

**БАКАЛАВРСЬКА РОБОТА**

**БР. Шз - 08.00.00.000 Шз**

**Група Шз-21-1**

**Терешко Яна**

**2025**

**Івано-Франківський національний технічний університет нафти і газу**

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

**Терешко Яна Русланівна**

(прізвище, ім'я, по батькові)

УДК 004.4  
(індекс)

## **БАКАЛАВРСЬКА РОБОТА**

**Програмний сервіс для розрахунку харчової цінності їжі**

(назва роботи)

**Інженерія програмного забезпечення**

(назва освітньої програми)

**121 - Інженерія програмного забезпечення**

(шифр і назва спеціальності)

**Робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело**

Здобувач освітнього рівня Терешко Я.Р.  
(підпис, ініціали та прізвище здобувача)

Науковий керівник Бандура Вікторія Валеріївна, к.т.н., доцент  
(підпис, прізвище, ім'я, по батькові, науковий ступінь, вчене звання керівника)

Допущено до захисту  
Завідувач кафедри

доц. Бандура В.В.  
(посада) (підпис) (дата) (ініціали та прізвище)

**Івано-Франківськ – 2025**



## 6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 28 квітня 2025 р.

Керівник \_\_\_\_\_

(підпис)

Завдання прийняв до виконання \_\_\_\_\_

(підпис)

## КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту	Примітка
1	Аналіз предметної області розробки програмних сервісів розрахунку харчової цінності їжі	11.05.2025	виконано
2	Особливості розробки сервісу на основі моделей глибокого навчання	20.05.2025	виконано
3	Моделі та методи глибокого навчання для додатку розрахунку харчової цінності їжі	28.05.2025	виконано
4	Реалізація сервісу для розрахунку харчової цінності їжі на основі розпізнавання зображень	03.06.2025	виконано
5	Модуль взаємодії з користувачем та кількісна оцінка порцій	16.06.2025	виконано
6	Оформлення пояснювальної записки дипломної роботи завідувачем кафедри	11.06.2025	виконано

Студент – дипломник \_\_\_\_\_

(підпис)

Керівник роботи \_\_\_\_\_

(підпис)

## АНОТАЦІЯ

Бакалаврська робота містить 83 сторінки, 31 рисунок, список використаних джерел із 31 найменуваннями.

**Мета роботи** - розробити інтелектуальний програмний сервіс для автоматичного розпізнавання їжі з фотографій та розрахунку її харчової цінності на основі методів глибокого навчання.

**Об'єкт дослідження** - процеси автоматизованого розпізнавання харчових продуктів і розрахунку їх нутрієнтної цінності.

**Предмет дослідження** - методи та моделі глибокого навчання, які застосовуються для візуального розпізнавання їжі, її інгредієнтів і оцінки порцій у мобільних та веб-додатках.

**В першому розділі** узагальнено ключові вимоги до сучасних систем оцінки харчування та виконано порівняльний аналіз провідних рішень, що дозволило сформулювати вимоги до власного сервісу

**В другому розділі** описано сучасні методи глибокого навчання для розпізнавання інгредієнтів та оцінки порцій, а також представлено релевантні датасети для тренування моделей.

**В третьому розділі** реалізовано архітектуру програмного сервісу з ключовими модулями, проведено тестування та підтверджено ефективність застосування обраних моделей на практиці.

**Висновок:** запропоновано архітектуру сервісу, що включає модулі розпізнавання інгредієнтів за допомогою візуально-мовних моделей, оцінки кількості їжі, розрахунку нутрієнтів та взаємодії з користувачем

**КЛЮЧОВІ СЛОВА:** РОЗПІЗНАВАННЯ ЇЖІ, ХАРЧОВА ЦІННІСТЬ, ГЛИБОКЕ НАВЧАННЯ, ВІЗУАЛЬНО-МОВНІ МОДЕЛІ, ОЦІНКА ПОРЦІЙ, АВТОМАТИЗАЦІЯ ХАРЧОВОГО МОНІТОРИНГУ, НЕЙРОННІ МЕРЕЖІ, АНАЛІЗ ЗОБРАЖЕНЬ.

## ANNOTATION

The bachelor's thesis contains 83 pages, 31 figures, a list of used sources with 31 names.

**The purpose of the work** is to develop an intelligent software service for automatic recognition of food from photographs and calculation of its nutritional value based on deep learning methods.

**The object of the study** is the processes of automated recognition of food products and calculation of their nutritional value.

**The subject of the study** is methods and models of deep learning that are used for visual recognition of food, its ingredients and portion estimation in mobile and web applications.

**The first section** summarizes the key requirements for modern nutrition assessment systems and performs a comparative analysis of leading solutions, which allowed us to form the requirements for our own service.

**The second section** describes modern deep learning methods for ingredient recognition and portion estimation, and also presents relevant datasets for training models.

**In the third section**, the architecture of the software service with key modules is implemented, testing is carried out and the effectiveness of the selected models in practice is confirmed.

**Conclusion:** a service architecture is proposed that includes modules for recognizing ingredients using visual-linguistic models, estimating food quantities, calculating nutrients, and interacting with the user.

**KEYWORDS:** FOOD RECOGNITION, NUTRITIONAL VALUE, DEEP LEARNING, VISION-LANGUAGE MODELS, PORTION ESTIMATION, AUTOMATED DIETARY MONITORING, NEURAL NETWORKS, IMAGE ANALYSIS.

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ .....	8
ВСТУП.....	9
<b>РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ РОЗРОБКИ ПРОГРАМНИХ СЕРВІСІВ РОЗРАХУНКУ ХАРЧОВОЇ ЦІННОСТІ ЇЖІ .....</b>	<b>12</b>
1.1. Актуальність розробки сервісу аналізу харчових продуктів .....	12
1.2. Особливості розробки сервісу на основі моделей глибокого навчання	14
1.3. Аналіз мобільних додатків розпізнавання їжі та автоматизованої оцінки харчування.....	16
1.3.1. Додаток MyFitnessPal .....	17
1.3.2. Додаток LoseIt.....	18
1.3.3. Додаток MyDietCam .....	21
1.3.4. Мобільний додаток FoodTracker .....	22
1.3.5. Додаток "DietLens" .....	24
1.3.6. Мобільний додаток Cronometer .....	26
Висновки до розділу .....	29
<b>РОЗДІЛ 2. МОДЕЛІ ТА МЕТОДИ ГЛИБОКОГО НАВЧАННЯ ДЛЯ РЕАЛІЗАЦІЇ ДОДАТКУ РОЗРАХУНКУ ХАРЧОВОЇ ЦІННОСТІ ЇЖІ.....</b>	<b>30</b>
2.1. Опис процесу розпізнавання інгредієнтів на основі глибокого навчання .....	30
2.2. Представлення методів автоматичної оцінки порцій.....	40
2.2.1. Методи оцінки за одним зображенням .....	41
2.2.2. Методи оцінки на основі багатьох зображень та відео.....	45
2.3. Набори даних, використані для розпізнавання їжі.....	48

					БР.ШЗ – 08.00.00.000 ПЗ				
Змн.	Арк.	№ докум.	Підпис	Дата	Програмний сервіс для розрахунку харчової цінності їжі  Пояснювальна записка	Літ.	Арк.	Акрушіє	
Розроб.		Терешко Я.Р.						6	
Перевір.		Бандура В.В.							
Реценз.									
Н. Контр.		Піх М.М.							
Затверд.		Бандура В.В.						ІФНТУНГ ШЗ-21-1	

Висновки до розділу .....	50
<b>РОЗДІЛ 3. РЕАЛІЗАЦІЯ СЕРВІСУ ДЛЯ РОЗРАХУНКУ ХАРЧОВОЇ ЦІННОСТІ ЇЖИ НА ОСНОВІ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ.....</b>	<b>52</b>
3.1. Огляд архітектури сервісу .....	52
3.2 Модуль розпізнавання інгредієнтів на основі візуально-мовних моделей.....	54
3.3. Модуль взаємодії з користувачем та кількісна оцінка порцій.....	57
3.4. Модуль розрахунку нутрієнтної цінності.....	60
3.4. Налаштування та опис сервісу .....	63
3.5. Опис наборів даних для оцінки ефективності сервісу .....	65
3.6. Експериментальне дослідження сервісу .....	67
3.7. Оцінка результатів .....	70
Висновки до розділу .....	76
<b>ВИСНОВКИ .....</b>	<b>78</b>
<b>ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....</b>	<b>80</b>
<b>БІБЛІОГРАФІЧНА ДОВІДКА</b>	

## ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

MAPE - Mean Absolute Percentage Error

SAM - Segment Anything Model

SGN - Structure-aware Generation Network

MSFM - Multi-Scale Fusion Module

FESM - Feature Enhancement and Suppression Module

Mrgcn - multi-relational graph convolutional network

DCNN - Convolutional Neural Networks

DCNN - Deep Convolutional Neural Networks

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						8
Змн.	Арк.	№ докум.	Підпис	Дата		

## ВСТУП

Дослідження діагностики харчування показали, що проблеми з діагностикою, такі як ожиріння, пов'язані з іншими хронічними захворюваннями, включаючи гіпертонію, нерегулярний рівень цукру в крові та підвищений ризик серцевих нападів. Основною причиною цих проблем є поганий спосіб життя та нездорові харчові звички, які можна контролювати за допомогою інтерактивних мобільних додатків для охорони здоров'я (mHealth). Однак традиційні системи моніторингу діагностики з використанням ручного ведення журналу харчування страждають від неточності, неповного звітування, витрат часу та низької дотриманності. Сучасні системи моніторингу діагностики вирішують ці проблеми шляхом автоматичної оцінки споживання харчових продуктів за допомогою методів машинного навчання.

### **Актуальність роботи**

У сучасному світі, де дедалі більше людей прагнуть до здорового способу життя, питання контролю раціону та споживаних нутрієнтів набуває особливої важливості. Водночас зростає попит на цифрові рішення, які дозволяють автоматизувати процес аналізу харчування та розрахунку харчової цінності. Традиційні методи потребують ручного введення даних і не забезпечують високої точності та зручності. Тому актуальним є створення інтелектуального сервісу, здатного розпізнавати їжу за зображенням і здійснювати точні розрахунки нутрієнтів, використовуючи сучасні досягнення в галузі глибокого навчання, комп'ютерного зору та обробки природної мови. Це дозволяє не лише підвищити якість контролю харчування, а й запропонувати користувачам нові можливості для самообслуговування та моніторингу здоров'я.

У сучасних умовах зростання інтересу до здорового способу життя, контролю харчування та профілактики захворювань, пов'язаних із

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						9
Змн.	Арк.	№ докум.	Підпис	Дата		

харчуванням, розробка інтелектуальних сервісів для аналізу їжі набуває особливої актуальності. Автоматизація процесів розпізнавання продуктів харчування, оцінки розміру порцій та розрахунку нутрієнтної цінності на основі глибокого навчання відкриває нові можливості для створення ефективних мобільних і веб-сервісів. У даній роботі розглянуто підходи до розробки такого сервісу, проаналізовано існуючі рішення, обґрунтовано вибір моделей і методів глибокого навчання, а також проведено реалізацію і тестування створеного сервісу.

Проблема точного та зручного розрахунку харчової цінності їжі є важливою для широкого кола користувачів: людей, які слідкують за дієтою, спортсменів, пацієнтів із харчовими обмеженнями. Існуючі додатки мають низку обмежень, пов'язаних із точністю розпізнавання продуктів, зручністю користування та адаптацією під локальні ринки продуктів. Використання технологій глибокого навчання, зокрема візуально-мовних моделей та багатокадрового аналізу, дозволяє суттєво покращити якість розпізнавання і точність оцінки харчової цінності. Актуальність даної роботи обумовлена потребою у розробці високоточних, доступних і користувацьки орієнтованих рішень для автоматизованого аналізу харчування.

**Мета роботи** - розробити інтелектуальний програмний сервіс для автоматичного розпізнавання їжі з фотографій та розрахунку її харчової цінності на основі методів глибокого навчання.

#### **Завдання дослідження**

1. Провести аналіз предметної області розрахунку харчової цінності та існуючих програмних рішень.
2. Дослідити сучасні методи глибокого навчання для розпізнавання зображень їжі та оцінки порцій.
3. Обґрунтувати вибір архітектури сервісу та методів реалізації.
4. Реалізувати модулі системи: розпізнавання інгредієнтів, оцінка порцій, обчислення нутрієнтів.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						10
Змн.	Арк.	№ докум.	Підпис	Дата		

5. Здійснити експериментальну перевірку ефективності роботи сервісу.

**Об'єкт дослідження** - процеси автоматизованого розпізнавання харчових продуктів і розрахунку їх нутрієнтної цінності.

**Предмет дослідження** - методи та моделі глибокого навчання, які застосовуються для візуального розпізнавання їжі, її інгредієнтів і оцінки порцій у мобільних та веб-додатках.

#### **Методи дослідження**

- Аналіз літературних джерел і огляд програмного забезпечення.
- Метод комп'ютерного зору для розпізнавання зображень.
- Глибокі нейронні мережі (CNN, ViT, multimodal models).
- Методика експериментального тестування та порівняльного аналізу.

#### **Наукова новизна**

Запропоновано комплексний підхід до розрахунку харчової цінності їжі, що поєднує глибоке візуально-мовне розпізнавання інгредієнтів з автоматичною оцінкою об'єму порцій на основі зображення, що забезпечує точність і зручність оцінки без ручного введення даних.

#### **Практичне застосування**

Результати дослідження можуть бути впроваджені в мобільні та веб-додатки для дієтологів, фітнес-тренерів, медичних працівників та звичайних користувачів, які прагнуть слідкувати за своїм харчуванням, а також як основа для подальшої розробки систем медичного моніторингу.

Бакалаврська робота містить 83 сторінки, 31 рисунок, 3 розділи список використаних джерел із 31 найменуванням.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		11

# РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ РОЗРОБКИ ПРОГРАМНИХ СЕРВІСІВ РОЗРАХУНКУ ХАРЧОВОЇ ЦІННОСТІ ЇЖІ

## 1.1. Актуальність розробки сервісу аналізу харчових продуктів

Харчування є фундаментальним аспектом здоров'я та добробуту людини, відіграючи вирішальну роль у профілактиці хронічних захворювань та підтримці оптимальної якості життя. Зважаючи на це, точний та регулярний аналіз споживання нутрієнтів є критично важливим для формування та дотримання здорового способу життя, а також для проведення дієтологічних досліджень. Термін "нутрієнти", або поживні речовини, використовується для позначення хімічних сполук, що містяться в їжі та напоях і є необхідними для забезпечення життєдіяльності, росту, розвитку та нормального функціонування організму людини. Традиційні методи збору даних про дієту, такі як харчові щоденники або опитувальники частоти споживання їжі, є трудомісткими, схильними до помилок пригадування та можуть створювати значне навантаження на респондентів.

У відповідь на ці виклики, розроблено численні технологічні рішення, зокрема мобільні додатки, спрямовані на оптимізацію процесу оцінки споживання харчових речовин. Однак, досягнення високої точності в автоматизованій оцінці дієти значною мірою залежить від ефективного вирішення двох ключових технічних завдань: ідентифікації харчових інгредієнтів у складних стравах та кількісної оцінки порцій спожитої їжі. Для того, щоб мобільний додаток міг надати комплексне рішення для повністю автоматичної оцінки харчування, він повинен інтегрувати ефективні механізми для обох цих компонентів.

У даній роботі представлено розробку та оцінку напівавтоматизованого сервісу для аналізу харчових продуктів, основним фокусом якого є автоматичне розпізнавання харчових інгредієнтів зі зображень. На відміну

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		12

від повністю автоматичних систем, що прагнуть оцінити як інгредієнти, так і порції, запропонований сервіс покладається на автоматичне розпізнавання інгредієнтів та ручне введення користувачем інформації про розмір порції. Цей підхід дозволяє зосередити технічні зусилля на найбільш складному етапі автоматизації – ідентифікації компонентів страви.

Запропонований сервіс інтегрує передові моделі глибокого навчання. Зокрема, використовується синергія візуально-мовних моделей, таких як VLP-2, для аналізу зображення страви та генерації відповідного текстового опису або вилучення візуальних ознак, та великих мовних моделей, таких як GPT-3.5. Результати демонструють, що комбіноване використання VLP-2 та GPT-3.5 дозволяє досягти високої точності в розпізнаванні різноманітних харчових інгредієнтів навіть у візуально складних композиціях страв, забезпечуючи значну універсальність системи щодо різних кухонь.

Для розрахунку поживної цінності страв використовується модель GPT-3.5, яка була попередньо налаштована або дотренована на структурованому наборі даних про склад продуктів, наприклад, наборі даних міністерства сільського господарства або аналогічних національних нутрієнтних базах. Ця модель обчислює вміст нутрієнтів на основі розпізнаних інгредієнтів та інформації про об'єм або вагу порцій, яка надається користувачем вручну.

Зважаючи на складність задачі та з метою детального аналізу, основний акцент даної роботи зосереджено саме на компоненті автоматичного розпізнавання інгредієнтів, як найбільш критичному та інноваційному етапі сервісу. Проведена порівняльна оцінка ефективності розробленого напівавтоматичного сервісу відносно існуючих мобільних додатків для оцінки дієти, які часто мають обмежені можливості автоматичного розпізнавання або покладаються виключно на пошук у базах даних. Оцінювалися такі метрики, як точність ідентифікації інгредієнтів, час

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		13

обробки зображення та потенціал для масштабування на широкий спектр страв.

Отримані результати свідчать про високий потенціал запропонованого напівавтоматичного сервісу у підвищенні точності та ефективності процесу моніторингу споживання нутрієнтів користувачами. Зокрема, продемонстровано його значну універсальність для ідентифікації інгредієнтів у стравах різних кухонь світу, що є суттєвою перевагою порівняно з системами, навченими лише на обмежених наборах даних. Це сприяє більш точному веденню харчових щоденників, роблячи їх більш доступними та зручними порівняно з ручними методами або менш досконалими автоматизованими системами, і є важливим кроком на шляху до створення повноцінних автоматичних систем оцінки дієти. Подальші дослідження включатимуть інтеграцію компонентів автоматичної оцінки порцій для створення повністю автоматизованого рішення.

## **1.2. Особливості розробки сервісу на основі моделей глибокого навчання**

У даній роботі представлено розробку мобільного додатка, призначеного для ефективного вирішення завдання автоматичного розпізнавання інгредієнтів. Як методологічну основу для нашого дослідження було проведено всебічний систематичний огляд літературних джерел. Метою цього огляду було узагальнити стан справ у галузі, ідентифікувати передові підходи, що застосовуються в існуючих мобільних додатках та дослідницьких системах для розпізнавання інгредієнтів та оцінки порцій їжі, починаючи від традиційних методів комп'ютерного зору до новітніх підходів на основі глибокого навчання.

Аналіз досліджень підтверджує, що основний фокус наукової спільноти та розробників рішень з автоматизованої оцінки дієти

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		14

зосереджений переважно на розпізнаванні інгредієнтів, тоді як питанню точної та автоматичної оцінки порцій приділяється значно менше уваги. Результати огляду дозволили ідентифікувати state-of-the-art підходи, проаналізувати їхню ефективність, виявити ключові переваги та обмеження різних методів, а також позиціонувати нашу реалізацію в контексті існуючих рішень.

Огляд технологічних основ розпізнавання інгредієнтів охоплює еволюцію методів від класичних підходів на основі ручного вилучення ознак до використання згорткових нейронних мереж (CNN) [1, 2], глибоких згорткових нейронних мереж (DCNN) [3, 4, 5], та, врешті, новітніх архітектур, таких як моделі трансформерів [6, 7], які продемонстрували високу ефективність у задачах візуального розпізнавання та обробки природної мови. Методи оцінки порцій були категоризовані за типом вхідних даних на методи, що використовують одне зображення, та методи, що вимагають багато зображень або відеопослідовність, з обговоренням відповідних стратегій та викликів. Складність обох завдань, розпізнавання інгредієнтів та оцінки порцій, стимулює безперервну розробку нових, більш досконалих підходів. Більш детальний аналіз кожного з розглянутих методів, разом з глибоким розумінням їхніх переваг та обмежень, представлено в даній роботі.

В роботі наведена архітектура сервісу розпізнавання інгредієнтів, детально описано запропонований сервіс для автоматичного розпізнавання харчових інгредієнтів. Висвітлюються обрані технології, включаючи моделі VLP-2 та GPT-3.5, надається обґрунтування їх вибору з огляду на завдання, а також представлено детальний опис архітектури та функціонування кожного компонента сервісу.

Проведено виклад експериментальної основи реалізації мобільного додатка. Деталізовано проведені експерименти, спрямовані на оптимізацію роботи сервісу. Проведено порівняльний аналіз ефективності розробленого

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		15

сервісу відносно референтних моделей, що використовуються в інших мобільних додатках або представлені в літературі. Обговорено важливість дотримання стандартизованих протоколів оцінки для коректного порівняння ефективності моделей різних архітектур на загальноприйнятих наборах даних.

Також надається інтерпретація результатів, що стосуються точності розпізнавання інгредієнтів, швидкості обробки та універсальності системи. Обговорюються виявлені обмеження та пропонуються потенційні шляхи вдосконалення архітектури та алгоритмів сервісу. Також визначаються перспективні напрямки для подальших досліджень та розробок, зокрема, інтеграція більш досконалих методів автоматичної оцінки порцій для створення повноцінної системи автоматизованого моніторингу дієти.

### **1.3. Аналіз мобільних додатків розпізнавання їжі та автоматизованої оцінки харчування**

У цьому підрозділі представлено аналіз існуючих описаних мобільних додатків, призначених для автоматизації розпізнавання об'єктів їжі та оцінки споживання нутрієнтів. З метою забезпечення прозорості та можливості відтворення результатів порівняльного аналізу, ми обмежуємо наш розгляд рішеннями, представленими в наукових публікаціях, де архітектура нейронних мереж, використані методології та набори даних для експериментальної оцінки є публічно доступними. Хоча також проаналізуємо такі пропріетарні комерційні додатки, такі як MyFitnessPal та LoseIt. Окрім аналізу конкретних додатків, цей розділ містить більш комплексний огляд фундаментальних підходів до розпізнавання харчових інгредієнтів та квантифікації порцій, починаючи від класичних методів комп'ютерного зору до сучасних підходів на основі глибокого навчання.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		16

### 1.3.1. Додаток MyFitnessPal

MyFitnessPal — це один із найпопулярніших мобільних додатків та веб-сайтів, призначених для відстеження харчування та фізичної активності. Він був запущений у 2005 році.

Головна мета MyFitnessPal — допомогти користувачам контролювати споживання калорій, макронутрієнтів (білків, жирів, вуглеводів) та мікронутрієнтів (вітамінів, мінералів) для досягнення своїх цілей щодо здоров'я та ваги (зниження, набір або підтримання ваги).

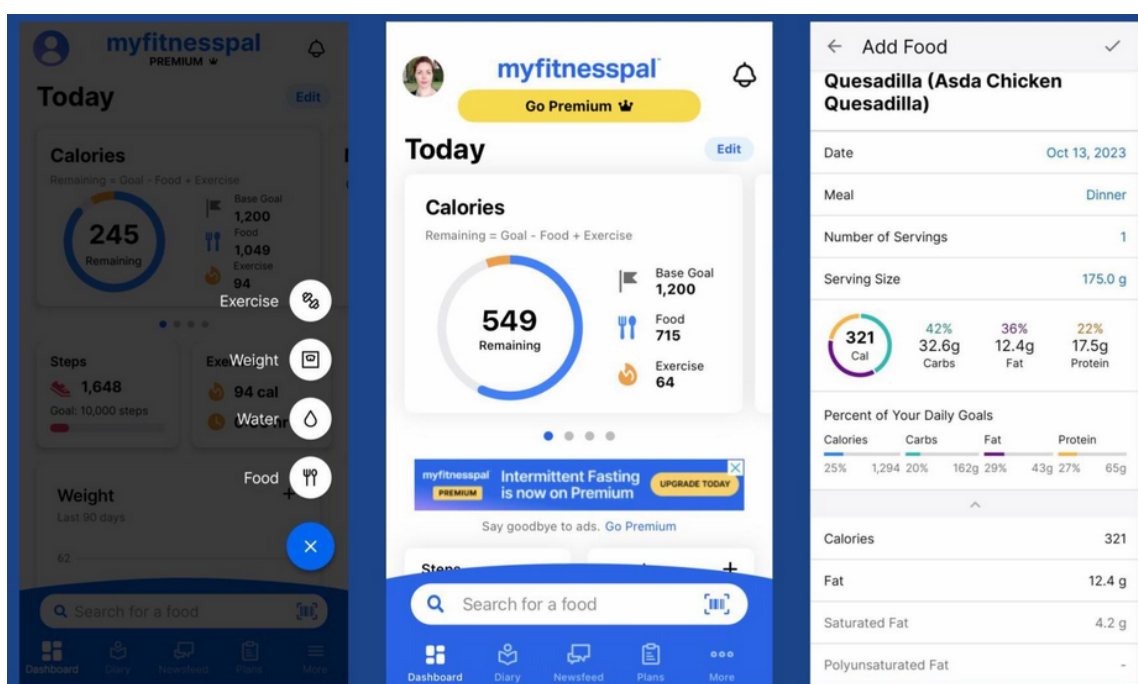


Рисунок 1.1 – Інтерфейс додатку MyFitnessPal

Наведемо основні функції та особливості MyFitnessPal:

- Додаток містить одну з найбільших у світі баз даних продуктів, що включає як загальні продукти, так і брендовані товари, а також страви з ресторанів. Ця база постійно оновлюється як самими користувачами, так і командою MyFitnessPal.
- Користувачі можуть швидко додавати упаковані продукти до свого щоденника харчування, просто відсканувавши штрих-код.

- Можливість пошуку продуктів у базі даних або ручного введення інформації про самостійно приготовані страви чи продукти без штрих-коду.

- Користувачі можуть реєструвати всі прийоми їжі протягом дня (сніданок, обід, вечеря, перекуси).

- Можливість зберігати часто вживані страви або створювати власні рецепти з автоматичним розрахунком їхньої поживної цінності.

- Додаток дозволяє реєструвати різні види фізичних вправ та розраховувати приблизну кількість спалених калорій. Він також може інтегруватися з іншими фітнес-трекерами та пристроями.

- Користувачі можуть встановлювати цілі щодо щоденного споживання калорій та бажаного розподілу макронутрієнтів.

- Надання звітів та графіків, що демонструють споживання нутрієнтів протягом дня, тижня чи місяця, а також прогрес у досягненні ваги.

- Можливість додавати друзів, ділитися прогресом та отримувати підтримку від спільноти користувачів через форуми.

MyFitnessPal працює за моделлю freemium, пропонуючи основний функціонал безкоштовно та розширені можливості через платну підписку (MyFitnessPal Premium). Завдяки своїй зручності, великій базі даних та широкому функціоналу, він є одним із лідерів на ринку додатків для відстеження здоров'я та фітнесу.

Важливо зазначити, що, на відміну від систем, які використовують автоматичне розпізнавання їжі за зображенням, MyFitnessPal переважно покладається на активну участь користувача у введенні даних шляхом пошуку в базі, сканування або ручного запису.

### 1.3.2. Додаток LoseIt

Lose It! — це ще один широко популярний мобільний додаток та онлайн-платформа, яка спеціалізується на допомозі користувачам у зниженні

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		18

ваги шляхом відстеження споживання калорій та фізичної активності. Додаток був запущений у 2008 році.

Основна філософія Lose It! полягає у встановленні індивідуального щоденного бюджету калорій для кожного користувача, розрахованого на основі його поточних даних (вага, вік, стать, рівень активності) та визначеної мети зниження ваги (темп зниження). Користувачі реєструють усю спожиту їжу та фізичну активність, щоб залишатися в межах встановленого бюджету.



Рисунок 1.2 – Інтерфейс додатку Lose It!

Розглянемо функції та особливості додатку Lose It!:

1. Містить велику базу даних харчових продуктів, страв та напоїв, хоча іноді вважається дещо меншою за базу MyFitnessPal, але все ще достатньо обширною.

2. Дозволяє швидко додавати упаковані продукти з використанням сканеру штрих-кодів.

3. Можливість шукати продукти в базі даних або вручну вводити інформацію про власні страви.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		19

4. Функція "Snap It": Це одна з відмінних особливостей Lose It!. Вона дозволяє користувачам зробити фотографію їжі, і додаток, використовуючи технології розпізнавання зображень, намагається ідентифікувати страву або продукт та запропонувати варіанти з бази даних для вибору та реєстрації. Важливо розуміти, що це не є повністю автоматичним процесом; система надає припущення, які користувач має підтвердити або відкоригувати, і функціональність обмежена розпізнаванням страв/продуктів, а не інгредієнтів.

5. Можливість зберігати улюблені страви та рецепти для швидкого доступу.

6. Інтеграція з багатьма популярними фітнес-трекерами та додатками, а також можливість ручного введення даних про тренування для коригування щоденного бюджету калорій.

7. Моніторинг зміни ваги, вимірів тіла та інших показників.

8. Можливість брати участь у групових викликах та взаємодіяти з іншими користувачами.

9. У платній версії (Lose It! Premium) надається більш детальний аналіз споживання макро- та мікронутрієнтів.

Подібно до MyFitnessPal, Lose It! є пропріетарним комерційним додатком. Хоча він включає елементи розпізнавання зображень через функцію "Snap It", його основним методом ведення щоденника залишається пошук у базі даних та ручне введення/підтвердження. Функція "Snap It" зазвичай працює на рівні розпізнавання цілих страв або продуктів, а не на рівні детального розпізнавання інгредієнтів, що є фокусом більш складних дослідницьких систем. Тому через свою закритість та орієнтацію на дещо інші методи введення даних, він, як правило, не включається до наукових оглядів, що аналізують специфічні автоматизовані підходи на основі комп'ютерного зору з відкритими архітектурами.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		20

### 1.3.3. Додаток MyDietCam

MyDietCam — це назва дослідницького прототипу мобільного додатку для автоматизованого розпізнавання їжі. Важливо відрізнити його від широкодоступних комерційних додатків, таких як MyFitnessPal або Lose It!, оскільки MyDietCam є результатом академічних досліджень, спрямованих на вивчення конкретних технічних підходів.

Основна мета MyDietCam — здійснити автоматичне розпізнавання страв (ідентифікацію типу страви) за фотографією, зробленою користувачем. Це один із кроків у напрямку автоматизації процесу оцінки харчування.



Рисунок 1.3 – Вигляд додатку MyDietCam

З точки зору технології, додаток MyDietCam використовує сучасні підходи на основі глибокого навчання. Для вилучення ключових візуальних ознак із зображень страв (наприклад, текстур, форм, кольорів) застосовується архітектура глибокої нейронної мережі Inception-Resnet-V2. Автори дослідження визначили її як найбільш ефективну для цього завдання у своєму контексті. Отримані візуальні ознаки потім передаються до класифікатора — алгоритму ARCIKELM (Adaptive Reduced Class Incremental

Guided Extreme Learning Machine), який і виконує фінальну ідентифікацію страви, зіставляючи ознаки з класами страв, відомих системі.

Ефективність MyDietCam була оцінена на кількох стандартних академічних наборах даних, що широко використовуються спільнотою комп'ютерного зору для бенчмаркінгу систем розпізнавання їжі, включаючи Food101, UECFOOD100 та UECFOOD256. За повідомленнями авторів, додаток продемонстрував показники точності розпізнавання 87.27% на Food101, 88.74% на UECFOOD100 та 76.51% на UECFOOD256.

Ключовим аспектом MyDietCam, який відрізняє його від більш складних систем (включно з тими, що фокусуються на розпізнаванні інгредієнтів), є його орієнтація саме на розпізнавання страв (наприклад, визначення, що на фотографії "борщ" або "паста Карбонара"), а не на детальний аналіз та ідентифікацію окремих інгредієнтів у складі цієї страви (наприклад, визначення всіх компонентів борщу). Також, згідно з наявними описами, додаток переважно не включає функціонал автоматичної оцінки порцій на основі аналізу самого зображення.

Таким чином, MyDietCam є важливим прикладом ранніх досліджень, які продемонстрували потенціал використання глибокого навчання для автоматизації етапу ідентифікації страв у мобільних додатках для оцінки харчування, але він не вирішує більш складні завдання розпізнавання інгредієнтів та оцінки порцій без участі користувача.

#### *1.3.4. Мобільний додаток FoodTracker*

FoodTracker — це назва дослідницького прототипу мобільного додатку, і є ще одним прикладом академічної розробки в галузі автоматизованого аналізу харчування, що не є комерційним продуктом, доступним широкому загалу користувачів.

Основне завдання, на якому зосереджений FoodTracker, — це автоматичне детектування об'єктів їжі на фотографії. На відміну від простого

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		22

розпізнавання, детектування включає не лише ідентифікацію (класифікацію) типу їжі, але й визначення її місця розташування на зображенні, зазвичай шляхом обведення об'єктів обмежувальними рамками. Це робить його корисним для сценаріїв, де на одній тарілці присутні кілька різних продуктів або компонентів страви.

З технічної точки зору, FoodTracker використовує підходи глибокого навчання, оптимізовані для розгортання на мобільних пристроях з обмеженими обчислювальними ресурсами. Додаток базується на архітектурі глибокої згорткової нейронної мережі (DCNN) Mobilenet, яка відома своєю ефективністю та компактністю. Ця DCNN інтегрована з фреймворком YOLOv2 (You Only Look Once version 2), популярною одностадійною архітектурою для детектування об'єктів. Таке поєднання дозволяє системі одночасно ідентифікувати та локалізувати декілька об'єктів їжі на зображенні за один прохід мережі, що сприяє відносно високій швидкості обробки.

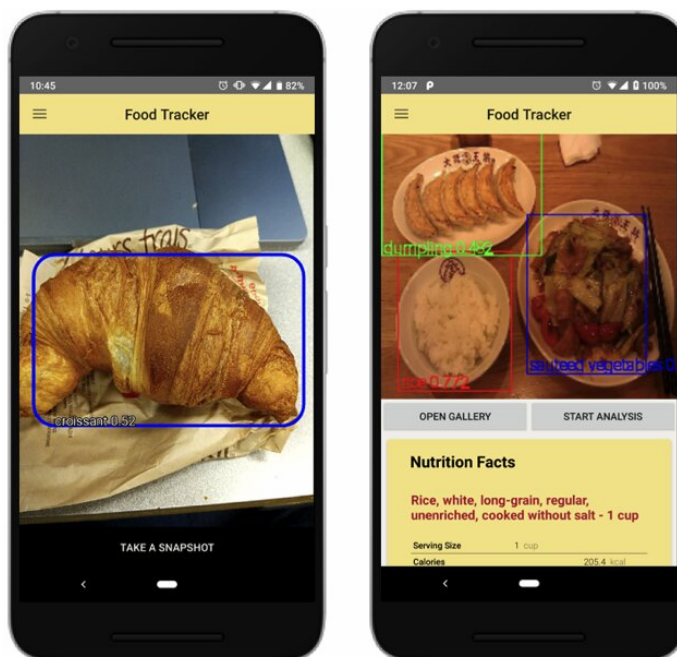


Рисунок 1.4 – Вигляд додатку FoodTracker

Ефективність FoodTracker була оцінена на стандартних наборах даних для розпізнавання та детектування їжі, зокрема UECFOOD100 та

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		23

UECFood256. Як метрика оцінки ефективності для завдання детектування використовувалася середня точність детектування (mAP - Mean Average Precision). За повідомленнями авторів [9], FoodTracker досягнув показників mAP 76.35% на наборі даних UECFOOD100 та 75.05% на UECFOOD256.

Слід зазначити, що, хоча FoodTracker здійснює локалізацію об'єктів, його основне завдання — це детектування та ідентифікація об'єктів їжі (окремих продуктів або страв), а не детальний аналіз та розпізнавання інгредієнтів, які можуть бути змішані у складній страві (наприклад, розпізнавання овочів, м'яса та круп у рагу). Також в описі цієї роботи не згадується функціонал автоматичної оцінки порцій на основі аналізу самого зображення.

Таким чином, FoodTracker є важливим прикладом дослідницької розробки, що демонструє застосування оптимізованих для мобільних пристроїв архітектур глибокого навчання (Mobilenet) та фреймворків детектування об'єктів (YOLOv2) для ефективною ідентифікації та локалізації об'єктів їжі на фотографіях.

### 1.3.5. Додаток "DietLens"

DietLens — це назва дослідницької системи/мобільного додатка, яка була представлена в науковій публікації Міна та співавторів [10]. Подібно до MyDietCam та FoodTracker, це наукова розробка, а не комерційний продукт, спрямований на дослідження можливостей автоматизації оцінки харчування за допомогою зображень.

Основне завдання DietLens полягає в автоматичному розпізнаванні страв за фотографією, зробленою користувачем. Метою є ідентифікація конкретного типу страви (наприклад, "піца Маргарита", "курка каррі") для подальшого використання цієї інформації в процесі обліку харчування.

З технічної точки зору, DietLens використовує підходи глибокого навчання для розпізнавання зображень страв. Зокрема, система базується на

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		24

добре відомій та широко використовуваній архітектурі згорткової нейронної мережі ResNet-50.

Важливий аспект DietLens, пов'язаний з оцінкою порцій: На відміну від систем, що прагнуть автоматично визначити кількість спожитої їжі на основі аналізу візуальних даних, DietLens не виконує автоматичну оцінку порцій із зображення. Замість цього, система покладається на попередньо визначені розміри порцій, які, ймовірно, зберігаються у пов'язаній базі даних і обираються користувачем вручну після розпізнавання страви. Це означає, що компонент кількісної оцінки в цьому додатку не є автоматизованим на основі зображень.

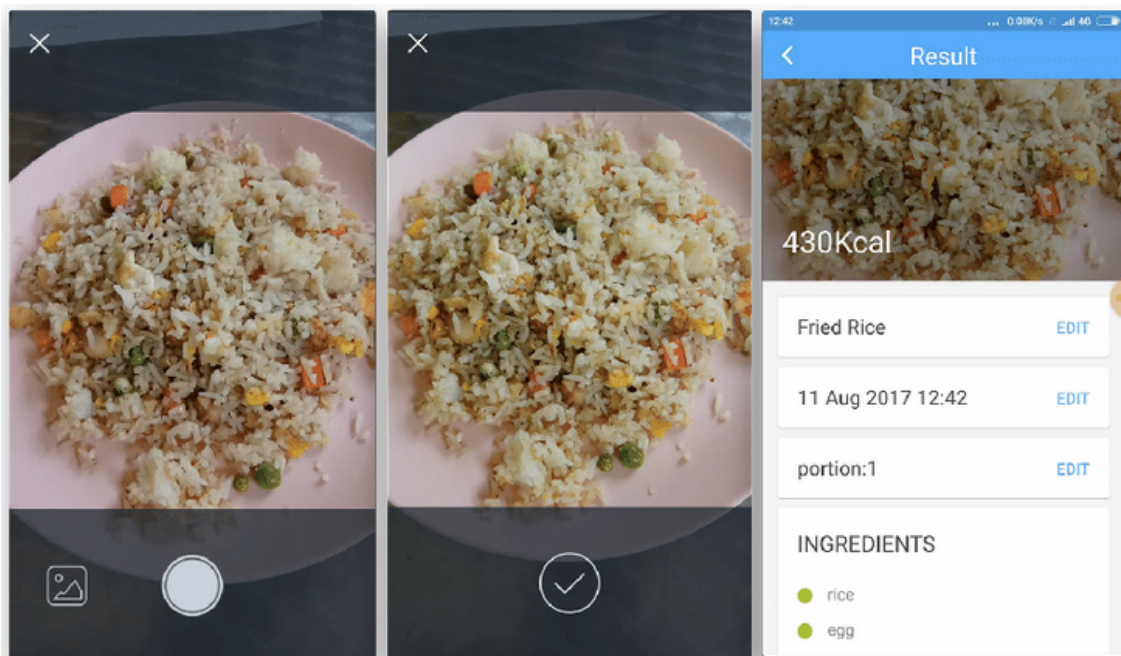


Рисунок 1.5 – Інтерфейс додатку DietLens

Ефективність DietLens була оцінена на різноманітному наборі даних, зібраному з Google, який охоплював страви різних кухонь світу, включаючи китайську, західну, індійську, малайзійську та інші. Крім того, система була протестована на більш складному та реалістичному наборі даних, зібраному від реальних користувачів (що складався з 4697 зображень страв).

На гетерогенному наборі даних з Google додаток показав показники точності Top-1 78.6% для західних страв та 64.2% для китайських страв. Це

ілюструє виклик, пов'язаний з адаптацією систем розпізнавання до візуальних відмінностей страв різних кухонь. Середня точність Top-1 для популярних страв на цьому наборі даних склала 75.2%. Однак, при тестуванні на наборі даних з реальних зображень користувачів, точність Top-1 моделі суттєво знизилася приблизно до 47.6%. Це підкреслює загальну проблему перенесення ефективності систем комп'ютерного зору з керованих лабораторних наборів даних на неконтрольовані умови реального світу.

Як і MyDietCam, DietLens орієнтований на розпізнавання страв (як цілих об'єктів), а не на розпізнавання інгредієнтів у їхньому складі. Його основним обмеженням як системи для повної автоматизації обліку харчування є відсутність можливості автоматичної оцінки порцій на основі зображення та значне зниження точності в умовах реального використання. Незважаючи на це, DietLens є цінним внеском у дослідження, оскільки він демонструє можливості глибокого навчання для розпізнавання страв у мультикультурному контексті та виявляє важливі проблеми, пов'язані з ефективністю на реальних даних.

### *1.3.6. Мобільний додаток Cronometer*

Cronometer — це мобільний додаток та веб-платформа, що використовується для відстеження харчування, фізичної активності та аналізу споживання нутрієнтів. Його ключова перевага та відмінність від багатьох інших подібних додатків, таких як MyFitnessPal або Lose It!, полягає у наданні надзвичайно деталізованої інформації про споживання широкого спектру мікронутрієнтів (вітамінів, мінералів, незамінних амінокислот, жирних кислот тощо), а не лише калорій та основних макронутрієнтів (білків, жирів, вуглеводів).

Основна мета Cronometer — допомогти користувачам отримати максимально точне та повне уявлення про якість свого раціону з точки зору всіх необхідних поживних речовин, ідентифікувати можливі дефіцити або

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		26

надлишки та оптимізувати своє харчування відповідно до індивідуальних потреб чи дієтичних цілей.

### Основні функції Cronometer:

- На відміну від баз даних, що значною мірою формуються користувачами (як у MyFitnessPal), база Cronometer переважно базується на верифікованих даних з надійних джерел, таких як лабораторні аналізи та державні нутрієнтні банки даних (наприклад, USDA). Це забезпечує вищу точність інформації про нутрієнтний склад продуктів.

- Додаток відстежує не тільки калорії, білки, жири та вуглеводи, але й повний спектр вітамінів (включаючи різні форми), мінералів, незамінних амінокислот, омега-3 та омега-6 жирних кислот, холестерин, цукор, клітковину тощо.

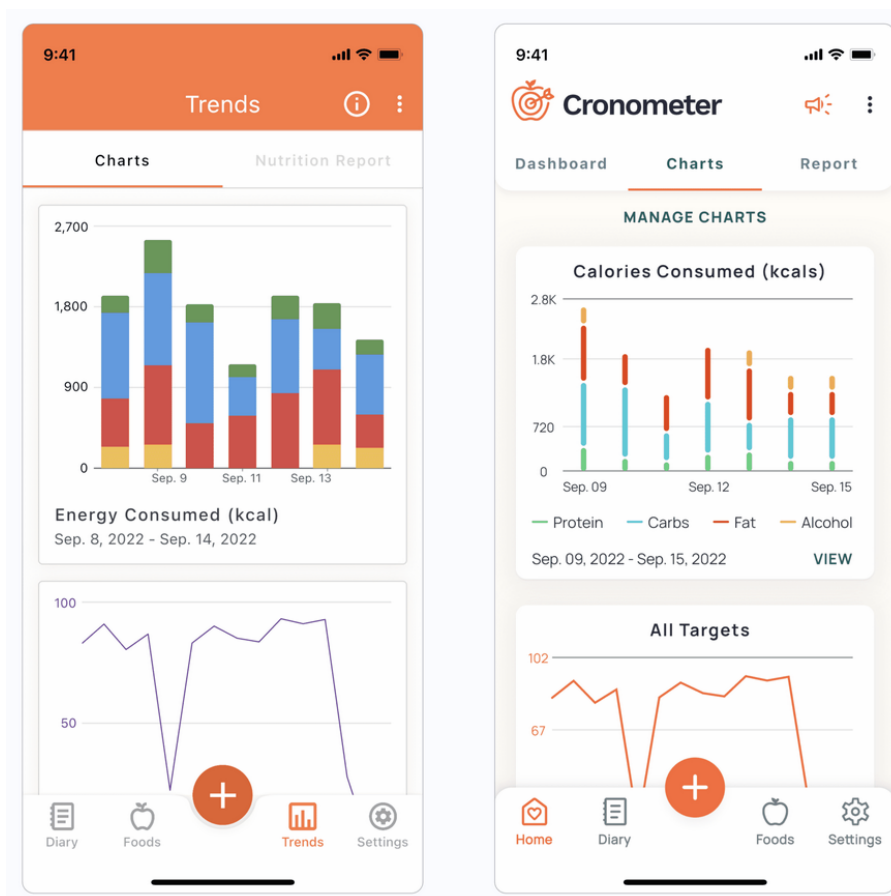


Рисунок 1.6 – Інтерфейс Cronometer

- Надає детальні візуальні звіти та графіки, що демонструють споживання кожного відстежуваного нутрієнта порівняно з рекомендованими добовими нормами (RDA або іншими стандартами, обраними користувачем).

- Підтримує ручне введення продуктів, пошук у базі даних та сканування штрих-кодів.

- Можливість легко розрахувати та зберегти нутрієнтний склад власних рецептів.

- Дозволяє реєструвати тренування та інтегрується з багатьма популярними фітнес-трекерами (наприклад, Apple Health, Google Fit, Garmin, Fitbit).

- Користувачі можуть встановлювати цілі не тільки щодо ваги та співвідношення макронутрієнтів, але й щодо споживання конкретних мікронутрієнтів.

- Зручний для людей, які дотримуються специфічних дієт (кетогенна, веганська, палео, низьковуглеводна тощо), завдяки деталізованому відстеженню всіх необхідних показників.

Cronometer працює за моделлю freemium, пропонуючи базовий функціонал безкоштовно та розширені можливості (як-от більш глибокі звіти, розширені налаштування макро- та мікроцілей) у платній підписці "Gold".

Цей додаток часто обирають користувачі, для яких точність та повнота даних про нутрієнтний склад мають першочергове значення – спортсмени, люди з певними станами здоров'я, що вимагають ретельного контролю дієти, дослідники, дієтологи та ті, хто прагне максимально оптимізувати своє харчування. Подібно до інших комерційних трекерів, його основним методом введення даних є ручний або напівавтоматичний пошук/сканування, а не автоматичне розпізнавання їжі за зображенням.

									Арк.
									28
Змн.	Арк.	№ докум.	Підпис	Дата	БР.ПЗ – 08.00.00.000 ПЗ				

Аналіз цих та інших робіт, деталізований у другому розділі, дозволяє зробити висновок, що, попри значний прогрес у розпізнаванні страв та детектуванні об'єктів їжі, завдання точного автоматичного розпізнавання інгредієнтів у складних стравах та ефективною оцінки порцій без прямої участі користувача залишаються суттєвими викликами. Більшість існуючих мобільних рішень або працюють на рівні страв, або використовують спрощені методи оцінки порцій. Наша робота спрямована на подолання цих обмежень, зокрема, зосереджуючись на підвищенні точності саме автоматичного розпізнавання інгредієнтів.

### **Висновки до розділу**

У розділі проведено всебічний аналіз предметної області розробки програмних сервісів для розрахунку харчової цінності їжі. Було обґрунтовано актуальність створення таких сервісів, зумовлену зростанням інтересу до здорового способу життя, необхідністю точного обліку споживаних нутрієнтів та потребою у зручних інструментах для моніторингу харчування.

Проведено огляд існуючих мобільних додатків для аналізу харчування: MyFitnessPal, LoseIt, MyDietCam, FoodTracker, DietLens та Cronometer. Виявлено їх сильні сторони, такі як велика база даних продуктів, функціональність автоматичного розпізнавання їжі та підтримка користувацьких цілей (зниження ваги, підтримання форми, медичні дієти). Разом з тим виявлено і типові обмеження: залежність точності розпізнавання від якості фотографій, обмеженість бази локальних продуктів, складність налаштувань для користувачів з індивідуальними потребами. Таким чином, аналіз показав доцільність розробки нового сервісу, який би поєднував найкращі практики існуючих рішень, враховував недоліки аналогів і забезпечував високу точність, адаптивність і зручність для кінцевого користувача за допомогою інноваційних технологій глибокого навчання

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		29

## РОЗДІЛ 2. МОДЕЛІ ТА МЕТОДИ ГЛИБОКОГО НАВЧАННЯ ДЛЯ РЕАЛІЗАЦІЇ ДОДАТКУ РОЗРАХУНКУ ХАРЧОВОЇ ЦІННОСТІ ЇЖІ

### 2.1. Опис процесу розпізнавання інгредієнтів на основі глибокого навчання

У розробці автоматичного додатка для обчислення харчових речовин на основі зображень важливо розуміти досягнення в розпізнаванні інгредієнтів. Тут ми розглядаємо різні підходи, зосереджені на розпізнаванні інгредієнтів, категоризуючи їх на згорткові нейронні мережі (CNN), глибокі згорткові нейронні мережі (DCNN), моделі трансформерів та передові архітектури, а також інноваційні техніки.

Одним з широко застосовуваних методів для розпізнавання інгредієнтів є багатоміткова класифікація. Цей метод передбачає прогнозування кількох вихідних міток або категорій, кожна з яких представляє різний інгредієнт [11].

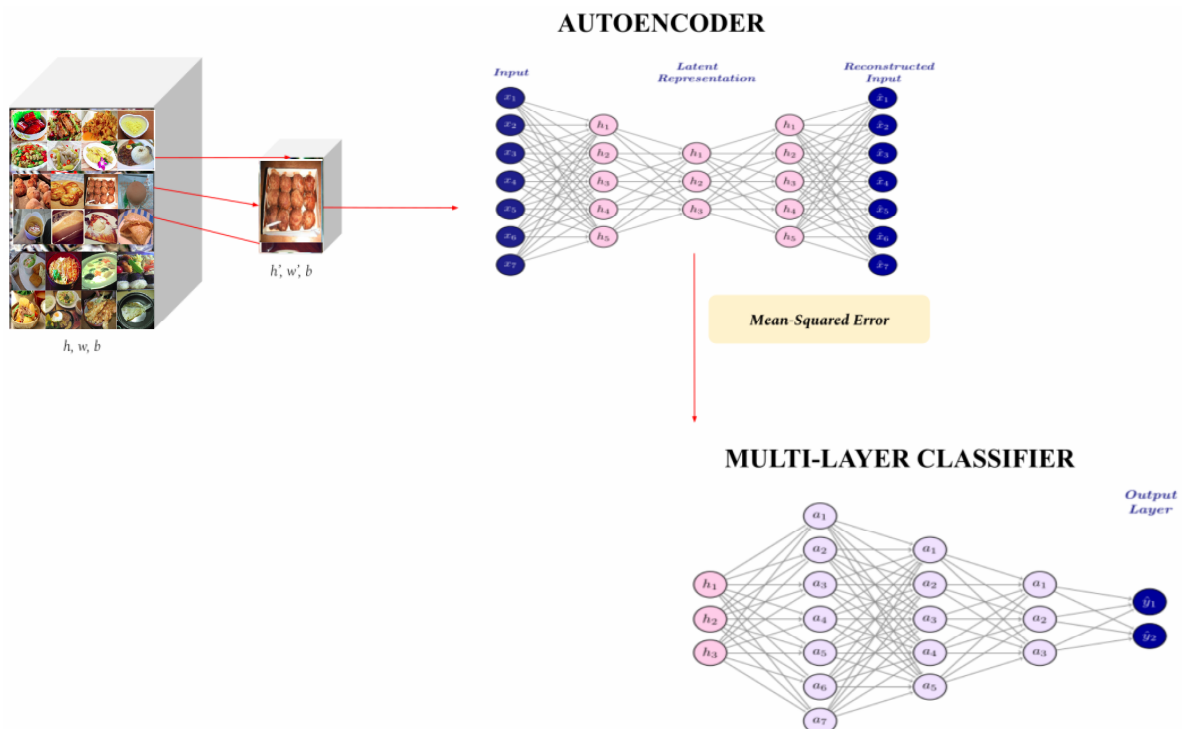


Рисунок 2.1 – Приклад архітектури моделі багатоміткової класифікації

Рисунок 2.1 ілюструє архітектуру моделі та маршрут потоку даних. Вона складається з двох основних компонентів: автокодувальника та класифікатора. Патчі (фрагменти) зі зменшеними просторовими та повними спектральними розмірностями вибираються з оригінальної сцени дистанційного зондування та передаються до мережі.

Зосереджуючись на згорткових нейронних мережах (CNN), в [1] представляється модель, адаптовану для багатоміткового прогнозування, спеціально розроблену для розпізнавання кількох харчових інгредієнтів з одного зображення. Цей підхід дозволяє моделі точно визначати список інгредієнтів у стравах, з якими вона раніше не стикалася, використовуючи два новостворені набори даних для покращення навчання та узагальнення. Подібним чином, в [2] застосовують попередньо навчену CNN, EfficientNetB2, для класифікації зображень їжі на наборі даних MedGRFood, підкреслюючи здатність CNN обробляти складні зображення їжі.

Запропонована система оцінки харчування потребує двох зображень їжі для розрахунку її нутрієнтного складу. Спочатку, за допомогою камери смартфона, зображення робляться з кутом зйомки близько 45 градусів від вертикальної осі їжі та на відстані близько 40 - 50 см. Референсна картка, розміщена поруч зі стравою, допомагає оцінити масштаб знятих зображень. Далі підсистема класифікації їжі, використовуючи методи глибокого навчання, розпізнає клас їжі. Одночасно підсистема оцінки об'єму їжі, використовуючи методи стереобачення, створює 3D реконструкцію їжі та оцінює її кількість (об'єм).

Для тренування та оцінки обох підсистем була створена відповідна база зображень їжі. База зображень складається з двох піднаборів даних:

- а) набір зображень, придатний для моделей класифікації на основі глибокого навчання,
- б) набір зображень, придатний для систем оцінки об'єму їжі на основі методів стереобачення.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						31
Змн.	Арк.	№ докум.	Підпис	Дата		

Нарешті, знаючи клас їжі та її об'єм, ми розраховуємо нутрієнтний склад знятої їжі, використовуючи базу даних USDA [28] та Грецький набір даних про склад продуктів.

Основні частини запропонованої системи оцінки харчування є:

- 1) набір зображень їжі та база даних про її поживну цінність,
- 2) підсистема класифікації,
- 3) підсистема оцінки об'єму, які показано на рис. 2.2.

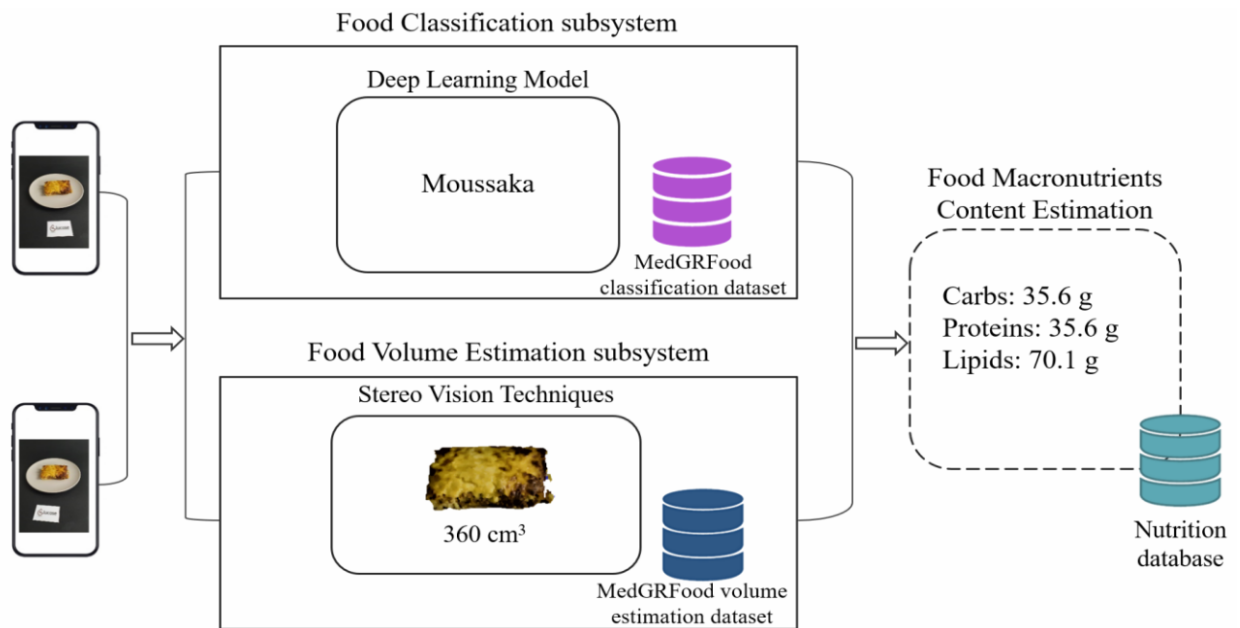


Рисунок 2.2 - Запропонована система оцінки харчування

Переходячи до глибоких згорткових нейронних мереж (DCNN), робота [3] пропонує новий глибокий навчальний фреймворк для пошуку рецептів кулінарії та розпізнавання інгредієнтів на зображеннях їжі. Цей підхід адаптує архітектури від VGG для одночасного навчання категоризації їжі та розпізнавання інгредієнтів, демонструючи покращення в обробці великих візуальних варіацій страв.

Рисунок 2.3 надає огляд фреймворку (або структури системи), який складається з двох модулів: розпізнавання інгредієнтів та zero-shot пошуку рецептів. Перший модуль формулює завдання розпізнавання інгредієнтів як

задачу багатозадачного навчання, використовуючи глибоку згорткову нейронну мережу (ГЗММ). Маючи зображення страви, модуль видає назву страви разом з гістограмою інгредієнтів. Розроблена ГЗММ може розпізнавати 172 категорії китайських страв та 353 інгредієнти. Наскільки нам відомо, досі не опубліковано результатів для розпізнавання інгредієнтів у такому великому масштабі.

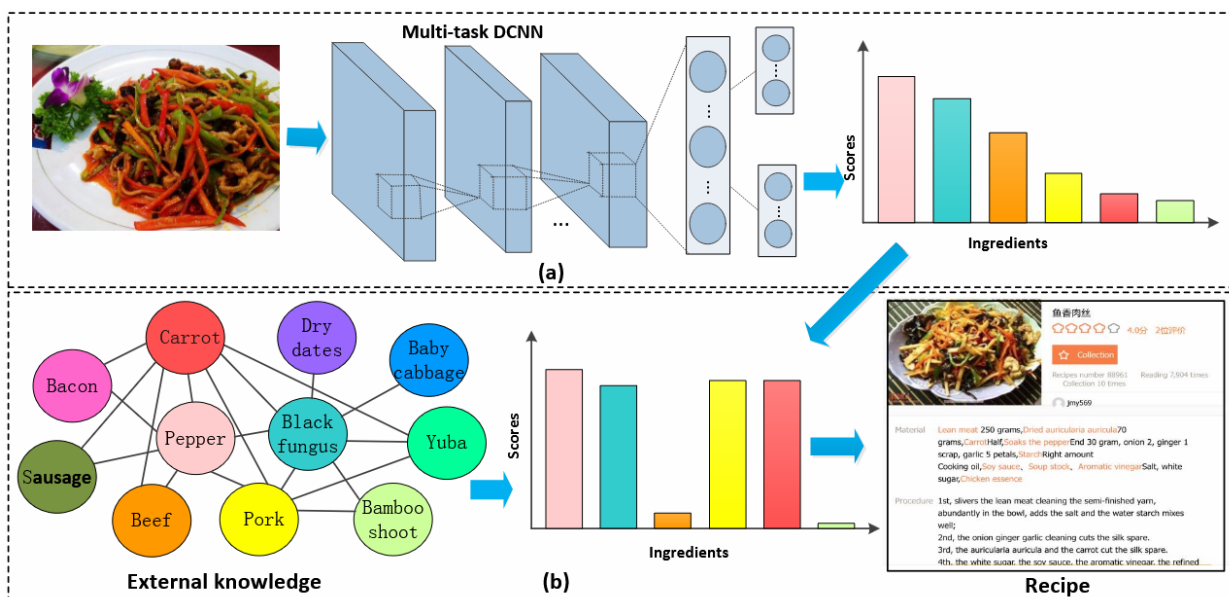


Рисунок 2.3 - Огляд фреймворку: (а) розпізнавання інгредієнтів, (б) zero-shot пошук рецептів.

Маючи зображення страви з невідомою категорією, фреймворк здійснює пошук рецепту для цієї страви. Рецепт оригінально китайською мовою та перекладений Google на англійську.

Другий модуль виконує zero-shot пошук рецептів шляхом зіставлення передбачених інгредієнтів з великим корпусом, що містить понад 60 000 рецептів. Корпус включає деякі категорії страв, а також інгредієнти, невідомі багатозадачній ГЗММ. Для підвищення ефективності пошуку з корпусу рецептів навчається граф, що кодує контекстний зв'язок між інгредієнтами. Використовуючи цей граф, застосовується умовне випадкове поле (УВП), щоб імовірно налаштувати розподіл ймовірностей інгредієнтів та

зменшити потенційну помилку розпізнавання, спричинену невідомою категорією страви.

В [4] представляють багатореляційну графову згорткову нейронну мережу (mRGCN) для розпізнавання інгредієнтів з нульовим навчанням. Цей новаторський підхід інтегрує кілька реляційних графів з DCNN для прогнозування класифікаторів для небачених інгредієнтів, показуючи хороші результати на китайських та японських наборах даних їжі, таких як UECFOOD-100 та VireoFood-172.

Рисунок 2.4 представляє огляд фреймворку, який на етапі тренування складається з двох модулів: багатоміткової глибокої згорткової нейронної мережі (ГЗММ) для навчання класифікатора відомих інгредієнтів та багатореляційної графової згорткової нейронної мережі (mRGCN) для передбачення класифікатора невідомих інгредієнтів.

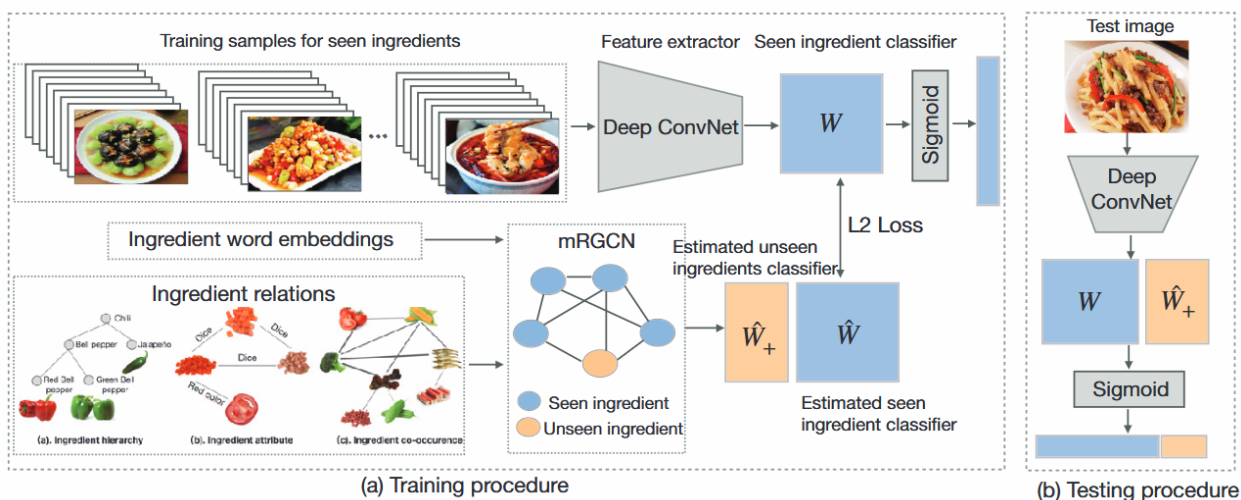


Рисунок 2.4 - Огляд фреймворку розпізнавання інгредієнтів з нульовим навчанням

Під час тренування запропонований фреймворк містить два основні модулі: багатоміткову глибоку згорткову нейронну мережу для розпізнавання відомих інгредієнтів та багатореляційну графову згорткову нейронну мережу (mRGCN) для передбачення класифікатора невідомих

інгредієнтів. Навчені класифікатори інгредієнтів, витягнуті з багатоміткової ЗММ (CNN), використовуються як "ground-truth" класифікатори, що супервізують навчання mRGCN. Таким чином, знання, отримані з відомих інгредієнтів, поширюватимуться через mRGCN для генерації класифікаторів для невідомих інгредієнтів. Під час тестування, для заданого зображення, запропонований фреймворк передбачатиме (розпізнаватиме) невідомий інгредієнт за допомогою оціненого класифікатора невідомих інгредієнтів.

Входом для mRGCN є векторні представлення слів-інгредієнтів, а також множинні графи відношень між інгредієнтами, а виходом mRGCN — передбачені класифікатори для кожного інгредієнта.

Шляхом мінімізації L2 відстані між передбаченими класифікаторами та "ground-truth" класифікаторами, знання, отримані з відомих інгредієнтів, поширюватимуться на невідомі інгредієнти та уможливають zero-shot розпізнавання інгредієнтів.

В [5] досліджується розпізнавання інгредієнтів їжі за допомогою багатозадачного та регіонального глибокого навчання з використанням DCNN для вирішення складнощів розпізнавання інгредієнтів через різноманітні зовнішні види, демонструючи переваги використання категорій їжі для покращення точності.

У дослідженні моделей трансформерів та передових архітектур, стаття [6] представляє зворотну систему приготування, яка генерує рецепти з зображень їжі. Цей підхід використовує нову архітектуру для прогнозування інгредієнтів без фіксованого порядку та складання інструкцій шляхом розгляду як зображення, так і інгредієнти одночасно. Він використовує ResNet-50 для вбудовування зображень і трансформери для декодування, оцінений на наборі даних Recipe1M, демонструючи покращення прогнозування інгредієнтів та здатність створювати якісні рецепти. Це демонструє інтеграцію візуального та текстового аналізу в генерації рецептів.

Досягнення в класифікації зображень їжі за допомогою моделі Vision Transformer, AlsmViT, представлені в [7]. Ця модель покращена завдяки технікам доповнення даних та посилення характеристик, включаючи Augmentplus, LayerScale та механізми багат шарового сприйняття. AlsmViT досягає значної точності 95.17% на наборі даних Food101 та 94.29% на наборі даних Vireo Food-172, перевершуючи інші методи самонавчання в класифікації зображень їжі.

Модель FL-Tran, представлена в [12], є покращенням багатоміткової класифікації зображень шляхом ефективної обробки дрібних об'єктів та виявлення прихованих характеристик. Ця модель інтегрує модуль багатомасштабної ф'южн (MSFM), модуль просторової уваги (SAM) та модуль посилення та пригнічення характеристик (FESM). FL-Tran вирішує проблему розпізнавання дрібних об'єктів і виявляє корисні характеристики, які часто затіняються більш помітними, перевершуючи існуючі методи на ключових наборах даних, таких як MS-COCO 2014, PASCAL VOC 2007 та NUS-WIDE.

В [13] застосовують багатомітковий підхід до навчання для розпізнавання змішаних страв, підкреслюючи регіональне розпізнавання та ефективність негативного семплінгу та цільових попередньо навчених моделей. Цей підхід, продемонстрований через експерименти на двох спеціально зібраних наборах даних, значно зменшує потребу в ручній маркуванні та покращує продуктивність порівняно з традиційними багатомітковими методами класифікації.

На основі інноваційних технік, в [14] інтегрується кулінарну логіку в розпізнавання інгредієнтів, використовуючи послідовне навчання. Цей метод покращує видобування характеристик за допомогою модуля подвійного потоку ф'южн характеристик та застосовує навчання з підкріпленням з гібридною функцією втрат, показуючи перспективи для покращення пізнання

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		36

їжі та потенціал для рекомендацій щодо дієти та допомоги в приготуванні їжі.

В [15] представляють метод, який поєднує мережу уваги ф'южн з модулем спільного навчання інгредієнтів їжі, спрямований на тонке розпізнавання їжі та інгредієнтів. Цей підхід особливо ефективний для складних китайських страв, досягаючи високої точності на наборі даних VIREO Food-172 шляхом зосередження на дискримінативних регіонах та вирішення проблеми нерівноважності інгредієнтів, закладаючи основу для ширших застосувань у управлінні дієтою та автоматизованих системах харчування.

Розширюючи сферу розпізнавання їжі, в [16] представляють мережу каскадної багатоуваги, керовану інгредієнтами (IG-CMAN). Використовуючи багатомасштабне, послідовне локалізацію інформативних регіонів від рівня категорії до рівня інгредієнтів, цей метод показує вражаючу продуктивність на популярних наборах даних їжі та новоствореному ISIA Food-200. Зосереджуючись на дискримінативних, пов'язаних з інгредієнтами регіонах, IG-CMAN покращує аналіз зображень їжі, обіцяючи досягнення в багатомодальному веденні щоденників їжі та персоналізованому охороні здоров'я.

В [17] обговорюється Food2Vec та FoodBERT, моделі для вбудовування інгредієнтів, які розширені до багатомодальних версій для допомоги в ідентифікації заміників інгредієнтів. FoodBERT найефективніший у рекомендаціях заміників, що свідчить про те, що його вбудовування точніше захоплюють знання про їжу, ніж традиційні моделі, відкриваючи шлях до персоналізованих дієтичних рекомендацій та покращеного розуміння харчових продуктів завдяки цільовому навчанню на зображеннях.

У поєднанні глибокого навчання та обробки природної мови, в [18] описується автоматизована методологія оцінки дієти, яка поєднує розпізнавання зображень підробленої їжі з уніфікацією назв харчових

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		37

продуктів. Досягаючи точності класифікації 92.18% та 93%, цей підхід є значним кроком у автоматизації оцінки дієти, зменшуючи ручну обробку даних та дозволяючи проводити оцінку в реальному часі через додатки для смартфонів. Це дослідження підкреслює потенціал розширення методології для включення автоматичного вимірювання ваги або об'єму їжі, спрямованого на повністю автоматизований процес оцінки дієти.

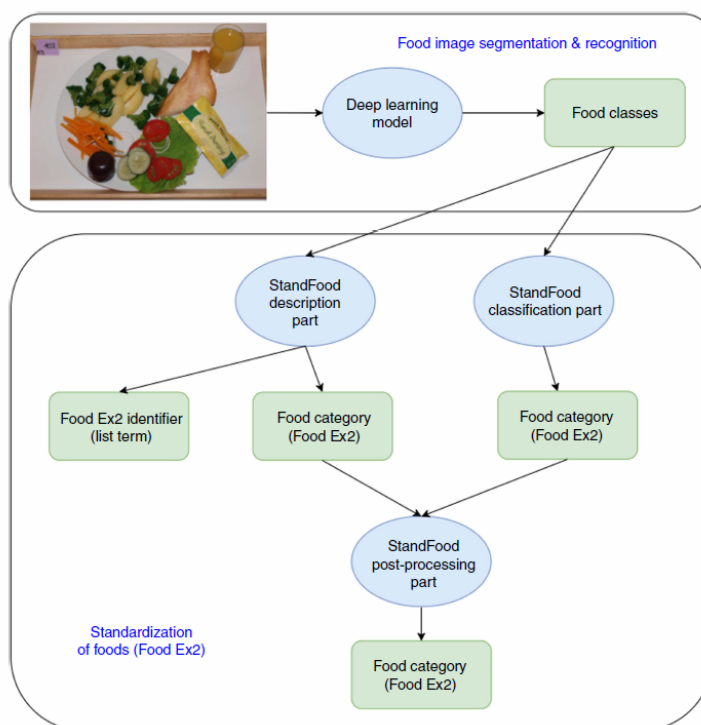


Рисунок 2.5 – Схема методології

Рисунок 2.5 показує блок-схему методології. По-перше, процес розпізнавання зображень їжі використовує зображення імітації їжі (fake-food image) для визначення класів (назв) всіх об'єктів їжі на зображенні. Ці назви їжі потім обробляються за допомогою методу StandFood для визначення дескрипторів FoodEx2 розпізнаних об'єктів їжі. Як тільки і назви їжі, і дескриптори ідентифіковано, розпізнані імітації їжі можуть бути зіставлені з даними про склад з FCDB. Кінцевим результатом, таким чином, є зображення імітації їжі, стандартизоване за допомогою унікальних дескрипторів, що

уможливорює перетворення споживання харчових продуктів у споживання нутрієнтів та сприяє автоматизованій оцінці дієти.

Процес розпізнавання зображень їжі використовує зображення імітації їжі для знаходження класів (назв) для всіх об'єктів їжі на зображенні. Вони потім обробляються за допомогою методу StandFood для визначення дескрипторів FoodEx2 розпізнаних об'єктів їжі. Як тільки і назви їжі, і дескриптори ідентифіковано, розпізнані імітації їжі можуть бути зіставлені з даними про склад з бази даних складу продуктів. Кінцевим результатом є зображення імітації їжі, стандартизоване за допомогою унікальних дескрипторів, що уможливорює перетворення споживання харчових продуктів у споживання нутрієнтів та сприяє автоматизованій оцінці дієти.

В [19] представляється FIRE, новаторський багатомодальний метод для генерації рецептів з зображень їжі. Використовуючи модель BLIP для генерації назви, Vision Transformer для видобування інгредієнтів та модель T5 для генерації кулінарних інструкцій, FIRE показує значні покращення порівняно з існуючими методами. Цей підхід дозволяє налаштовувати рецепти та автоматизувати процес приготування, задаючи напрямок для майбутніх досліджень у метриках правильності рецептів, різноманітності рецептів за допомогою графів знань та вирішення проблеми галюцинацій у генерації рецептів.

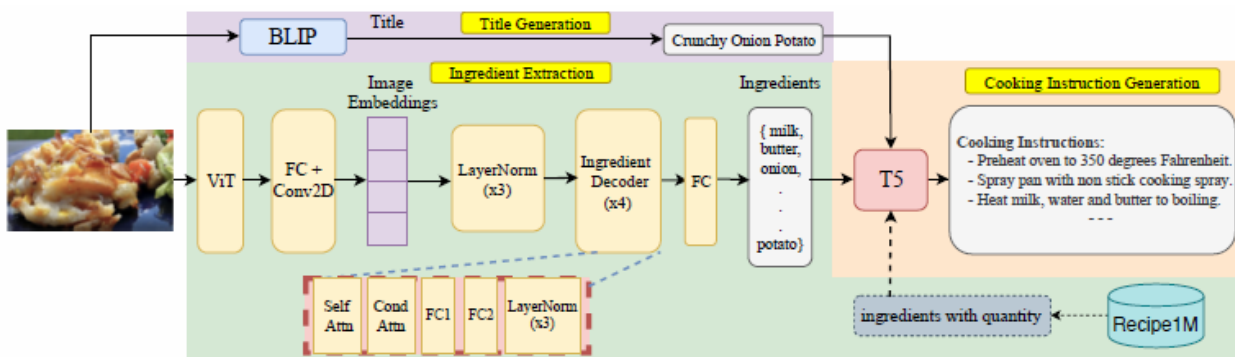


Рисунок 2.6 - Запропонована архітектура для вилучення інгредієнтів та генерації назви рецепту й інструкцій з приготування із зображення їжі

Інгредієнти із зазначенням кількості подаються тільки на етапі тренування).

Для покращення технік розпізнавання інгредієнтів, автори в [20] пропонують новий фреймворк на основі CNN для сегментації інгредієнтів на зображеннях їжі. Цей підхід не вимагає піксельної анотації та використовує стандартизовану біологічну ієрархічну структуру інгредієнтів для розробки моделі класифікації одиничних інгредієнтів, значно покращуючи ідентифікацію інгредієнтів у зображеннях їжі, як показано на наборі даних FoodSeg103.

У контексті генерації рецептів та розпізнавання їжі, в [21] представляють мережу генерації структури (SGN) для створення кулінарних інструкцій з зображень їжі та інгредієнтів. Цей метод підкреслює корисність поєднання CNN та RNN для видобування характеристик та класифікації, особливо при відсутності структурної анотації в довгих абзацах рецептів.

Вирішуючи проблеми розпізнавання їжі та обчислення калорій, в [22] представляють метод, який поєднує сегментацію зображень з класифікатором прямого розповсюдження для покращення точності класифікації. Використовуючи такі техніки, як виявлення помітних регіонів та багатомасштабна сегментація, цей підхід не лише ідентифікує харчові продукти на зображеннях, а й обчислює калорійність на основі об'єму їжі та вмісту поживних речовин, демонструючи покращення продуктивності класифікації.

## 2.2. Представлення методів автоматичної оцінки порцій

Автоматичне обчислення об'єму їжі представляє багато викликів, таких як різноманітність складів їжі, розширення асортименту інгредієнтів та різні методи приготування. Якість фотографій, використаних для оцінки об'єму їжі, є критичною, причому чіткі та добре освітлені зображення зазвичай

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		40

дають точніші оцінки, ніж розмиті або погано освітлені. На сьогоднішній день було розроблено кілька методів для точної оцінки об'єму їжі, починаючи від простого підрахунку пікселів до передових технік 3D реконструкції зображень. Ці методи поділяються на дві категорії: методи з одним зображенням та методи з багатьма зображеннями/відео. Методи з одним зображенням оцінюють об'єм їжі за допомогою лише одного фото. Ці техніки простіші в використанні, ніж методи з багатьма зображеннями, оскільки не вимагають робити кілька знімків з різних ракурсів. Однак ця простота може призвести до меншої точності оцінки об'єму порівняно з методами, які використовують кілька видів.

### *2.2.1. Методи оцінки за одним зображенням*

Методи оцінки за одним зображенням використовують лише одне фото, спрощуючи процес завдяки усуненню необхідності робити кілька знімків з різних ракурсів. Ця зручність, однак, часто йде коштом точності порівняно з методами на основі багатьох зображень або відео.

Серед найпростіших технік для оцінки порцій є традиційні ручні методи, які використовують лінійки та регульовані клини. Ці підходи, хоча й прості, вимагають безпосередньої взаємодії зі стравою, що може обмежити їх застосування в неінвазивних умовах.

Розвиток технологій привів до появи більш складних методів оцінки порцій їжі за одним зображенням. Зокрема, мобільна доповнена реальність та віртуальна реальність, пропонують динамічні та інтерактивні способи оцінки розмірів порцій. Ці методи доповнюються техніками візуальної оцінки [25], які, незважаючи на їхню схильність до людської помилки, надають цінні інсайти в оцінку порцій шляхом суб'єктивної оцінки. Автоматичні методи оцінки об'єму їжі пропонують значні переваги для осіб, які керують хронічними захворюваннями, сприяючи моніторингу дієти без необхідності безпосередньої участі експерта, що прискорює результати порівняно з

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						41
Змн.	Арк.	№ докум.	Підпис	Дата		

традиційними підходами, які зазвичай вимагають надсилання зображень їжі дієтологу. Традиційний процес вимагає постійної участі дієтолога, що обмежує їхню здатність швидко задовольняти потреби великої кількості пацієнтів. Однак значною проблемою автоматизованих систем є відсутність стандартизації, без чітких експертних рекомендацій щодо прийнятних рівнів помилок. Точність і зручність використання різняться залежно від техніки оцінки, які загалом поділяються на методи з одним зображенням та методи з багатьма зображеннями. Хоча методи з одним зображенням простіші в використанні, вони часто жертвують точністю, на відміну від методів з багатьма зображеннями, які досягають більшої точності завдяки використанню зображень з різних ракурсів.

Генеративно-суперницькі мережі (GAN) досліджувалися в [26] для оцінки енергетичної цінності порцій їжі з зображень з одним видом. Вони ввели концепцію "енергетичного розподілу" всередині кожного зображення, використовуючи набір даних з детальними анотаціями, включаючи маски сегментації та енергетичні значення. Їхній підхід на основі GAN показує обнадійливі результати з середнім рівнем помилки оцінки енергії 10.89% 10.89%, що підкреслює потенціал GAN у дієтичній оцінці.

Рисунок 2.7 демонструє приклади зображень у наборі даних Nutrition5k. На карті глибини об'єкти, що знаходяться ближче до камери, представлені синім кольором, тоді як об'єкти, що знаходяться далі, показані червоним. Колірна шкала слугує візуальним індикатором, що відображає різноманітні відстані від кожної точки до камери, з одиницею вимірювання в сантиметрах. Як метод оцінки харчування на основі зображень, наш DPF-Nutrition був оцінений на 3.5 тис. зображень їжі з набору даних Nutrition5k. Набір даних Nutrition5k надає попередньо визначене розділення, що поділяє дані на тренувальний та тестовий піднабори із співвідношенням 5:1. Існує невелика кількість помилкових зображень, як показано на рисунку 2.8, які

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						42
Змн.	Арк.	№ докум.	Підпис	Дата		

мають потенціал ввести в оману тренування моделі, що може призвести до субоптимальної ефективності.

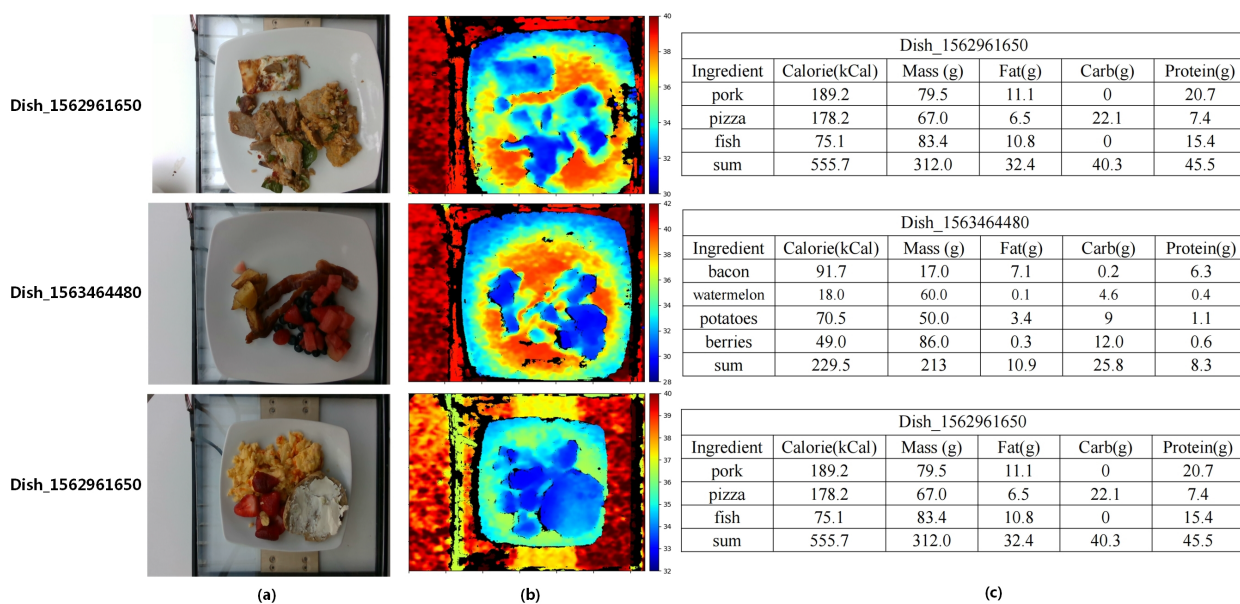


Рисунок 2.7 - Приклади зображень з набору даних Nutrition5k. (а) RGB зображення. (б) Карти глибини. (в) Нутрієнтні анотації.

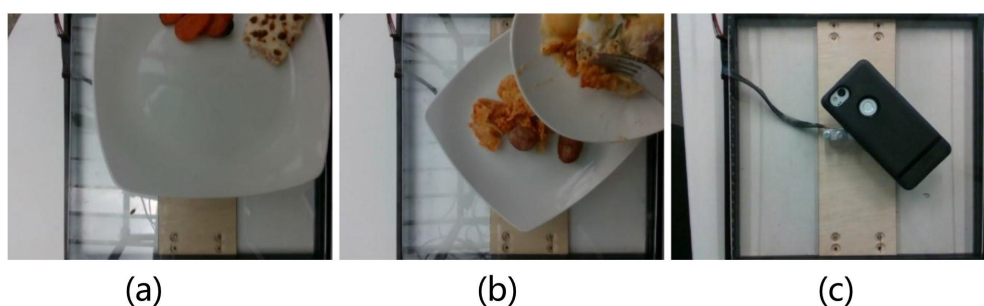


Рисунок 2.8 - Приклади некоректних зображень. (а) Їжа не повністю потрапила в кадр. (б) Страви перекриваються. (в) Зображення, що не містить їжі

В [27] пропонують метод, який використовує техніку аналізу зображень для оцінки калорійності з зображень їжі. Цей метод унікальний тим, що не покладається на ідентифікацію конкретних харчових продуктів. Натомість він використовує візуальні схожості між різними

характеристиками зображень, такими як кольорові гістограми, кольорові корелограми та характеристики SURF, для ранжування зображень за схожістю. Калорійність потім оцінюється за допомогою лінійної оцінки на основі цих рангів, використовуючи набір даних із 6512 зображень їжі з калорійністю, оціненою експертами, з сервісу FoodLog.

Робота [28] представляє FoodSAM, розвиток моделі Segment Anything Model (SAM) для сегментації зображень їжі. FoodSAM покращує обмеження SAM, поєднуючи грубі семантичні маски з масками, згенерованими SAM, для покращення якості сегментації. Він застосовує сегментацію екземплярів для розпізнавання інгредієнтів як окремі об'єкти та включає паноптичну сегментацію з детектором об'єктів для ідентифікації неїстівних об'єктів. Крім того, FoodSAM включає сегментацію за запитом для зображень їжі, підтримуючи різні типи запитів. Як перша структура, яка досягає сегментації екземплярів, паноптичної та сегментації за запитом для харчових зображень, FoodSAM показує сильну продуктивність у великих експериментах, підкреслюючи потенціал SAM у сфері аналізу зображень їжі.

В [29] представляють систему Menu-Match, розроблену для ідентифікації та оцінки калорійності ресторанних страв з зображень. Створюючи базу даних конкретних позицій меню ресторану, Menu-Match використовує передові технології комп'ютерного зору, порівнюючи вхідні зображення з відомими позиціями за допомогою таких характеристик, як кольорові гістограми, HOG, SIFT, LBP та фільтри MR8. Цей новаторський підхід спрощує процес оцінки калорійності, перетворюючи його на завдання ідентифікації, що полегшує ведення щоденника їжі.

В [30] представляють покращену структуру кодувальника-декодувальника для оцінки енергії їжі з одного монокулярного зображення, вирішуючи проблему видобування обмеженої енергетичної інформації. Цей метод перетворює зображення в формат, де енергетична інформація про їжу більш доступна, сприяє перевіреному набору даних високої якості

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		44

дієтологами, включаючи зображення, маски сегментації та калорійні значення. Демонструючи значне покращення порівняно з попередніми методами, їхній підхід зменшує середню абсолютну похибку процента (MAPE) та середню абсолютну похибку (MAE) більше ніж на 10% та 30 ккал відповідно, що свідчить про його потенціал у автоматичній оцінці дієти на основі зображень.

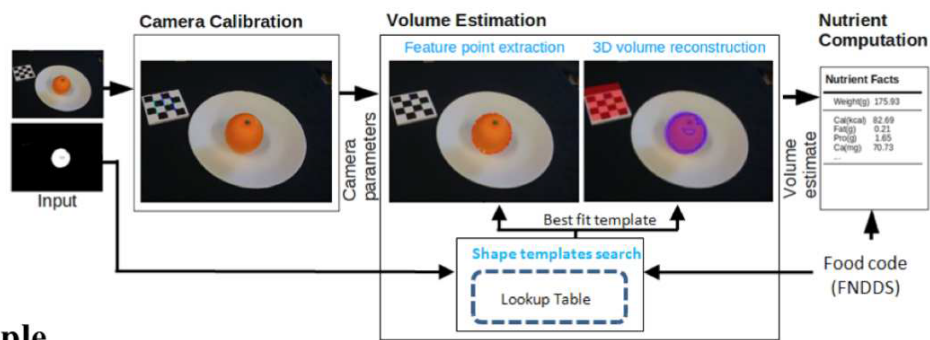
Ще одним важливим аспектом оцінки порцій є точне виділення області їжі зі зображення. В [31] застосовують TransalNet та KernelCut для уточнення кордону між харчовим продуктом та фоном на зображеннях їжі. Застосований до набору з 1027 зображень з набору даних UNIMIB2016, цей метод поєднує найсучасніші методи виявлення салієнтності на основі DNN з теорією графів для покращення точності виділення області їжі.

### 2.2.2. Методи оцінки на основі багатьох зображень та відео

Методи на основі багатьох зображень та відео значно покращують точність оцінки порцій їжі завдяки використанню зображень з різних ракурсів. Використання 3D геометричних моделей та шаблонів форм є поширеною стратегією для визначення розмірів порцій. Ці методи залежать від точної класифікації та сегментації харчових продуктів, що дозволяє ефективно застосовувати геометричні моделі для визначення точних розмірів порцій.

Оскільки продукти харчування можуть мати значні відмінності у формах, не існує єдиної геометричної моделі, яка була б придатною для всіх типів харчових продуктів. Коректна мітка класифікації їжі та маска сегментації на зображенні самі по собі є недостатніми для 3D реконструкції об'єкта їжі, отже, використання геометричних моделей дозволить оцінити об'єм. При цьому ми можемо використовувати мітку їжі для вибору відповідного класу геометричних моделей для оцінки об'єму за одним зображенням, як показано на рисунку 2.9.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		45



### Example

Input datasets		Results	
<p><b>Meal Image</b></p>	<p><b>Segmented Images</b></p> <p>Food Code            - 11112110 (Milk)            - 53105500 (Chocolate cake)</p>	<p><b>Best-Fit Geometric Model</b>            - Cylinder, Sphere, Prism</p>	<p><b>Feature Points</b></p>
		<p><b>Estimated volumes</b></p>	

Рисунок 2.9 - Оцінка порції за допомогою геометричних моделей

В [36] представляють метод, який використовує 3D моделювання та оцінку пози для оцінки об'єму їжі з зображень. Ця техніка ідентифікує та кількісно оцінює харчові продукти всередині одного зображення, покращуючи оцінку вмісту поживних речовин завдяки точній оцінці об'єму.

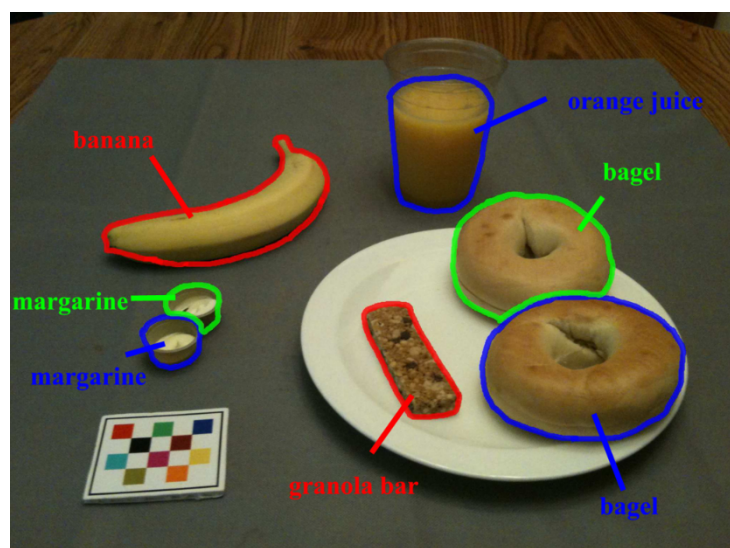


Рисунок 2.10 - Приклад зображення їжі з шахівницею у сцені

Кожен об'єкт їжі сегментується та ідентифікується за допомогою системи оцінки харчування.

Системи доповненої реальності (AR), такі як Eat AR та Serv AR , використовують орієнтирні маркери або стандартні об'єкти всередині кадру зображення. Ці системи обчислюють об'єм їжі, накладаючи 3D форми на захоплену сцену, тим самим забезпечуючи точне вимірювання об'єму. Методи щільної 3D реконструкції моделей створюють детальні 3D моделі з багатьох точок зору. Цей підхід дозволяє надзвичайно точно оцінювати об'єм їжі.

У іншому підході в [41] розробили новий LiDAR-систему машинного зору, яка включає спеціально розроблений роликівий конвеєр для забезпечення неруйнівного онлайн-вимірювання об'єму бататів. Використовуючи передові техніки аналізу, включаючи множинну лінійну регресію та нейронні мережі, система досягла вражаючої точності 97.9% у оцінці об'єму, демонструючи її потенціал для високопродуктивних сільськогосподарських застосувань.

Ці методи на основі багатьох зображень та відео, разом із найновішими технологіями сегментації та оцінки енергії, сприяють розвитку сфери дієтичної оцінки. Вони пропонують точніші та зручніші варіанти для оцінки порцій їжі, відповідаючи потребам осіб, які стежать за своїм харчуванням. Завдяки використанню передових технологій машинного навчання та комп'ютерного зору ці підходи значно покращують точність ідентифікації та вимірювання порцій їжі. Це покращення особливо корисне, оскільки воно мінімізує помилки, характерні для старіших, більш ручних технік. Крім того, інтеграція цих передових технологій у інструменти дієтичної оцінки має потенціал зробити відстеження споживання їжі простішим та ефективнішим для користувачів. Це, у свою чергу, може спонукати людей дотримуватися відстеження свого харчування, сприяючи кращим харчовим звичкам та інформованим рішенням щодо їжі.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						47
Змн.	Арк.	№ докум.	Підпис	Дата		

### 2.3. Набори даних, використані для розпізнавання їжі

Продуктивність методів виділення характеристик і класифікації в значній мірі залежить від детально зібраної колекції зображень, які в нашому випадку є зображеннями їжі. Такі набори даних, як UECFOOD-100, Food-101, UECFOOD256, UNCIT-FD1200 і UNCIT-FD889, врешті-решт використовуються як еталони для порівняння продуктивності існуючих підходів з новими класифікаторами. Такі набори даних можуть бути відмінними за характеристиками, такими як загальна кількість зображень у певному наборі даних, тип кухні та включені категорії їжі.

Наприклад, UECFOOD-100 містить 100 різних сортів категорій їжі, і кожна категорія їжі має обмежувальну рамку, яка вказує на місце розташування харчового продукту на фотографії. Категорії їжі в цьому наборі даних головним чином належать до популярних страв у Японії. Подібно, UECFOOD-256 — це ще один варіант UECFOOD-100. Однак він відрізняється за кількістю зображень, оскільки містить 256 зображень їжі різних видів. Food-101 містить 101 000 реальних зображень, класифікованих у 101 категорію їжі. Він включає різноманітні, але візуально схожі класи їжі. Подібно, набір даних RFID містить 1098 зображень їжі з 61 різних категорій. Колекція RFID наразі має три екземпляри 101 швидкої їжі.

UNCIT-FD1200 складається з 4754 зображень їжі 1200 видів страв, зроблених з реальних трапез. Кожну тарілку з їжею знімають кілька разів, і загальний набір даних представляє як геометричну, так і фотометричну варіативність. Подібно, набір даних UNICT-FD889 містить 3583 зображення [45] 889 різних реальних тарілок з їжею, зроблених за допомогою мобільних пристроїв в неконтрольованих сценаріях (наприклад, різні фони та умови освітлення). Крім того, кожне зображення страви в UNICT-FD889 знімають кілька разів, щоб забезпечити геометричну та фотометричну варіативність (зміни в обертанні, масштабі та точці зору).

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		48



Таким чином, результати проведеного огляду свідчать про нагальну потребу у створенні широких та універсальних (або репрезентативних) наборів даних про їжу з метою покращення розпізнавання харчових продуктів та підвищення ефективності. Ця необхідність зумовлена тим, що регіонально-специфічні продукти харчування або набори даних з меншою кількістю категорій їжі можуть знизити точність та ефективність методів класифікації та вилучення (ознак).

### Висновки до розділу

У розділі розглянуто ключові моделі та методи глибокого навчання, необхідні для розробки додатку розрахунку харчової цінності їжі. Детально описано процес розпізнавання інгредієнтів за допомогою глибоких нейронних мереж, де основну роль відіграють архітектури згорткових нейронних мереж (CNN) та сучасні трансформерні моделі. Визначено основні етапи обробки зображень: попередня обробка, детекція об'єктів, класифікація продуктів і подальша оцінка їх харчової цінності.

Аналіз методів автоматичної оцінки порцій показав існування двох основних підходів:

- оцінка за одним зображенням, яка є простішою у реалізації, але менш точною через обмеженість просторової інформації;
- оцінка на основі багатьох зображень або відео, що дозволяє отримати тривимірну інформацію про їжу і суттєво підвищити точність розрахунку об'єму та ваги порцій.

Особливу увагу приділено аналізу доступних наборів даних для розпізнавання їжі, таких як Food-101, UECFOOD-256, VIREO-172 та інших. Набори даних оцінено за їх різноманітністю, якістю анотацій та придатністю для навчання моделей, що працюють у реальних умовах.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		50

Підсумовуючи, у розділі обґрунтовано вибір технологій і методів для реалізації додатку: застосування глибоких нейронних мереж для розпізнавання інгредієнтів, використання багатоканальних методів для точної оцінки порцій та важливість якісних датасетів для ефективного навчання моделей. Ці результати створюють основу для розробки високоточних і зручних у використанні сервісів автоматизованого аналізу харчування.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		51

## РОЗДІЛ 3. РЕАЛІЗАЦІЯ СЕРВІСУ ДЛЯ РОЗРАХУНКУ ХАРЧОВОЇ ЦІННОСТІ ЇЖИ НА ОСНОВІ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

### 3.1. Огляд архітектури сервісу

У даному розділі детально описано архітектуру та функціонування запропонованого сервісу для оцінки харчування на основі зображень. Основний фокус розробленого сервісу зосереджено на завданні автоматичного розпізнавання харчових інгредієнтів у складних стравах, тоді як етап кількісної оцінки порцій реалізовано шляхом ручного введення даних користувачем. Такий підхід визначає систему як "напівавтоматизовану" і дозволяє зосередити дослідницькі та інженерні зусилля на вирішенні найбільш складного аспекту автоматизації – якісної ідентифікації компонентів страви.

Загальна схема запропонованого сервісу представлена на рисунку 3.1. Процес обробки даних починається із захоплення візуальної інформації про страву. Користувач ініціює процес, завантажуючи зображення страви з пам'яті мобільного пристрою або здійснюючи фотографування безпосередньо за допомогою інтегрованої функції камери. Якість та ракурс зйомки можуть впливати на подальші етапи розпізнавання.

Отримане зображення подається на вхід першого модуля сервісу, який відповідає за аналіз візуального вмісту. На цьому етапі використовується передова візуально-мовна модель, така як VLIP-2. Модель VLIP-2 здійснює глибокий аналіз вхідного зображення та генерує його детальний текстовий опис (image caption). Цей текстовий опис слугує проміжним представленням, ефективно переводячи складну візуальну інформацію у структурований або напівструктурований текстовий формат, який є більш придатним для подальшої обробки потужними мовними моделями.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		52

Сгенерований текстовий опис страви далі передається до другого ключового модуля сервісу, реалізованого на базі великої мовної моделі, такої як GPT-3.5. Шляхом застосування методів обробки природної мови (Natural Language Processing, NLP) модель GPT-3.5 виконує завдання екстракції (вилучення) інгредієнтів із текстового опису. На цьому етапі модель ідентифікує та виокремлює назви окремих харчових компонентів, які були згадані або описані моделлю BLIP-2. Точність цього кроку залежить як від якості початкового опису, так і від лінгвістичних можливостей моделі GPT-3.5 і її здатності розрізняти інгредієнти серед іншого тексту.

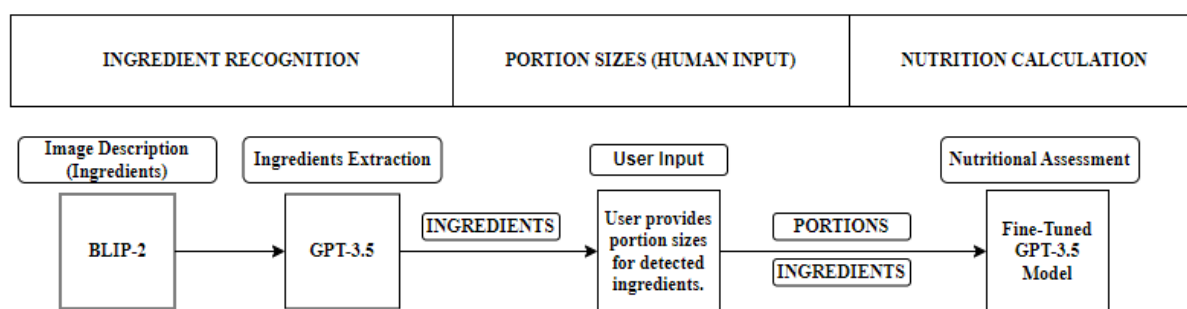


Рисунок 3.1 – Візуалізація архітектури сервісу та його основних компонентів

Ідентифікований перелік інгредієнтів автоматично відображається в інтерфейсі мобільного додатка. На цьому етапі система переходить у напівавтоматичний режим: користувач отримує можливість верифікувати розпізнані інгредієнти (додати відсутні, видалити некоректно визначені) та, що найважливіше, ввести інформацію про розмір спожитої порції для кожного інгредієнта (наприклад, в грамах, мілілітрах, штуках або використовуючи стандартні мірки). Незважаючи на те, що цей крок вимагає активної участі користувача, він є необхідним для отримання кількісних даних, які не можуть бути надійно визначені лише за одним 2D зображенням без спеціалізованих методів оцінки порцій.

Після введення користувачем даних про кількість/об'єм порцій для кожного інгредієнта, здійснюється автоматичний розрахунок нутрієнтного профілю страви. Для виконання цього розрахунку використовується модель GPT-3.5, яка була попередньо дотренована або налаштована на інтеграцію з верифікованими базами даних складу продуктів, такими як USDA FoodData Central та/або іншими релевантними національними чи міжнародними нутрієнтними банками даних. Модель інтерпретує пару "інгредієнт - кількість" та звертається до відповідних даних у базі для агрегування інформації про поживну цінність. Розрахований нутрієнтний профіль є комплексним та деталізованим, надаючи інформацію як про основні макронутрієнти (білки, жири, вуглеводи), так і про широкий спектр мікронутрієнтів (вітаміни, мінерали), що містяться у спожитій порції страви.

Отримані дані про розпізнані інгредієнти, введені користувачем порції та деталізований нутрієнтний профіль страви зберігаються в інтегрованій базі даних (наприклад, на базі FirebaseDB). Це дозволяє користувачеві вести історію свого харчування, відстежувати прогрес, а також створює основу для подальшого аналізу накопичених даних, хоча детальний аналіз цих даних виходить за межі даного дослідження.

Загальна схема сервісу, що ілюструє послідовність цих етапів, представлена на рисунку 3.1. Далі в цьому розділі буде представлено детальний опис кожного з ключових компонентів запропонованого сервісу (модулів VLP-2, GPT-3.5 та їхньої взаємодії), їхню технічну реалізацію та обґрунтування вибору саме цих моделей.

### **3.2 Модуль розпізнавання інгредієнтів на основі візуально-мовних моделей**

Детальний опис ключових компонентів сервісу, що відповідають за розпізнавання інгредієнтів, представлено нижче. Першим етапом сервісу є

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						54
Змн.	Арк.	№ докум.	Підпис	Дата		

перетворення візуальної інформації із зображення страви в текстове представлення – генерація опису.

На початковому етапі сервісу застосовується передова модель візії-мовлення BLIP-2 (Bootstrapping Language-Image Pre-training). Ця модель є мультимодальною, ефективно інтегруючи можливості обробки візуальної інформації (аналіз зображень) та генерації тексту (робота з природною мовою). Архітектурно BLIP-2 поєднує візуальний енкодер та мовну модель (зокрема, у використовуваній версії інтегровано Vicuna LLM – велику лінгвістичну модель) для виконання завдань, що вимагають глибокого розуміння зв'язку між зображеннями та текстом, таких як генерація підписів (captioning) або відповіді на запитання за зображенням (Visual Question Answering, VQA).

Однією з ключових властивостей BLIP-2, що є критично важливою для нашого сервісу, є її запитуваність (queryability). Це уможливорює застосування технік інженерії запитів (prompt engineering) – ретельного формулювання текстового запиту, що подається моделі разом із зображенням. Шляхом формулювання специфічного запиту ми можемо спрямувати генерацію тексту моделлю таким чином, щоб отримати інформацію, максимально релевантну для подальшого вилучення інгредієнтів, тим самим підвищуючи точність та якість генерованого опису. З метою отримання лаконічного та сфокусованого на інгредієнтах опису, фінальна версія запиту, що застосовувалася на цьому етапі сервісу, мала наступний вигляд:

"Перелічіть основні харчові інгредієнти, показані на зображенні. Надайте лаконічну відповідь без додаткових коментарів щодо страви або її оточення."

Сгенерований моделлю BLIP-2 текстовий опис страви, що являє собою перелік інгредієнтів у форматі природної мови, є вхідними даними для наступного етапу обробки. На цьому етапі застосовується ще одна потужна

						Арк.
					БР.ПЗ – 08.00.00.000 ПЗ	55
Змн.	Арк.	№ докум.	Підпис	Дата		

велика лінгвістична модель – GPT-3.5. Роль GPT-3.5 полягає у парсингу (аналізі) наданого текстового опису з метою ідентифікації та прецизійного вилучення (екстракції) з нього конкретних сутностей – назв харчових інгредієнтів. Цей крок є критично важливим для мінімізації ймовірності генерації галюцинацій – вигадування моделлю неіснуючих інгредієнтів, які не були згадані в описі від VLP-2 або фактично відсутні на зображенні.

Шляхом ретельної розробки запитів до GPT-3.5 було досягнуто точного вилучення інгредієнтів у строго структурованому форматі переліку. Такий формат є зручним для подальшої автоматизованої програмної обробки та інтеграції з іншими компонентами сервісу. Запит, який ми використовували для екстракції інгредієнтів із тексту, був сформульований наступним чином:

"З наданого речення, {resp}, вилучіть харчові інгредієнти, форматуючи вихід як список (напр., [інгредієнт1, інгредієнт2, інгредієнт3]). Наприклад, якщо надано 'Рецепт включає яблука, апельсини та курку', відповідь повинна бути [яблуко, апельсин, курка], концентруючи увагу виключно на справжніх харчових компонентах."

Де {resp} є плейсхолдером для текстового опису, отриманого від моделі VLP-2. Наявність прикладу у запиті (few-shot prompting) додатково сприяє спрямуванню моделі на бажаний формат виводу та тип сутностей. Цей процес прецизійного вилучення інгредієнтів забезпечує безперервність та ефективність переходу до наступного етапу нашого сервісу, що включає взаємодію з користувачем для введення порцій та подальший розрахунок поживної цінності.

Приклади зображень страв, які можуть бути оброблені цим конвеєром на етапах генерації опису та вилучення інгредієнтів, показані на рисунках 3.3 та 3.4 ілюструючи різноманітність вхідних візуальних даних.

## Meal Nutrition Calculator


Your Unique ID: b034cda2-1386-4fe2-8b2f-1110b4b20d2b

Choose an image...

Drag and drop file here  
Limit 200MB per file • JPG, JPEG

[Browse files](#)

test7.jpg 65.5KB



Uploaded Image.

Ingredient Portions

Enter the portion size for Shrimp (in grams)  
0

Enter the portion size for Broccoli (in grams)  
0

Enter the portion size for Rice (in grams)  
0

Enter name of the ingredient to add

Enter name of the ingredient to remove

[UPDATE](#)

[SUBMIT](#)

Рисунок 3.2 - Приклад зображення страви, що містить рис, броколі та креветки

Meal Nutrition Calculator


Your Unique ID: 0045ac5f-8700-4527-0188-0488ca5334d

Choose an image...

Drag and drop file here  
Limit 200MB per file • JPG, JPEG

[Browse files](#)

50284.jpg 171.1KB



Ingredient Portions

Enter the portion size for Salmon (in grams)  
0

Enter the portion size for Lemon (in grams)  
0

Enter the portion size for Asparagus (in grams)  
0

Enter name of the ingredient to add

Enter name of the ingredient to remove

[UPDATE](#)

[SUBMIT](#)

Meal Nutrition Calculator


Your Unique ID: b034cda2-1386-4fe2-8b2f-1110b4b20d2b

Choose an image...

Drag and drop file here  
Limit 200MB per file • JPG, JPEG

[Browse files](#)

18182.jpg 224.2KB



Ingredient Portions

Enter the portion size for Pulled pork (in grams)  
0

Enter the portion size for Fried potatoes (in grams)  
0

Enter the portion size for Broccoli (in grams)  
0

Enter name of the ingredient to add

Enter name of the ingredient to remove

[UPDATE](#)

[SUBMIT](#)

Рисунок 3.3 - Приклад зображення страви, що містить картоплю, спаржу, лосось, свинину та картоплю фрі

### 3.3. Модуль взаємодії з користувачем та кількісна оцінка порцій

Після успішного завершення етапу автоматичного вилучення інгредієнтів із зображення за допомогою моделей VLP-2 та GPT-3.5,

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		57

управління сервісом переходить до модуля взаємодії з користувачем. Цей етап є ключовим для напівавтоматизованої природи запропонованої системи, оскільки він дозволяє користувачеві не лише верифікувати результати автоматичного розпізнавання, але й надати критично важливу кількісну інформацію про спожиту порцію страви.

Ідентифікований на попередньому етапі перелік інгредієнтів відображається в спеціалізованому інтерфейсі мобільного додатка. На цьому кроці користувачу надається можливість внести необхідні коригування до переліку розпізнаних компонентів та, що найважливіше в контексті нутрієнтної оцінки, ввести кількісні параметри спожитої порції для кожного ідентифікованого інгредієнта.

На поточній стадії реалізації системи як основна метрика для введення порцій використовується маса в грамах. Вибір цієї метрики зумовлений її прямою кореляцією з даними в більшості стандартних нутрієнтних баз даних (таких як USDA) та відносною простотою інтеграції на даному етапі розробки. Однак, для підвищення зручності використання системи у майбутньому, передбачається реалізація підтримки альтернативних одиниць вимірювання (наприклад, об'єм у мілілітрах, стандартні мірки, штуки) та функціоналу автоматичної конвертації між ними на основі даних про щільність продуктів та типові розміри порцій.

Для підвищення точності фінального нутрієнтного розрахунку та компенсації потенційних помилок автоматичного розпізнавання, реалізовано функціонал ручного редагування переліку інгредієнтів. Це надає користувачеві можливість додати інгредієнти, які могли бути не виявлені автоматизованою системою (наприклад, приховані під іншими компонентами або візуально складні для розпізнавання), або видалити компоненти, які були некоректно ідентифіковані (наприклад, "галюцинації" моделі або помилки класифікації). Такий людський контроль на етапі введення даних про

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		58

інгредієнти є важливим для забезпечення коректності вхідних даних для подальшого розрахунку поживної цінності.

Після верифікації користувачем розпізнаних інгредієнтів та введення кількісних даних про порції для кожного з них, він може ініціювати подальшу обробку даних. Активація функції, позначеної як "ОНОВИТИ" (або аналогічної за змістом), зазвичай запускає внутрішню обробку та валідацію введених даних, готуючи їх до нутрієнтного розрахунку. Коли користувач готовий отримати фінальний нутрієнтний профіль наданої страви, він активує функцію "НАДІСЛАТИ". Ця дія передає фінальний, верифікований користувачем перелік інгредієнтів із зазначенням їхніх кількісних параметрів для подальшої обробки до відповідного сервісу або модуля системи, де здійснюється безпосередньо нутрієнтна оцінка.

## Meal Nutrition Calculator

Your Unique ID: b034cda2-1386-4fe2-8b2f-1110b4b20d2b

Choose an image...

Drag and drop file here  
Limit 200MB per file • JPG, JPEG

Browse files

test1.jpg 223.8KB



Uploaded image.

Ingredient Portions

Enter the portion size for Chicken (in grams)

100 - +

Enter the portion size for Sweet Potato Fries (in grams)

100 - +

Enter the portion size for Salad (in grams)

50 - +

Enter name of the ingredient to add

Enter name of the ingredient to remove

UPDATE

SUBMIT

Submitted successfully!

Рисунок 3.4 - Інтерфейс користувача додатка, що демонструє кнопки "ОНОВИТИ" та "НАДІСЛАТИ"

Отримані в результаті нутрієнтної оцінки дані про склад страви можуть бути збережені в базі даних додатка для подальшого поглибленого аналізу,

									Арк.
									59
Змн.	Арк.	№ докум.	Підпис	Дата	БР.ПЗ – 08.00.00.000 ПЗ				

зокрема для виявлення індивідуальних дієтичних тенденцій, моніторингу споживання певних нутрієнтів з плином часу або для інших дослідницьких цілей, що стосуються здоров'я та харчування. Однак, детальний аналіз та використання накопичених нутрієнтних даних виходить за межі даного дисертаційного дослідження, яке зосереджене на розробці та оцінці самого сервісу розпізнавання та кількісної оцінки.

Інтерфейс користувача мобільного додатка, через який здійснюється взаємодія на цьому етапі, візуально представлено на рисунку 3.2. Загалом, інтерфейс відображає етап взаємодії користувача, де після завантаження фотографії та автоматичного розпізнавання, користувач верифікує інгредієнти, вводить їхні порції та надсилає дані для розрахунку поживної цінності. Аспекти розробки цього інтерфейсу, включаючи численні ітерації дизайн-процесу, необхідні для забезпечення інтуїтивно зрозумілого, ефективного та плавного досвіду взаємодії з користувачем при введенні даних про інгредієнти та порції, також виходять за межі даного дисертаційного дослідження, присвяченого, в першу чергу, технічним аспектам автоматизованого аналізу зображень та даних.

### 3.4. Модуль розрахунку нутрієнтної цінності

Етап обчислення нутрієнтної цінності є фінальним кроком у конверсії оцінки харчування, що здійснюється після успішного розпізнавання інгредієнтів та введення користувачем даних про їхні порції. На цьому етапі реалізовано використання моделі GPT-3.5, яка була спеціально налаштована або дотренована для взаємодії з верифікованими базами даних складу продуктів, зокрема з набором даних USDA FoodData Central та потенційно іншими релевантними джерелами нутрієнтної інформації.

Модель GPT-3.5 на цьому етапі функціонує як інтелектуальний інтерфейс до бази даних. Вона обробляє структурований вхідний перелік

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						60
Змн.	Арк.	№ докум.	Підпис	Дата		

інгредієнтів та їхні відповідні кількісні параметри (вагу в грамах або інші метрики, конвертовані у стандартизований формат), отримані з попереднього етапу взаємодії з користувачем. На основі цих даних модель здійснює запит до нутрієнтної бази даних, агрегує інформацію про вміст поживних речовин для кожного компонента з урахуванням його кількості та генерує комплексний нутрієнтний профіль для всієї страви.

Важливою проблемою, що виникає при роботі з будь-якими базами даних складу продуктів, є їхня потенційна невичерпність. Нутрієнтні дані можуть бути неповними, оскільки нові продукти, специфічні способи їх приготування, регіональні страви або рідкісні інгредієнти можуть бути відсутніми або недостатньо представленими в наборі даних, що використовувався для налаштування моделі GPT-3.5. Це може призвести до неповного або некоректного розрахунку поживної цінності, якщо система не може зіставити ідентифікований інгредієнт із записами в базі даних.

Для подолання цієї проблеми та поступового збагачення знань системи реалізовано механізм зворотного зв'язку та оновлення бази даних, що включає людину в контурі (human-in-the-loop). У випадку, коли ідентифікований користувачем інгредієнт (або його конкретна варіація чи спосіб приготування) не може бути зіставлений з наявними даними в нутрієнтній базі, інформація про цей "невідомий" інгредієнт автоматично фіксується. Ці дані надсилаються до спеціалізованого сховища (наприклад, у базу даних Firebase) у категорію "Невідомі інгредієнти" для подальшого аналізу.

Ці "невідомі інгредієнти" підлягають верифікації та дослідженню спеціалістами, наприклад, дієтологами або медичною командою. Завданням експертів є ідентифікувати інгредієнт, знайти його достовірні нутрієнтні дані в інших надійних джерелах (наукові публікації, офіційні нутрієнтні банки даних інших країн тощо) та внести відповідну інформацію до основної бази даних складу продуктів. Після оновлення бази даних здійснюється повторне

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		61

налаштування або дотренування моделі GPT-3.5 на актуалізованому наборі даних. Це гарантує, що при подальших запитах цей інгредієнт буде коректно розпізнано моделлю, а його нутрієнтна цінність буде врахована при фінальному розрахунку нутрієнтного профілю страви, таким чином підвищуючи вичерпність та точність системи з часом.

У випадку успішної ідентифікації всіх інгредієнтів у базі даних та отримання відповідних нутрієнтних даних, модель розраховує повний нутрієнтний профіль для страви. Цей профіль є комплексним, надаючи детальні дані про вміст макронутрієнтів (білків, жирів, вуглеводів), а також широкий спектр мікронутрієнтів (вітамінів, мінералів) у спожитій порції. Отриманий нутрієнтний профіль може бути збережений у внутрішній базі даних системи для ведення історії харчування користувача або ж безпосередньо візуалізований в інтерфейсі додатка для миттєвого доступу до інформації.

```

— Fatty acids, total monounsaturated (g):6.08
— Fatty acids, total polyunsaturated (g):3.19
— Fatty acids, total saturated (g):4.24
— Folate, DFE (µg):30
— Folate, food (µg):30
— Folate, total (µg):30
— Glutamic acid (g):2.68
— Glycine (g):1.22
— Histidine (g):0.53
— Iron, Fe (mg):1.31
— Isoleucine (g):0.91
— Leucine (g):1.34
— Lysine (g):1.47
— Magnesium, Mg (mg):20
— Manganese, Mn (mg):0.03
— Methionine (g):0.48
— Niacin (mg):6.64
  
```

Рисунок 3.5 - Фрагмент нутрієнтного профілю страви

В рамках даної роботи отриманий нутрієнтний профіль страви зберігається у внутрішній системі переважно для цілей верифікації та аналізу

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						62
Змн.	Арк.	№ докум.	Підпис	Дата		

ефективності сервісу нутрієнтного розрахунку, а не для поглибленого аналізу дієтичних тенденцій користувачів, що виходить за межі поточної роботи. Фрагмент типового нутрієнтного профілю, що генерується системою на цьому етапі, представлено на рисунку 3.5.

### 3.4. Налаштування та опис сервісу

Виконання виведення моделі (model inference), що є найбільш обчислювально інтенсивним етапом у конвеєрі обробки зображень та даних, здійснювалося на виділеному сервері. Цей сервер був конфігурований з достатніми обчислювальними ресурсами, імовірно, включаючи прискорення на базі графічних процесорів (GPU), необхідне для ефективної роботи моделей глибокого навчання (BLIP-2, GPT-3.5).

З погляду програмної реалізації, архітектура розробленої системи є модульною та складається з трьох ключових компонентів: компонента бази даних, серверної частини (backend) та клієнтської частини (frontend). Компонент бази даних відповідає за стійке зберігання вхідних зображень (опціонально), розпізнаних інгредієнтів, введених користувачем порцій, розрахованих нутрієнтних профілів та інформації про "невідомі інгредієнти" для подальшого аналізу та оновлення. Серверна частина реалізує основну бізнес-логіку системи, включаючи управління послідовністю виконання етапів сервісу, обробку запитів від клієнтської частини, взаємодію з компонентом бази даних та головне – організацію викликів до сервера виведення моделі. Клієнтська частина забезпечує графічний інтерфейс користувача для взаємодії із системою (завантаження зображень, верифікація інгредієнтів, введення порцій) та візуалізації результатів.

Як основна мова програмування для розробки компонентів системи використовувався Python. Вибір Python зумовлений його широкою екосистемою бібліотек та фреймворків, що є стандартами де-факто у галузях

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		63

машинного навчання, обробки даних та веб-розробки. Клієнтська частина (frontend) була розроблена за допомогою Streamlit, що є пакетом на Python, призначеним для спрощеного та швидкого створення веб-додатків та інтерактивних дашбордів, що особливо зручно для прототипування, демонстрації моделей та розробки інтерфейсів для збору даних.

Взаємодія між компонентами системи (клієнтська частина, серверна частина, сервер виведення моделі) реалізована за допомогою програмних інтерфейсів (API - Application Programming Interface). Було розроблено два основні API: перший API забезпечує зв'язок між клієнтською та серверною частинами, приймаючи вхідні дані від користувача та надсилаючи відповіді. Другий API слугує для взаємодії серверної частини безпосередньо із сервером, на якому розміщено та виконується виведення моделей глибокого навчання. Така архітектура, заснована на розподілених компонентах, що взаємодіють через API, забезпечує високу модульність системи, чіткий поділ відповідальності між частинами та уможливує незалежне масштабування кожного компонента відповідно до навантаження.

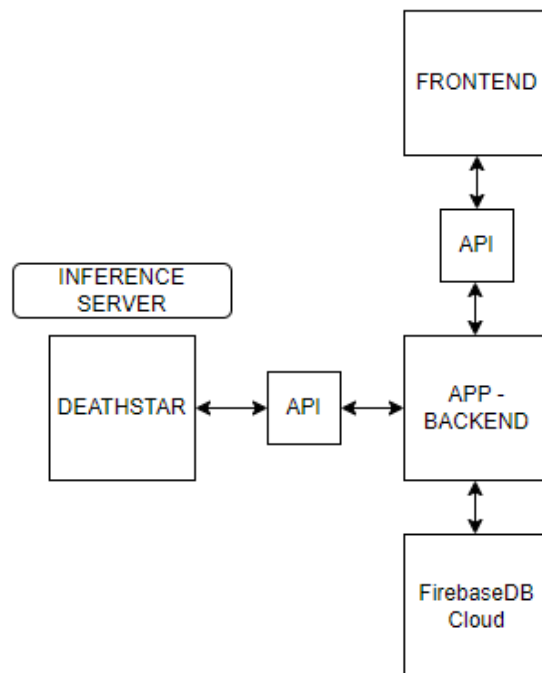


Рисунок 3.6 – Системна архітектура додатку

Зокрема, це надає значну гнучкість у розгортанні, дозволяючи легко переключати сервер виведення моделі на іншу апаратну платформу (наприклад, з більш потужними або економічними прискорювачами) у разі необхідності підвищення продуктивності, зниження витрат або зміни вимог до ресурсів, без суттєвих змін у роботі інших частин системи.

Розгортання додатка та його компонентів здійснювалося на серверах AWS (Amazon Web Services), що надає доступ до гнучких, масштабованих та надійних хмарних ресурсів. Загальна архітектура системи, що ілюструє взаємодію її компонентів та потоки даних, представлена на рисунку 3.6.

### **3.5. Опис наборів даних для оцінки ефективності сервісу**

Для оцінки ефективності нашого сервісу, зокрема компонента розпізнавання інгредієнтів, ми обрали набори даних UECFOOD100 та UECFOOD256. Це рішення було обумовлено тим, що додатки для харчування "MyFoodCam" і "FoodTracker", згадані в нашому розділі "Пов'язані роботи", також тестували свою продуктивність за допомогою цих наборів даних.

Набір даних UECFOOD100 містить 13 920 зображень у 100 категоріях, зосереджених на японській їжі, зокрема на популярних стравах у Японії. Зображення, зроблені за допомогою мобільних камер, роблять цей набір даних ідеальним для тестування мобільного додатка для харчування в реальних умовах. Приклади цих зображень представлені на рисунку 3.8.

Водночас набір даних UECFOOD256 містить 30 290 зображень у 256 категоріях, також зосереджених на японській їжі, зроблених за допомогою мобільних камер, деякі зображення показані на рисунку 3.9. Він активно використовується як бенчмарк для тестування продуктивності моделей машинного навчання, зокрема глибоких нейронних мереж, у задачі класифікації зображень страв. Зображення в UECFOOD256, як правило,

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		65

демонструють окремі страви, що робить його придатним для навчання моделей, які ідентифікують страву як єдиний об'єкт, а не окремі інгредієнти в її складі

Крім того, ми використали внутрішній набір даних. Цей набір даних включає 38 зображень страв з різних, переважно західних, кухонь. Було вручну позначено інгредієнти на цих зображеннях, які були зроблені за допомогою смартфона в природних умовах освітлення, що максимально наближає їх до реальних ситуацій. Велика частина цих зображень показана на рисунку 3.7.



Рисунок 3.7 - Зображення страв із внутрішнього набору даних

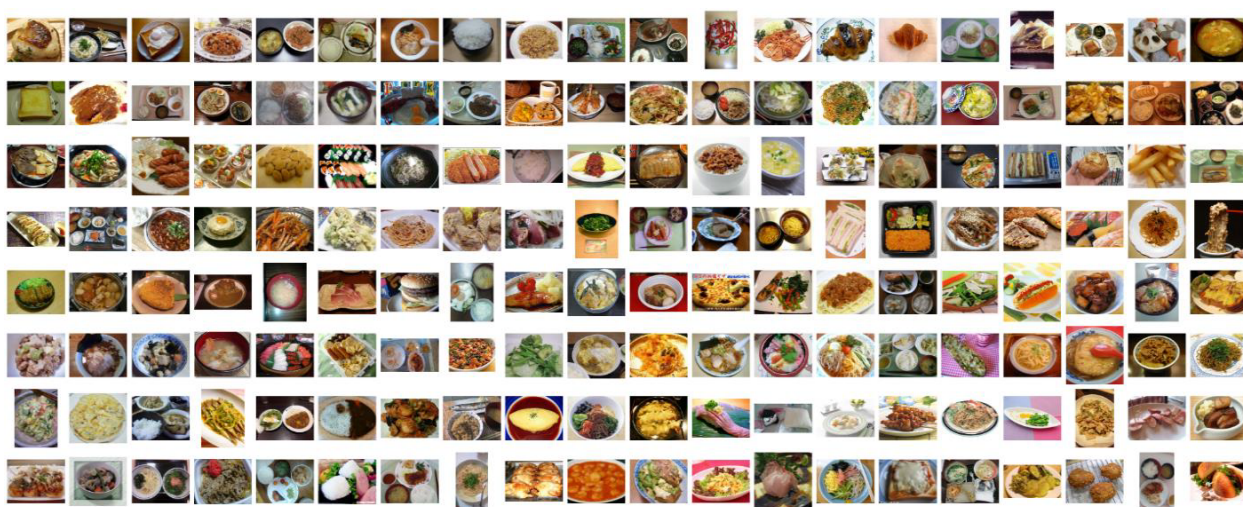


Рисунок 3.8 - Зображення страв з набору даних UECFOOD100



Рисунок 3.9 - Зображення страв з набору даних UECFOOD256

### 3.6. Експериментальне дослідження сервісу

У наших експериментах ми організували процес у два основні етапи. На першому етапі ми зосередилися на налаштуванні гіперпараметрів моделі VLLM-2/Vicuna. Завдяки всебічному тестуванню різних конфігурацій та інженерії запитів ми визначили оптимальні гіперпараметри для великої мовної моделі (LLM) та найефективніші запити для генерування описів зображень та вилучення інгредієнтів. Ми також внесли конкретні коригування в інші гіперпараметри для подальшого уточнення виходу. Ураховуючи те, що модель мови використовується для генерування описів, що призводить до варіативності виходу при кожному виконанні, ми зосередилися на налаштуванні параметрів температури та max\_tokens для управління цією варіативністю. Параметр температури контролює рівень непередбачуваності у прогнозах моделі — нижчі значення призводять до меншої варіативності, тоді як параметр max\_tokens встановлює максимальну довжину відповідей моделі. Ми знизили температуру для зменшення варіативності відповідей і скоротили max\_tokens для прискорення процесу генерування.

Зокрема, ми налаштували ключові гіперпараметри, температуру та `max_length`, до 0,5 та 800 відповідно, щоб мінімізувати непередбачуваність виходу LLM та покращити швидкість генерування. Деталі цієї оптимальної конфігурації гіперпараметрів можна знайти в таблиці 4.1.

Таблиця 4.1 - Оптимальна конфігурація гіперпараметрів Vicuna-LLM

Гіперпараметр	Оптимальне значення
Температура	0,5
<code>top_p</code>	0,9
Репетиційний штраф	1,0
Штраф за довжину	1
Максимальні нові токени	400
Максимальна довжина	800

Для завдання генерування описів страв після інженерії запитів ми використовували такий запит: "Перелічіть основні харчові інгредієнти, показані на зображенні. Надайте лаконічну відповідь без додаткових коментарів щодо страви або її оточення." Цей запит був розроблений для отримання прямолінійних описів, зосереджених на інгредієнтах страви.

Після цього ми використали GPT-3.5 для точного та ефективного вилучення інгредієнтів з генерованих описів. Запит, який ми використовували для цього етапу, був таким: "З наданого речення, {resp}, вилучіть харчові інгредієнти, форматуючи вихід як список (напр., [інгредієнт1, інгредієнт2, інгредієнт3])." Приклад для ясності був таким: "Якщо вхідне 'Рецепт включає яблука, апельсини та курку', вихід повинен бути [яблуко, апельсин, курка]", підкреслюючи зосередження на справжніх харчових інгредієнтах.

На другому етапі наших експериментів ми зосередилися на оцінці нашого напівавтоматичного конвеєра за допомогою наборів даних UECFOOD100, UECFOOD256 та нашого власного внутрішнього набору даних. Ми застосували наш конвеєр розпізнавання інгредієнтів до всього

внутрішнього набору даних. Важливо зазначити, що наша модель не була попередньо навчена або налаштована на цьому наборі даних, і для кожного зображення ми генерували одне прогнозування. Цей підхід дав нам точність класифікації 96%96% та точність виявлення 86%86%. Щоб отримати загальний показник точності, ми помножили ці два показники, що дало нам середню точність 83%83%. Цей результат був дуже хорошим.

Після тестування на внутрішньому наборі даних ми продовжили оцінку нашого конвеєра на наборах даних UECFOOD100 та UECFOOD256. Ураховуючи, що порівнянні додатки, такі як "MyDietCam" і "FoodTracker", навчали свої моделі на цих наборах даних і оцінювали їх на приблизно 10% даних, ми вирішили тестувати нашу модель на подібній вибірці 10%, щоб забезпечити послідовність з тестовими середовищами, використаними іншими. Однак наш метод тестування відрізнявся за кількома ключовими аспектами.

По-перше, на відміну від моделей у першому розділі, наша модель не навчалася та не налаштовувалася на цих наборах даних. В результаті класи, які прогнозувала наша модель, іноді відрізнялися від тих, що були в наборах даних, частково тому, що ці набори даних складаються переважно з японських страв, включаючи класи, які не представлені англійською мовою.

Щоб забезпечити справедливе та точне тестове середовище, незважаючи на ці відмінності, ми вжили конкретних заходів. Ми випадково вибрали 10% від кожного класу для створення збалансованого набору даних і уникнути ризику включення послідовних зображень, які можуть представляти один і той же предмет з різних ракурсів. Це випадкове зразкове відбирання було спрямоване на введення стохастичності. Крім того, ураховуючи ймовірність невизначеності виходу великих мовних моделей, ми вирішили обмежити наші прогнози однією спробою на зображення. Цей підхід міг би трохи недооцінити справжню продуктивність нашої моделі, але ми вважали це необхідним. Незважаючи на наші зусилля щодо випадкового

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		69

зразкового відбирання, ручна перевірка виявила, що іноді траплялися випадки, коли одне й те саме зображення було знято з різних ракурсів.

Ураховуючи, що показник точності розпізнавання інгредієнтів використовувався в першому розділі для перевірки продуктивності, ми прийняли той самий показник для нашої оцінки. Цей послідовний показник дозволяє проводити пряме порівняння наших результатів з встановленими в галузі.

### 3.7. Оцінка результатів

Ми використовуватимемо результати мобільних додатків, описаних у першому розділі, як наш базовий рівень. Зокрема, "MyDietCam" досягла точності розпізнавання їжі 88.74% на наборі даних UECFOOD100 та 76.51% на наборі даних UECFOOD256. Водночас "FoodTracker" повідомила про середню точність виявлення 76.35% для UECFOOD100 та 75.05% для UECFOOD256. Ці результати встановлюють базовий рівень для нашого порівняння, що дозволяє нам безпосередньо оцінити, як наш конвеєр порівняно з цими встановленими показниками.

Оцінка моделі на наборах даних, на яких вона не навчалася або не налаштовувалася, становить значну проблему через розбіжності в прогнозованих класах. Ці класи можуть не збігатися або можуть представляти варіації один одного. Для вирішення цієї проблеми ми використовували семантичну схожість для точної оцінки продуктивності. Порівнюючи еталонну істину з прогнозами моделі, ми можемо вважати прогноз правильним, якщо інгредієнти є варіаціями один одного або семантично схожими. Цей підхід був простим для нашого невеликого внутрішнього набору даних, який складався лише з 38 зображень з позначеними інгредієнтами вручну. Ми оцінили модель за допомогою двох основних показників: точності класифікації інгредієнтів (відсоток правильно

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		70

класифікованих інгредієнтів) та точності виявлення інгредієнтів (відсоток правильно виявлених інгредієнтів). Загальний бал, отриманий шляхом множення цих двох показників, склав остаточної бал.

Однак ручна оцінка стає значно складнішою з більшими наборами даних, особливо з такими, що представляють складні кухні, як UECFOOD100 та UECFOOD256, які також не містять позначених інгредієнтів. Основні проблеми включали адаптацію до наборів даних UECFOOD без попереднього навчання, подолання розбіжностей у розмірах наборів даних та обмежені знання в оцінці семантичної схожості та конвертації назв страв в інгредієнти. Ці проблеми надихнули на потребу в комплексній оцінювальній рамці. Ця рамка дозволила б проводити порівняльний аналіз з іншими моделями розпізнавання їжі, зосереджуючись на оцінці продуктивності, порівнянні моделей, диференціацію фокусу розпізнавання та оцінку здатності до узагальнення.

Для вирішення цих проблем ми запропонували використовувати GPT-4, налаштований за допомогою інженерії запитів, для оцінки та порівняння здатності до узагальнення та продуктивності нашої моделі розпізнавання інгредієнтів з іншими. Ця стратегія спрямована на пом'якшення розбіжностей у фокусах наборів даних та фонових навчань шляхом використання здатності GPT-4 до оцінки семантичної схожості та конвертації назв страв в списки інгредієнтів. Це забезпечує точніше порівняння, коли прямі збіги неможливі через відсутність конкретних інгредієнтів або класів їжі.

Ми обрали GPT-4 за його передові здібності до семантичного порівняння та його постійно оновлювану базу знань, що майже гарантує актуальність та точність. Детерміновані результати, забезпечені точною інженерією запитів, пропонують новий підхід до оцінки. Ми сформулювали метод, за яким для кожного завдання класифікації зображень еталонна істина та прогноз надаються в заданому форматі. Це дозволяє GPT-4 оцінювати семантичну схожість і, коли необхідно, порівнювати інгредієнти страв або

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		71

їжі. Запит, який ми використовували, був таким: "У мене є завдання класифікації зображень. Я надам вам еталонну істину та мій прогноз у такому форматі: назва зображення | еталонна істина | прогноз. Для кожного зображення вкажіть "так", якщо вони семантично схожі, інакше "ні". Крім того, якщо еталонна істина стосується страви або їжі, визначте деякі інгредієнти та порівняйте їх з моїм прогнозом. Якщо існує семантична схожість, вважайте прогноз правильним. Розрахуйте відсоток відповідей "так" зі всіх наданих відповідей."

Цей відсоток слугує нашим показником точності розпізнавання інгредієнтів, що узгоджується з показниками оцінки, використаними іншими додатками. Це забезпечує справедливе порівняння між нашим конвеєром і їхніми. Завдяки інтеграції GPT-4 у нашу оцінювальну рамку ми вирішуємо виявлені проблеми та встановлюємо новий стандарт для оцінки моделей розпізнавання їжі та інгредієнтів, що дозволяє ефективно порівнювати продуктивність на будь-якому наборі даних.

У нашому внутрішньому наборі даних ми досягли загальної точності 83%, причому наша точність класифікації склала 96%, а точність виявлення — 86%. Потім ми застосували наш конвеєр до набору даних UECFOOD100, використовуючи раніше описану оцінювальну рамку, і отримали точність розпізнавання інгредієнтів 81%. Для порівняння, "MyDietCam" досягла точності 88.74%, тоді як "FoodTracker" — 76.35%.

Потім ми тестували наш конвеєр на наборі даних UECFOOD256 і досягли точності розпізнавання інгредієнтів 75%.

Для порівняння, "MyDietCam" показала 76.51%, а "FoodTracker" — 75.05%. Ці тести показують, як наш конвеєр конкурує з встановленими стандартами в галузі, як показано в таблиці 3.2. Деякі результати роботи пропонованого сервісу на кожному наборі даних показані в таблицях 3.3 - 3.5.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		72

Таблиця 3.3 - Результати пропонованого сервісу в різних наборах даних порівняно з базовим рівнем

Набір даних	Пропонований сервіс	MyDietCam	FoodTracker
Внутрішній	83%	Н/Д	Н/Д
UECFood100	81%	88.74%	76.36
UECFood256	75%	76.51%	75.05

Нижче наведені приклади страв з наборів даних UECFOOD100 та UECFOOD256, які використовувалися для оцінки продуктивності нашої моделі. Ці складні страви можуть бути настільки складними, що навіть еталонна істина може не охопити всю складність страви.

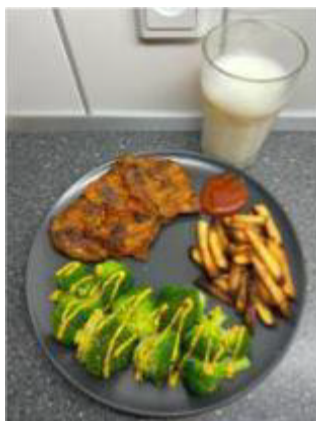


Рисунок 3.10 – Страва із внутрішнього набору даних, що містить броколі, картоплю фрі, молоко та смажену курку



Рисунок 3.11 - Страва із внутрішнього набору даних, що містить локшину, креветки та зелену цибулю

Таблиця 3.4 – Фрагмент результатів роботи сервісу на внутрішньому наборі даних

Страва	Інгредієнти (Еталонна істина)	Прогноз	Точність класифікації інгредієнтів	Точність виявлення інгредієнтів
Рисунок 3.10	смажена курка, картопля фрі, молоко, броколі	картопля фрі, смажена курка, броколі	100%	75%
Рисунок 3.11	локшина, креветки, зелена цибуля	локшина, креветки, зелена цибуля	100%	100%



Рисунок 3.12 - Страва з набору даних UECFOOD100, що містить рамен



Рисунок 3.13 - Страва з набору даних UECFOOD100, що містить рис та суп місо



Рисунок 3.14 - Страва з набору даних UECFOOD256, що містить гоя чанпуру



Рисунок 3.15 - Страва з набору даних UECFOOD256, що містить свинячу відбивну з рисом

Таблиця 3.5 – Фрагмент результатів роботи сервісу на наборі даних UECFOOD100

Страва	Еталонна істина	Прогноз	Результат
Рисунок 3.12	локшина рамен	миска, локшина рамен, варене яйце	правильно
Рисунок 3.13	рис, суп м'ясо	рис, суп м'ясо рис, овочі, м'ясо, соєвий соус, імбир, часник	правильно

Таблиця 3.6 – Фрагмент результатів роботи сервісу на наборі даних UECFOOD256

Страва	Еталонна істина	Прогноз	Результат
Рисунок 3.14	гоя чанпуру	миска, локшина рамен, варене яйце	правильно
Рисунок 3.15	свиняча відбивна на рисі	рис, суп м'ясо рис, овочі, м'ясо, соєвий соус, імбир, часник	неправильно

Отже, запропонований сервіс досяг загальної точності 83% на внутрішньому наборі даних, який включає як точність виявлення, так і класифікації. Якщо зосередитися виключно на класифікації, точність сягає 96%. Однак через невеликий розмір цього набору даних ми не можемо покладатися виключно на ці результати для оцінки загальної продуктивності та здатності до узагальнення нашого конвеєра. На наборі даних UECFOOD100 наша точність склала 81%, перевершивши базовий рівень "FoodTracker" 76.35%, але не досягнувши 88.74% "MyDietCam". Для набору даних UECFOOD256 наша точність впала до 75%, що є зрозумілим через збільшення кількості класів з 100 до 256. Наша продуктивність порівнянна з встановленими базовими рівнями, що є обнадійливим. Однак ці порівняння

потребують обережного розгляду. Наші показники точності отримані з наборів даних, на яких наш сервіс не навчався або не налаштовувався.

Для внутрішнього набору даних неточності переважно виникали через генерування описів VLIP-2. З наборів даних UECFOOD100 та UECFOOD256 помилки виникали як через VLIP-2, так і через оцінки GPT-4. Розбіжності в прогнозованих класах могли призвести до хибно негативних результатів. Незважаючи на те, що деякі прогнози, позначені GPT-4 як неправильні, могли бути правильними при ручній перевірці, ми вирішили не коригувати ці результати, щоб забезпечити послідовність оцінки.

Ці результати свідчать про сильну здатність нашої моделі до узагальнення на різних наборах даних, особливо з варіаціями мов та небаченими класами. Подальше покращення точності можливе шляхом налаштування моделі або збільшення кількості прогнозів. Якби наша модель була налаштована на 80–90% цих наборів даних, як інші моделі, наша точність, ймовірно, значно перевершила б базові рівні. Це підкреслює вражаючу здатність моделей візії-мовлення до узагальнення.

Ще одним ключовим спостереженням є перевага розпізнавання інгредієнтів над розпізнаванням конкретних страв для точної оцінки харчової цінності. Страви можуть сильно різнитися, тому зосередження на інгредієнтах дозволяє ширше розпізнавання з меншою кількістю класів і меншими обчислювальними вимогами порівняно з традиційними моделями CNN для розпізнавання страв. Цей підхід покращує універсальність додатка для різних кухонь, сприяючи точному розпізнаванню інгредієнтів навіть у іноземних стравах.

### Висновки до розділу

У цьому розділі описано процес практичної реалізації сервісу для розрахунку харчової цінності їжі на основі розпізнавання зображень.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		76

Представлено архітектуру системи, яка включає модуль розпізнавання інгредієнтів, модуль взаємодії з користувачем та оцінки порцій, модуль розрахунку нутрієнтної цінності, а також підсистеми налаштування та управління даними.

Модуль взаємодії з користувачем дозволяє зручно вносити додаткову інформацію про порції та отримувати миттєві оцінки ваги їжі на основі автоматичного аналізу вхідних даних. Розроблений модуль розрахунку нутрієнтної цінності здійснює швидкий підрахунок калорійності, вмісту білків, жирів, вуглеводів та інших нутрієнтів на основі розпізнаних інгредієнтів і їх кількісних характеристик.

Окрему увагу приділено налаштуванню сервісу та опису використаних наборів даних для оцінки його ефективності. У рамках експериментального дослідження проведено тестування сервісу, результати якого продемонстрували високу точність розпізнавання інгредієнтів та достовірність розрахунків харчової цінності.

Проведена оцінка результатів підтвердила доцільність обраних методів і моделей. Розроблений сервіс показав конкурентоспроможність серед аналогічних рішень, що відкриває перспективи для його подальшого вдосконалення і застосування у практиці здорового харчування.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
						77
Змн.	Арк.	№ докум.	Підпис	Дата		

## ВИСНОВКИ

В дипломній роботі було представлено розробку та початкову оцінку сервісу для оцінки харчування на основі зображень, зосередженого на автоматичному розпізнаванні інгредієнтів.

Проаналізовано особливості розробки сервісів із використанням моделей глибокого навчання, які відкривають можливості для автоматизованого розпізнавання продуктів за зображеннями, підвищуючи точність і зручність для користувачів. Окремо розглянуто питання навчання моделей, необхідність великих обсягів даних та підбір відповідних архітектур нейронних мереж.

Розроблено модуль розпізнавання інгредієнтів, що використовує візуально-мовні моделі для підвищення точності класифікації продуктів на зображеннях. Це забезпечило можливість інтерпретації складних візуальних сцен за допомогою комбінування зображень та текстових описів

Проведена оцінка продемонструвала високу ефективність розробленого модуля розпізнавання інгредієнтів на тестових наборах даних UECFOOD100, UECFOOD256 та на внутрішньому наборі даних. Незважаючи на те, що моделі не проходили специфічного дотренування або налаштування безпосередньо на цих наборах даних (окрім базового попереднього навчання), система продемонструвала значну здатність до узагальнення (генералізації), що свідчить про її потенційну придатність для аналізу страв різних кухонь світу та варіативної візуальної складності.

Однак, було ідентифіковано суттєве обмеження, пов'язане з повнотою нутрієнтних баз даних та, відповідно, охопленням розпізнаваних інгредієнтів, які можуть бути зіставлені з даними про поживну цінність. Відсутність деяких інгредієнтів або їх специфічних варіацій у базових даних безпосередньо обмежує точність фінальної нутрієнтної оцінки. З метою підвищення точності розрахунків для широкого спектру страв, існує нагальна

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		78

потреба у збільшенні охоплення розпізнаваних інгредієнтів шляхом подальшого дотренування або адаптації моделей розпізнавання на більш різноманітних та репрезентативних наборах даних інгредієнтів.

Важливим аспектом для точної оцінки реального споживання нутрієнтів є врахування випадків, коли користувачі не споживають усю приготовану або подану порцію страви. Для вирішення цієї проблеми планується надати користувачеві можливість сфотографувати залишки страви після їжі. У такому випадку фактичне споживання нутрієнтів буде розраховуватися як різниця між оціненою початковою порцією та оціненою порцією залишків.

					БР.ПЗ – 08.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		79

## ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Dounavi K and Tsoumani O. Mobile health applications in weight management: a systematic literature review. Am J Prev Med 2019; 56: 894–903.
2. Lose It! plans to take weight loss app international in 2015 | MobiHealthNews. - <https://www.mobihealthnews.com/41377/loseit-plans-to-take-weight-loss-app-international-in-2015>
3. Oppermann A. Artificial Intelligence vs. Machine Learning vs. Deep Learning: What's the difference?, <https://builtin.com/artificial-intelligence/ai-vs-machine-learning>.
4. Höchsmann C and Martin CK. Review of the validity and feasibility of image-assisted methods for dietary assessment. Int J Obes 2020; 44: 2358–2371.
5. MyFitnessPal Review | PCMag. – <https://www.pcmag.com/reviews/myfitnesspal>
6. Sahoo D, Hao W, Ke S, et al. FoodAI: Food Image Recognition via Deep Learning for Smart Food Logging. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, AK, USA: Association for Computing Machinery, 2019, p. 2260–2268.
7. Renehan AG, Tyson M, Egger M, et al. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet 2008; 371: 569–578.
8. Our New Look: The Details Behind Cronometer's Makeover. – <https://cronometer.com/blog/our-new-look/>
9. Azar KMJ, Lesser LI, Laing BY, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013; 45: 583–589.

					БР.ІІЗ – 08.00.00.000 ІІЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		80

10. Ferrara G, Kim J, Lin S, et al. A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. *JMIR Mhealth Uhealth* 2019.
11. Alshurafa N, Lin AW, Zhu F, et al. Counting bites with bits: expert workshop addressing calorie and macronutrient intake monitoring. *J Med Internet Res* 2019
12. Every day big data statistics 2.5 quintillion bytes of data created daily. <http://www.vcloudnews.com/every-day-big-data-statistics-2-5-quintillion-bytes-of-data-created-daily/>, 2015.
13. K. Aizawa and M. Ogawa. Foodlog: Multimedia tool for healthcare applications. *IEEE MultiMedia*, 22(2):4–8, 2015.
14. Wang, X., Kumar, D., Thome, N., Cord, M., and Precioso, F. (July, January 29). Recipe recognition with large multimodal food dataset. *Proceedings of the 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Turin, Italy.
15. Fang, S., Liu, C., Zhu, F., Delp, E.J., and Boushey, C.J. (2015, January 14–16). Single-View Food Portion Estimation Based on Geometric Models. *Proceedings of the 2015 IEEE International Symposium on Multimedia (ISM)*, Miami, FL, USA. DOI: 10.1109/ISM.2015.67
16. Herranz, L., Xu, R., and Jiang, S. (July, January 29). A probabilistic model for food image recognition in restaurants. *Proceedings of the 2015 IEEE International Conference on Multimedia and Expo (ICME)*, Turin, Italy. DOI: 10.1109/ICME.2015.7177464
17. V. Bettadapura, E. Thomaz, A. Parnami, G. D. Abowd, and I. Essa. Leveraging context to support automated food recognition in restaurants. In *Applications of Computer Vision (WACV)*, 2015 IEEE Winter Conference on, pages 580–587. IEEE, 2015

18. Meltzer ST Sources of error and bias in the assessment of dietary intake: 24-hour recalls and food records. University of Cape Town, South Africa, 1997.
19. Blanchard CM, Chin MK, Gilhooly CH, et al. Evaluation of PIQNIQ, a novel mobile application for capturing dietary intake. J Nutr 2021; 151: 1347–1356.
20. Martinon P, Fraticelli L, Giboreau A, et al. Nutrition as a key modifiable factor for periodontitis and main chronic diseases. J Clin Med 2021; 10: 97.
21. Institute for Public Health (IPH). National Health and Morbidity Survey (NHMS) 2019 2019.
22. FoodTracker: An AI-powered food detection mobile application. – <https://techxplore.com/news/2019-09-foodtracker-ai-powered-food-mobile-application.html>
23. Hamman RF, Wing RR, Edelstein SL, et al. Effect of weight loss with lifestyle intervention on risk of diabetes. Diabetes Care 2006; 29: 2102–2107.
24. Kannel WB, d’Agostino R and Cobb JL. Effect of weight on cardiovascular disease. Am J Clin Nutr 1996; 63: 419S–422S.
25. R.A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A.J. Davison, P. Kohli, J. Shotton, S. Hodges, A. Fitzgibbon, “KinectFusion: real-time dense surface mapping and tracking,” in: IEEE International Symposium on Mixed and Augmented Reality, Basel, Switzerland, October 2011, pp. 127-136.
26. OpenSURF [Online]: <https://code.google.com/p/opensurf/> [38] M. Anthimopoulos, J. Dehais, S. Shevchik, R.H. Botwey, D. Duke, P. Diem, S. Mougiakakou, “Computer vision-based carbohydrate estimation for type 1 diabetic patients using smartphones,” in J. Diabetes Sci Technol Apr. 2015.
27. D. Rhyner, H. Loher, J. Dehais, M. Anthimopoulos, S. Shevchik, R.H. Botwey, D. Duke, C. Stettler, P. Diem, and S. Mougiakakou, “Carbohydrate Estimation by a Mobile Phone-Based System Versus Self-Estimations of

- Individuals With Type 1 Diabetes Mellitus: A Comparative Study, “ in J. Medical Internet Research, vol. 18, no. 5, 2016.
- 28.L. Bally, J. Dehais, C.T. Nakas, M. Anthimopoulos, M. Laimer, D. Rhyner, G. Rosenberg et al. "Carbohydrate Estimation Supported by the GoCARB system in Individuals With Type 1 Diabetes: A Randomized Prospective Pilot Study." in Diabetes Care, 2016
- 29.Sharp D and Allman-Farinelli M. the feasibility and validity of mobile phones to assess dietary intake. Nutrition 2014; 30: 1257–1266. DOI: 10.1016/j.nut.2014.02.020
30. Eldridge AL, Piernas C, Illner A-K, et al. Evaluation of new technology-based tools for dietary intake assessment-an ILSI Europe dietary intake and exposure task force evaluation. Nutrients 2018; 11: 55.
31. H-S Chen, W. Jia, Y. Yue, Z. Li, Y.-N. Sun, J.D. Fernstrom, M. Sun, “Model-based measurement of food portion size for image-based dietary assessment using 3D/2D registration,” in Measurement Science and Technology, vol. 24, pp. 1-11, 2013

					БР.ІІЗ – 08.00.00.000 ІІЗ	Арк.
						83
Змн.	Арк.	№ докум.	Підпис	Дата		

## БІБЛІОГРАФІЧНА ДОВІДКА

**Тема дипломної роботи:** “Програмний сервіс для розрахунку харчової цінності їжі”

Обсяг пояснювальної записки: 83 аркуші.

Дата закінчення роботи: 11 червня 2025 р.

Підпис студента \_\_\_\_\_