurable and must not be interpreted as medical thresholds.

Boundary proof: each S_i lies in [0,100]. A convex combination of bounded values is also bounded by the same interval; therefore HCS is guaranteed to remain within [0,100].

Fig. 7. Health Compliance Score computation

Calorie Protein Diet
adherence adequacy Hydration activity: quality
Weighted HCS
0 <HCS < 100
Longitudinal dashboard
indicator (not a clinical outcome)

Fig. 5. Health Compliance Score computation and bounded aggregation.

SYSTEM ARCHITECTURE

The proposed system follows a four-tier architecture: client presentation, gateway and ingestion, knowledge and data, and multi-LLM agentic engine. The separation is intended to make numerical computation, persistence, retrieval, generation, and actions independently testable.

The client is a responsive Progressive Web Application with dashboard cards, health rings, meal history, camera input, conversational input, and mobile navigation. The gateway provides REST APls, authentication, input normalization, and Socket.lIO events. The data layer stores users, meals, water, workouts, Al logs, food records, and evidence chunks. The agentic engine coordinates local RAG, optional cloud-model adapters, visual reasoning, deterministic calculators, and typed actions.

DATA MODEL AND KNOWLEDGE LAYER

The relational model is intentionally separated from the knowledge corpus. Transactional tables contain user and activity state; food records provide canonical nutrient values; evidence chunks store source text, domain labels, provenance, and retrieval metadata.

ISSN: 3139-0862

The core entities are users, meals, water_logs, workout_logs, ai_coaching_logs, food_database, and evidence_chunks. The food database should include canonical names, aliases, serving units, calories, macronutrients, and source provenance. The evidence layer should include publication year, source organization, domain, evidence text, and version identifiers.

A production deployment can replace prototype storage with PostgreSQL or another transactional database and use a vector index for evidence retrieval. USDA FoodData Central is one example of a structured nutrition-data source that publishes downloadable data and documentation.

TABLE IV — CORE DATA DICTIONARY

Entity Principal fields Purpose
id, age, sex,
height_cm, Profile and target
users : ;
weight_kg, goal, generation
activity
user_id, food_id,
meals quantity, unit, kcal, Dietary transactions
protein, carbs, fat
user_id, volume_ml, ; ;
water_logs Hydration history
timestamp
user_id, activity, a .
workout_logs Activity history
duration, kcal
. ; intent, query, model, _
ai_coaching_logs Al traceability
sources, latency
name, aliases, .
Verified food
food_database ;
resolution
serving, nutrients,
source
domain, source,
RAG knowledge
evidence_chunks year, text,
layer
embedding

Fig. 6. Relational and knowledge-store schema

Fig. 6. Relationaland knowledge-store schema.

MULTIMODAL DIETARY INGESTION

The multimodal pipeline accepts a meal image or a natural-language description. Visual reasoning generates candidate food identities and approximate portions, while a verified food database resolves nutrient values. This division avoids asking a generative model to hallucinate exact calorie values directly from pixels.

For image input, the pipeline performs image quality checks, visual candidate generation, optional spatial segmentation, portion estimation, food-database matching, and nutrition calculation. For text input, the system performs normalization, delimiter splitting, quantity-unit extraction, alias matching, and database lookup.

Image-based dietary assessment remains an imperfect measurement problem. A systematic review of 159 screened studies and 78 _ included image-based food-recognition studies found a common pipeline of segmentation, classification, volume estimation, and nutrient lookup, while emphasizing persistent challenges in portion estimation and validation [7]. A meta-analysis of image-assisted dietary assessment also reported uncertainty in agreement with reference methods [8]. Therefore, NutriCoach Pro should expose confidence and uncertainty instead of presenting visual estimates as exact measurements.

Fig. 4. Dual-path multimodal dietary ingestion architecture

aN Multimodal visual
reasoning
Meal photo
Candidate foods Verified
+ portions DB
aN NLP parser + va
unit extraction
Text / voice

Fig. 7. Dual-path multim odal dietary ingestion architecture.

ALGORITHM 1 — DETERMINISTIC MULTI-ITEM FOOD PARSING

Input: natural-language dietary statement X. Output: structured list F of food records.

Step 1: Normalize case, punctuation, Unicode fractions, and common unit variants. Step 2: Split using commas, conjunctions, plus signs, semicolons, and line breaks. Step 3: Extract quantity q and unit u using regular expressions. Step 4: tokenize the food phrase. Step 5: compute Jaccard token overlap between the input and candidate food aliases. Step 6: add an alias boost when a canonical synonym is matched. Step 7: select the highest-scoring candidate subject to a confidence threshold. Step 8: resolve calories and nutrients from the food database. Step 9: return structured records and uncertainty flags.

Jaccard similarity is J(A,B) = |AN B| / |AU B|. Candidate selection is f* = argmax_f score(f,X). If the best candidate does not meet the threshold, the system asks a clarification question rather than silently selecting a low-confidence food.

“

Example input: “ 2 eggs, 1 banana and 250 ml milk.” Expected structured output contains three food entities with quantities 2, 1, and 250 ml. Nutritional values are resolved from the database rather than generated by the language model.

Fig. 2. End-to-end multimodal ingestion and response pipeline

Fig. 8. End-to-end multimodal ingestion and response pipeline.

CLINICAL KNOWLEDGE BASE

The evidence repository is divided into seven domains. Domain 1 covers energy balance and calorie management. Domain 2 covers protein adequacy and amino-acid

ISSN: 3139-0862

considerations. Domain 3 covers carbohydrate quality, dietary fibre, and lipid quality. Domain 4 covers hydration and fluid balance. Domain 5 covers chrononutrition and time-restricted eating. Domain 6 covers satiety and energy-density substitutions. Domain 7 covers nutrition considerations for PCOS and metabolic conditions.

WHO guidance supports dietary patterns rich in vegetables, fruit, legumes, nuts, and whole grains and recommends limiting free sugars to less than 10% of total energy, with additional benefit suggested below 5%. WHO also recommends limiting saturated fat and trans fat and reducing sodium intake. These values are used as evidence-grounded educational rules, not individualized prescriptions.

Time-restricted eating is represented as an optional dietary pattern rather than a universal intervention. Recent meta-analyses report mixed and context-dependent effects across weight, metabolic outcomes, and_ hunger, reinforcing the need for evidence-grounded responses. individualized and

For PCOS, the 2023 international guideline emphasizes healthy lifestyle management and does not identify one universally superior diet or exercise pattern. NutriCoach Pro therefore uses condition-aware educational content and escalation rather than diagnosis or treatment.

TABLE V — EVIDENCE-DOMAIN ORGANIZATION

Domain	Primary content	Safety posture
D1 Energy balance	BMR, TDEE, calorie targets	Estimate; no guaranteed outcome
D2 Protein	Protein adequacy, distribution, amino acids	Configurable; not universal
D3 Carbs/lipids	Fibre, whole grains, unsaturated fats	General dietary guidance
D4 Hydration	Fluid balance, activity, heat	Context-dependent estimate
D5 Chrononutrition	TRE, eating windows	Optional; mixed evidence
Energy density,	Educational	
D6 Satiety sa: .	substitutions suggestions	
D7 Metabolic PCOS, insulin Escalation; not conditions resistance, T2D diagnosis/treatment		

LOCAL RAG AND MULTI-LLM ENGINE

The local RAG pipeline follows Query — embedding — retrieval = evidence filtering = context construction — generation — provenance. The local store is designed to minimize unnecessary transmission of sensitive health information. The exact latency depends on hardware, embedding model, index size, and query length; therefore, latency claims require an explicit benchmark environment.

Cloud adapters can connect to external LLM providers for advanced reasoning or multimodal capabilities. Provider abstraction isolates application logic from any single vendor and allows the deployment to choose between local-only, hybrid, or cloud-assisted operation.

The model is not treated as a source of truth. If no sufficiently relevant evidence is retrieved, the assistant should state that the evidence is unavailable or provide a general educational answer with an explicit limitation. For high-risk clinical questions, the system should route to safety messaging rather than autonomous treatment advice.

Fig. 3. Local retrieval-augmented generation pipeline

Fig. 9. Local retrieval-augm ented generation pipeline.

AGENTIC ACTION CONTROLLER

Natural-language actions such as “ log two eggs for breakfast” or“ add 500 mlof water” are converted into typed tool calls. The LLM produces an intent and arguments, but the server validates the schema, checks authorization, executes a known function, commits the transaction, and then emits a real-time update.

This design prevents the language model from directly SQL. A ~ valid path is: natural-language command -— intent classifier — typed constructing arbitrary action schema — permission and validation — database transaction = WebSocket event. Invalid or ambiguous requests are rejected or clarified.

Representative events include stats_update_<uid> and water_update_<uid>. Acknowledgements and idempotency keys should be added in production to avoid duplicate writes during reconnects.

Fig. 5. Agentic action-control loop with validation boundary

Natural-language command
Intent classifier
Reject / clarify ‘Typed action schema WebSocket event
Permission + validation
Database transaction

ISSN: 3139-0862

Fig. 10. Agentic action-control loop with validation boundary.

REAL-TIME SYNCHRONIZATION

Flask-SocketIO provides bidirectional application events. When a meal, water record, or goal is updated, the server recalculates relevant dashboard aggregates and emits a user-scoped event. The client updates the affected components without a full page reload.

The real-time channel should be treated as an optimization rather than the source of truth. The database remains authoritative, and clients should reconcile state after reconnect. Production deployments should include event identifiers, acknowledgements, retry limits, duplicate suppression, and connection-health monitoring.

SECURITY, PRIVACY, ETHICS, AND CLINICAL SAFETY

Nutrition and health information can be sensitive personal data. The architecture therefore applies data minimization, encrypted transport, password hashing, least-privilege access, audit logging, retention controls, and explicit consent. A local-processing mode can reduce exposure of sensitive records to external model providers.

HIPAA should not be claimed automatically for a consumer application. HHS states that HIPAA Rules apply to covered entities and business associates, with specific contractual and safeguarding obligations in covered relationships [18]. GDPR Article 9 treats data concerning health as a special category of personal data subject to additional conditions for processing [19]. Accordingly, the manuscript describes NutriCoach Pro as privacy-oriented and designed to support applicable compliance requirements rather than claiming certification.

The application must not diagnose disease, prescribe medication, alter insulin dosage, or replace a clinician. Pregnancy, eating disorders, kidney disease, insulin-treated diabetes, severe cardiovascular disease, acute symptoms, and other high-risk situations should trigger conservative messaging and professional-care escalation.

Visual food recognition should disclose uncertainty. Clinical answers should provide evidence provenance. User-specific recommendations should distinguish general education from medical advice.

EXPERIMENTAL METHODOLOGY

The evaluation is organized into five experiments. Experiment A measures food-entity and quantity extraction on a held-out natural-language meal corpus. Experiment B measures unsupported clinical claims using a manually annotated health-question benchmark. Experiment C measures local-RAG and _ end-to-end latency under controlled workloads. Experiment D evaluates image-based food recognition and calorie estimation against weighed-food or reference nutrition records. Experiment E measures user task completion, correction rate, and perceived usability.

For NLP, precision, recall, and F1 are reported for food identity, quantity, and unit extraction. Calorie estimation uses mean absolute error (MAE) and mean absolute percentage error (MAPE) where a non-zero reference is available. For safety, the unsupported-claim rate is defined as the number of responses containing a materially unsupported clinical claim divided by the total evaluated responses.

Latency is reported using median, P95, and P99 rather than a single mean. The benchmark record must include CPU/GPU, RAM, operating system, software versions, model version, embedding model, corpus size, query count, and date. Image evaluation should report both simple single-food and mixed-meal subsets.

Fig. 8. Experimental evaluation framework

Held-out NLP Clinical QA Latency Image meal User task set set set

\

Metrics: F1 + MAE + unsupported-claim ra

P50/P95 latency * image error + task success

producibility: sample size » model version * hardware + date + protocbl

Fig. 11. Experimental evaluation framework.

PROTOTYPE-REPORTED RESULTS AND REPRODUCIBILITY

The project specification supplied for this manuscript reports an NLP extraction F1 of 96.4%, mean absolute calorie error of 12.3 kcal, local-RAG latency of 42.1 ms, WebSocket synchronization latency of 28.4 ms, and an unsupported clinical-hallucination rate below 1.2%. These are retained as project-reported prototype values, not independently verified results.

A publication-grade results section must reproduce each figure with a frozen dataset and protocol. In particular, the hallucination comparison values against GPT-40 and GPT-3.5 should only be reported if the exact model versions, prompts, sample size, annotation rubric, and evaluator agreement are documented.

TABLE VI — PROTOTYPE-REPORTED VALUES PENDING INDEPENDENT REPRODUCTION

Metric Reported value Evidence status

; Project-reported;

NLP extraction F1 96.4%

reproduce

Mean absolute Project-reported;

; 12.3 kcal

calorie error reproduce

Project-reported;

Local RAG latency 42.1 ms

reproduce

WebSocket ,

oo Project-reported;

synchronization 28.4 ms

reproduce

latency

Unsupported Project-reported;

<1.2%

ISSN: 3139-0862

clinical-claim rate reproduce
Cloud-model
Protocol and model
baseline 18.6% / 24.5% . .
oo versions required
hallucination rates

DISCUSSION

The principal contribution is architectural rather than the invention of a new nutrition equation. Existing nutrition applications demonstrate the utility of self-monitoring, while food-image research demonstrates the potential and

LIMITATIONS OF AUTOMATED DIETARY ASSESSMENT. NUTRICOACH

Pro combines these capabilities with RAG and an explicit action-control layer.

The architecture provides a useful separation of concerns: deterministic code answers numerical questions, structured databases supply nutrient values, retrieval supplies evidence, the action controller performs validated operations, and the LLM produces user-facing explanations. This separation supports auditability and makes it possible to evaluate components independently.

The system nevertheless cannot eliminate uncertainty. Food composition depends on recipe, brand, preparation, and portion size; visual estimation can be wrong; energy expenditure is estimated; evidence can become outdated; and external model behavior can change. These limitations should be exposed in the user interface rather than hidden behind conversational fluency.

LIMITATIONS

First, the prototype does not establish clinical efficacy. Second, the HCS is a software adherence metric and has not been clinically validated. Third, the proposed protein, hydration, and fasting parameters are configurable heuristics and should not be interpreted as universal requirements. Fourth, the project-reported benchmark values require reproduction. Fifth, image-based portion estimation remains uncertain. Sixth, local RAG latency is deployment-dependent. Seventh, the knowledge base requires versioning and periodic evidence review.

FUTURE WORK

Future work should include a larger, diverse, culturally representative food corpus with Indian and regional foods; weighed-food validation; uncertainty calibration; multilingual support for English, Kannada, and Hindi; on-device vision; wearable integration; glucose and activity data; human-in-the-loop' dietitian review; longitudinal adherence studies; privacy-preserving federated learning; and a formal clinical-safety evaluation.

An important next step is an ablation study comparing deterministic-only, LLM-only, RAG-only, and complete agentic configurations. Another is a red-team evaluation of prompt injection, malicious database actions, unsafe clinical requests, and conflicting evidence. Finally, the should

clinical-validation studies before any medical deployment claim.

CONCLUSION

NutriCoach Pro presents a modular architecture for system undergo prospective usability and conversational nutrition and metabolic-health tracking that combines deterministic metabolic computation, structured food data,
retrieval-augmented generation, agentic tool execution, and multimodal dietary assessment,
real-time synchronization. The architecture is designed around a simple safety principle: the language model should explain and orchestrate, but it should not become the authoritative source of numerical nutrition data or autonomous Clinical treatm ent.

The proposed evaluation framework makes the system measurable at the levels of extraction, estimation, evidence grounding, latency, action correctness, and user interaction. The current project-reported prototype values provide a starting point for validation, but they should not be interpreted as independently established clinical or scientific
performance until reproduced under a
documented protocol.

Al-ASSISTED DEVELOPMENT AND RESEARCH INTEGRITY

Al-assisted coding and writing tools may be used during implementation and manuscript preparation. The authors remain responsible for system architecture, source verification, experimental design,
interpretation,
citations, and final
numerical
results,
submission. Al
assistance does not substitute for clinical review, data validation, or reproducible experimentation.

ALGORITHM 1 — DETERMINISTIC MULTI-ITEM FOOD PARSING

Input: user text X; verified food database D; similarity threshold 8.

1. Normalize case, punctuation,
fractions.
2. Split X on
commas, conjunctions,
semicolons, and line breaks.
units,
and Unicode
plus signs,
3. For each segment, extract quantity q and unit u.
4. Tokenize the remaining food phrase and generate candidate foods from D.
5. Compute Jaccard overlap J(A,B)=|AN B[/|AU B| and apply alias boost where applicable.
6. Select f*=argmax score(f,X).

ISSN: 3139-0862

clarification.
If score<6,
request
7. Resolve nutrient values from D and scale by q/u.
8. Persist the validated
corresponding dashboard event.
Output: structured food
confidence metadata.
SUGGESTED ABLATION STUDY
TABLE VII —
records'. with
transaction and emit the
source and
Configuration
Purpose
Primary metrics
Baseline for
Deterministic only ; ; F1, MAE
parsing/calculation
Measures Fl,
LLM only unconstrained unsupported-claim
generation rate
Measures evidence Unsupported-claim
RAG only
grounding
rate, latency
RAG + typed tools
Measures safe
action execution
Action accuracy,
duplicate rate
End-to-end All metrics + task
Complete system ;
evaluation success

REFERENCES

[1] J. J. Noubiap et al., " Worldwide trends in metabolic syndrome from 2000 to 2023: a systematic review and modelling analysis,"
resting energy
American Journal of Clinical Nutrition, vol. 51,
Nature Communications, vol. 17, art.
no. 573, 2026, doi: 10.1038/s41467-025-67268-5.
[2] M. D. Mifflin, S. T. St Jeor, L. A. Hill, B. J. Scott, S. A. Daugherty, and Y. O. Koh, " A new predictive equation for expenditure in healthy individuals,"
no. 2, pp.
241- 247, 1990, doi: 10.1093/ajen/51.2.241.
[3] R. W. Morton et al, "A
meta-analysis and meta-regression of the effect of protein
systematic
review,
supplementation on resistance training-induced gains in muscle mass and strength,"
Medicine, vol.
British Journal of Sports
no. 6, pp. 376- 384, 2018, doi: 10.1136/bjsports-2017-097608.
[4] World Health Organization, " Healthy diet," WHO, updated 2025- 2026.
[5] World Health Organization, " Information note about intake of sugars recommended in the WHO guideline for adults and children," WHO, 2015.
[6] P. Lewis et al., " Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Advances in Neural Information Processing Systems, vol. 33, 2020.
[7] K. V. Dalakleidi, M. Papadelli, |. Kapolos, and K. Papadimitriou, " Applying Image-Based Food-Recognition Systems on Dietary Assessment: A Systematic Review," Advances in Nutrition, vol. 13, no. 6,pp. 2590- 2619, 2022, doi: 10.1093/advances/nmac078.
[8] S. Chang et al., " Validity of image-based dietary assessment methods: A_ systematic review and meta-analysis," Clinical Nutrition, 2020, doi: 10.1016/j.clnu.2020.08.002.
[9] L. Zhang, A. Misir, H. Boshuizen, and M. Ocké, " A Systematic Review and Meta-Analysis of Validation Studies Performed on Dietary Record Apps," Advances _ in Nutrition, vol. 12, no. 6, pp. 2321- 2332, 2021, doi: 10.1093/advances/nmab0O58.
"
[10] R. Paramastri et al., " Use of mobile applications to improve nutrition behaviour: A systematic review," Computer Methods and Programs in Biomedicine, vol. 192, art. no. 105459, 2020, doi: 10.1016/j.cmpb.2020.105459.
[11] D. Pala et al.," Smartphone applications for nutrition support: A systematic review of the target outcomes and main functionalities," International Journal of Medical Informatics, vol. 184, art. no. 105351, 2024, doi: 10.1016/j.ijmedinf.2024.105351.
[12] K. Villinger et al., " The effectiveness of app-based mobile interventions on _ nutrition behaviours and nutrition-related health outcomes: A systematic review and meta-analysis," Obesity Reviews, vol. 20, no. 10, pp. 1465- 1484, 2019, doi: 10.1111/obr.12903.
[13] H. J. Teede et al.," Recommendations from the 2023 International Evidence-Based Guideline for the Assessment and Management of Polycystic Ovary Syndrome," European Journal of Endocrinology, vol. 189, no. 2, pp. G43- G64, 2023, doi: 10.1093/ejendo/Ilvad096.
[14] A. D. Silva et al., " Time-restricted eating increases hunger in adults with overweight and obesity: A systematic review and meta-analysis of randomized controlled studies, "Nutrition Research, vol. 138, pp. 76- 88, 2025, doi: 10.1016/j.nutres.2025.04.003.
[15] D. Qi, X. Nie, and J. Zhang," A Systematic Review and Meta-analysis of the Impacts of Time-Restricted Eating on Metabolic Homeostasis," Angiology, vol. 76, no. 6, pp. 511- 524, 2025, doi: 10.1177/00033197241228046.
[16] S. Bedi et al.," Testing and Evaluation of Health Care Applications of Large Language Models: A Systematic Review," JAMA, published online Oct. 15, 2024, doi:

ISSN: 3139-0862

10.1001/jama.2024.21700.
[17] S. Basu and B. Huynh, “ Mitigating hallucinations in healthcare Al: a systematic review of evidence-based strategies,” BMC Health Services Research, vol. 26, art. no. 1115, 2026, doi: 10.1186/s12913-026-14851-1.
[18] U.S. Department of Health and Human Services, “ Covered Entities and Business Associates,” HHS HIPAA guidance, accessed 2026.
[19] European Union,“ Regulation (EU) 2016/679 (General Data Protection Regulation), Article 9," EUR-Lex, 2016.
[20] U.S. Department of Agriculture, Agricultural Research Service, “ FoodData Central: Downloadable Data and Documentation,” 2026.

AUTHOR AND RESEARCH-INTEGRITY STATEMENT

listed as students of Malnad College of Engineering, Hassan, Karnataka, India. The corresponding author for this manuscript is Kubera B (kkubera307@gmail.com; 8660770443).

Al-assisted coding and writing tools may be used during implementation and manuscript preparation. The authors remain responsible for architecture, source verification, experimental design, numerical results, interpretation, citations, and final submission. Prototype-reported benchmark values’ in this manuscript must be independently reproduced before being described as validated scientific performance.