

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
ФАКУЛЬТЕТ БІОМЕДИЧНОЇ ІНЖЕНЕРІЇ

(повна назва інституту/факультету)

кафедра БІОМЕДИЧНОЇ КІБЕРНЕТИКИ

(повна назва кафедри)

«На правах рукопису»

УДК 004.93'12 + 616.8-009.11

«До захисту допущено»

Завідувач кафедри БМК

_____ Євген. НАСТЕНКО
(підпис) (ініціали, ПРІЗВИЩЕ)

“ _____ ” грудня _____ 2021р.

Магістерська дисертація

на здобуття ступеня магістра

за освітньо-професійною програмою «Комп'ютерні технології в біології та медицині»

зі спеціальності 122 «Комп'ютерні науки»

на тему: Сценарії використання технології відстежування обличчя на основі методу гістограм орієнтованих градієнтів

Виконав (-ла): студент (-ка) II курсу, групи БС-01мп

КОВАЛЕНКО МАКСИМ ВЛАДИСЛАВОВИЧ

(прізвище, ім'я, по батькові)

(підпис)

Науковий керівник:

к.ф.-м.н., доц. каф. БМІ Соломін Андрій Вячеславович

(посада, науковий ступінь, вчене звання, прізвище, ім'я, по ініціали)

(підпис)

Консультант з розділів магістерської дисертації:

(назва розділу) (посада, вчене звання, науковий ступінь, прізвище, ініціали)

(підпис)

Рецензент: _____

(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Засвідчую, що у цій магістерській дисертації немає запозичень з праць інших авторів без відповідних посилань.

Студент (-ка) _____
(підпис)

Київ – 2021 року

Нормо контролер

Галина КОРНІЄНКО

Коваленко М. В. Сценарії використання технології відстежування обличчя на основі методу гістограм орієнтованих градієнтів : магістерська дис. : 122 Комп'ютері науки / Коваленко Максим Владиславович. – Київ, 2021. – 100 с.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет біомедичної інженерії
Кафедра біомедичної кібернетики

Рівень вищої освіти – другий (магістерський)

Спеціальність – 122 «Комп'ютерні науки»

Освітньо-професійна програма «Комп'ютерні технології в біології та медицині»

ЗАТВЕРДЖУЮ

Завідувач кафедри БМК

_____ Євген НАСТЕНКО

« 31 » серпня 2021 р.

ЗАВДАННЯ
на магістерську дисертацію студенту

КОВАЛЕНКО МАКСИМ ВЛАДИСЛАВОВИЧ

(прізвище, ім'я, по батькові)

1. Тема дисертації Сценарії використання технології відстежування обличчя на основі методу гістограм орієнтованих градієнтів

науковий керівник дисертації

Соломін Андрій В'ячеславович, к.ф.-м.н., доц. каф. БМІ

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «09» листопада 2021 р. №3703-с

2. Термін подання студентом дисертації 01-04 грудня 2021 року

3. Об'єкт дослідження: Системи предиктивного введення тексту для допомоги паралізованим пацієнтам

4. Вихідні дані: результати дипломної роботи бакалавра; додаток для спілкування з пацієнтами з технологією предиктивного тексту

5. Перелік завдань, які потрібно розробити: 1) Ознайомитись із оптимальними алгоритмами відстежування погляду та розпізнавання образів і методами їх реалізації, 2) Розробити систему предиктивного введення тексту, 3) Реалізувати програмний додаток, що надає пацієнту можливість набирати текстові повідомлення за допомогою руху його очей

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: 26 рисунків, 12 таблиць, презентація на 18 слайдів.

7. Орієнтовний перелік публікацій: 1. Kovalenko M.V. Predictive text typing system for communication with paralyzed patients / Kovalenko M.V., Solomin A.V. // Modern engineering and innovative technologies. – Germany, 2021.– Issue №17 2. Kovalenko M.V. Improving the virtual keyboard for communication with paralyzed patients / Kovalenko M.V., Solomin A.V. // Technique and technology of the future ‘2021’. – Karlsruhe, Germany, 2021

8. Консультанти розділів дисертації*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Магістерської дисертації	-		

9. Дата видачі завдання 31 серпня 2021 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Отримати завдання на МД	31 серпня 2021р.	
2	Практика	01.09.2021р. по 26.10.2021р.	
3	Опис результатів практики в розділах МД	1-8 листопада 2021р.	
4	Проектування системи предиктивного введення тексту	9-12 листопада 2021р..	
5	Розробка системи	13-18 листопада 2021р	
6	Тестування системи	18-23 листопада 2021р	
7	Оформлення розділу «Стартап проект за темою дисертації»	24-29 листопада 2021р	
8	Подання в електронному вигляді МД та анотації до неї на перевірку номоконтролера та UNICHECK .	1-3 грудня 2021р	
9	Предзахист МД та допуск до захисту дисертації	4-6 грудня 2021р.	
10	Подання МД рецензенту. Отримання рецензії.	7-15 грудня 2021р.	
11	Подання пакету документів по МД до захисту в ЕК ¹	15-17 грудня 2021р.	
12	Захист МД в ЕК	20-22 грудня 2021р.	

Студент

(підпис)

Максим КОВАЛЕНКО

(ім'я, ПРИЗВИЩЕ)

Науковий керівник

(підпис)

Андрій СОЛОМІН

(ім'я, ПРИЗВИЩЕ)

* Якщо визначені консультанти. Консультантом не може бути зазначено наукового керівника магістерської дисертації.

¹ не пізніше ніж за один тиждень до затвердженої дати захисту МД в ЕК

Реферат

Магістерська дисертація за темою «Сценарії використання технології відстежування обличчя на основі методу гістограм орієнтованих градієнтів» виконана студентом кафедри біомедичної кібернетики ФБМІ Коваленком Максимом Владиславовичем зі спеціальності 122 «Комп'ютерні науки» за освітньо-професійною програмою «Комп'ютерні технології в біології та медицині» та складається зі: вступу; 4 розділів («Аналіз джерел з предметної області», «Проектування складових частин системи», «Розробка комунікаційної системи», «Реалізація програмного додатку з функцією предиктивного набору тексту»), розділу зі стартап проекту, висновків до кожного з цих розділів; загальних висновків; списку використаних джерел, який налічує 70 джерело та додатки. Загальний обсяг роботи 100 сторінок.

Актуальність теми. Актуальність пов'язана з можливістю створення для паралізованих пацієнтів інформаційного каналу спілкування з оточенням шляхом розробки програмного додатку, що не вимагає спеціального дорогого обладнання та інвазивних процедур.

Мета дослідження. Розробка універсального програмного додатку-помічника для хворих із різними ступенями паралічу тіла, з використанням технології відстежування погляду та предиктивного введення тексту.

Об'єкт дослідження. Системи предиктивного введення тексту.

Предмет дослідження. Використання систем предиктивного введення тексту для допомоги у спілкуванні із паралізованими пацієнтами.

Методи дослідження. Комп'ютерний зір, рекурентні нейронні мережі, метод гістограм орієнтованих градієнтів, n-грами.

Ключові слова. паралізований пацієнт, канал спілкування, предиктивне введення тексту, НОГ, рекурентні нейронні мережі

Публікації. За результатами виконаної роботи були опубліковані 2 наукові публікації:

1. **Kovalenko M.V.** Predictive text typing system for communication with paralyzed patients / Kovalenko M.V., Solomin A.V. // Modern engineering and innovative technologies. – Germany, 2021.– Issue №17, Part 3, p. 19-22.

2. **Kovalenko M.V.** Improving the virtual keyboard for communication with paralyzed patients // Kovalenko M.V., Solomin A.V. // Technique and technology of the future '2021'. – Karlsruhe, Germany, 2021, p. 14-16.

Abstract

Master's dissertation on "Scenarios for the use of facial tracking technology based on the method of histograms of oriented gradients" performed by *Maksym Kovalenko*, a student of the Department of Biomedical Cybernetics FBMI by specialty 122 "*Computer Science*" in the educational and professional program "*Computer Technology in Biology and Medicine*" and consists of an introduction, 4 sections ("Analysis of the subject area", "Design of system components", "Development of communication system", "Implementation of a software application with the function of predictive text"), section with startup project, conclusions to each chapter, general conclusions and list of references that includes 70 points. The paper amounts to 100 pages.

Topic's relevance. The relevance is related to the possibility of creating an information channel for paralyzed patients to communicate with the environment by developing a software application that does not require special expensive equipment and invasive procedures.

Research objective. Development of a universal software application-assistant for patients with various degrees of paralysis of the body, using the technology of eye tracking and predictive text input.

Object of study. Predictive text input systems.

Subject of study. Using predictive text input systems to help communicate with paralyzed patients.

Research methods. Computer vision, recurrent neural networks, method of oriented gradients, n-grams.

Keywords. Paralyzed patient, communication channel, predictive text, HOG, recurrent neural network

Publications. According to the results of the work 2 scientific papers were published:

1. **Kovalenko M.V.** Predictive text typing system for communication with paralyzed patients / Kovalenko M.V., Solomin A.V. // Modern engineering and innovative technologies. – Germany, 2021.– Issue №17, Part 3, p. 19-22.

2. **Kovalenko M.V.** Improving the virtual keyboard for communication with paralyzed patients // Kovalenko M.V., Solomin A.V. // Technique and technology of the future '2021'. – Karlsruhe, Germany, 2021, p. 14-16.

Зміст

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ	12
ВСТУП.....	13
РОЗДІЛ 1 АНАЛІЗ ДЖЕРЕЛ З ПРЕДМЕТНОЇ ОБЛАСТІ	17
1.1 Розуміння ПРОБЛЕМАТИКИ	17
1.1.1 Що таке параліч?.....	17
1.1.2 Які існують види паралічу?.....	17
1.1.3 Що викликає параліч?.....	18
1.2 ТЕХНОЛОГІЯ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ	19
1.2.1 Історія розвитку технології.....	19
1.2.2 Методи розпізнавання обличчя.....	21
1.2.3 Традиційні методи.....	22
1.2.4 Переваги та недоліки технології.....	24
1.3 ТЕХНОЛОГІЯ ВІДСТЕЖУВАННЯ ПОГЛЯДУ	24
1.3.1 Типи трекерів.....	24
1.3.1.1. Спеціальні контактні лінзи.....	25
1.3.1.2. Оптичне відстежування.....	25
1.3.1.3. Вимірювання електричного потенціалу.....	26
1.3.2. Застосування технології	27
1.4 ПРОГНОЗУВАННЯ ТЕКСТУ.....	28
1.4.1 Зародження технології	29
1.4.2 Використання сучасних алгоритмів та нейронних мереж.....	30
Висновки до розділу 1	31
РОЗДІЛ 2 ПРОЕКТУВАННЯ СКЛАДОВИХ ЧАСТИН СИСТЕМИ.....	32
2.1. ТЕОРЕТИЧНА ОСНОВА СИСТЕМИ.....	32
2.2 ПОВЕДІНКА ТА ВИМОГИ СИСТЕМ ПРОГНОЗУВАННЯ ТЕКСТУ.....	33
2.2.1 Обмеження людей із вадами руху.....	33
2.2.2 Вимоги до систем передбачення тексту та їх користувачів	34

2.2.3 Поведінка систем прогнозування тексту.....	35
2.3 АНАЛІЗ СИСТЕМ ПРОГНОЗУВАННЯ.....	36
2.3.1 Розмір текстового блоку.....	37
2.3.2 Підхід довготривалої короткочасної пам'яті (LSTM).....	40
2.4 ПОВЕДІНКА ТА ВИМОГИ СИСТЕМ ВІДСТЕЖУВАННЯ ПОГЛЯДУ	42
2.4.1 Фіксації.....	43
2.4.2 Розмір зіниць.....	44
2.4.3 Якість даних при вимірюванні.....	44
2.4.4 Висновки щодо якості даних.....	46
2.5 ВИЗНАЧЕННЯ ВИМОГ ДО СИСТЕМИ.....	47
Висновки до розділу 2	48
РОЗДІЛ 3 РОЗРОБКА КОМУНІКАЦІЙНОЇ СИСТЕМИ	50
3.1 ОБРОБКА ПРИРОДНОЇ МОВИ (NLP) З N-ГРАМАМИ І МАРКОВСЬКИМИ МОДЕЛЯМИ	50
3.1.1 Попередня обробка тексту.....	51
3.1.2 Токенізація	52
3.1.3 Вилучення стоп слів.....	53
3.2 ВИМОГИ ДО РОЗМІРУ СЛОВНИКА	53
3.3 ВИМОГИ ДО КОРИСТУВАЦЬКОГО ІНТЕРФЕЙСУ	54
3.4 ПРОЕКТУВАННЯ ПРОГРАМНОГО ПРОДУКТУ	57
3.4.1. Декомпозиція 1-ого рівня.....	58
3.4.2. Декомпозиція 2-ого рівня.....	60
3.4.3. Діаграма дерева вузлів.....	63
3.4.4. Діаграма послідовності.....	64
3.4.4. Діаграма станів.....	65
Висновки до розділу 3	66
РОЗДІЛ 4 РЕАЛІЗАЦІЯ ПРОГРАМНОГО ДОДАТКУ З ФУНКЦІЄЮ ПРЕДИКТИВНОГО НАБОРУ ТЕКСТУ.....	67
4.1 ТЕОРЕТИЧНЕ ОБґРУНТУВАННЯ РОБОТИ ПРОГРАМИ	67

4.2 ЕТАПИ РОБОТИ ДОДАТКУ	70
4.2 ВЗАЄМОДІЯ З ПРОГРАМНИМ ДОДАТКОМ	74
Висновки до розділу 4	78
РОЗДІЛ 5 СТАРТАП АНАЛІЗ ПРОЕКТУ	80
5.1 РЕЗЮМЕ ПРОЕКТУ	80
5.2 ОРГАНІЗАЦІЯ ПРОЕКТУ	81
5.3 КАНВА БІЗНЕС-МОДЕЛІ ПРОЕКТУ	81
5.4 КЛЮЧОВІ ВИДИ ДІЯЛЬНОСТІ ПРОЕКТУ	82
5.5 ЦІННІСНІ ПРОПОЗИЦІЇ ТА СПОЖИВАЧІ	84
5.6 ОБГРУНТУВАННЯ РЕСУРСІВ ТА ВИТРАТ ПРОЕКТУ	87
Висновки до розділу 5	90
ЗАГАЛЬНІ ВИСНОВКИ.....	91
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	93

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ

HOG – (від Histogram of oriented gradients) гістограма орієнтованих градієнтів

NLP – (від Natural Language Processing) обробка природньої мови

LSTM – (від Long Short Term Memory) довга короткочасна пам'ять

ВСТУП

Актуальність роботи пов'язана з проблемою створення каналу спілкування для хворих з втратою мови при обтяженні обмеженою рухливістю.

Існує безліч факторів які можуть обмежити здатність людини до вільного спілкування. Більшість з них зазвичай пов'язані із різноманітними хворобами, що унеможлиблюють процес спілкування хворого із навколишнім світом. Серед них такі хвороби, як інсульт, бічний аміотрофічний склероз (БАС) або інші травми та неврологічні порушення, які можуть викликати параліч, пошкоджуючи нервові шляхи, які з'єднують мозок з іншими частинами тіла

У цій роботі розглянуті сценарії використання сучасної технології відстежування погляду з метою полегшення процесу комунікації для людей з обмеженими фізичними можливостями. Зокрема для пацієнтів з частковим, або повним паралічем тіла, що не можуть спілкуватися з оточуючим світом за допомогою голосу, рухів тіла та набору тексту на звичайних клавіатурах комп'ютерів та смартфонів.

Відстежування обличчя та напряму погляду є одними з небагатьох варіантів що лишаються, коли мова заходить про способи повернути паралізованим хворим можливість виражати свої думки. Раніше для цього використовувалися таблички із літерами, на які вказував лікар, очікуючи отримати від пацієнта сигнал для підтвердження вибору. Сьогодні ми маємо змогу автоматизувати цей процес, а також набагато пришвидшити його за допомогою рекурентних нейронних мереж та технології предиктивного набору тексту. У поєднанні ці технології мають змогу максимально спростити для паралізованих хворих спілкування з оточуючими та повернути їм свободу вираження своєї думок.

Мета і завдання дослідження

Попереднім дослідженням автора була розробка алгоритму відстежування погляду для комунікації з паралізованими пацієнтами. Результати дослідження висвітлені у дипломній роботі бакалавра, яка є комплексною, і розроблялась від керівництвом доцента кафедри Біомедичної Інженерії Соломіна Андрія В'ячеславовича (нинішній науковий керівник).

Мета даної роботи полягає у проектуванні та розробці системи предиктивного набору тексту у поєднанні із віртуальною клавіатурою та системою відстежування погляду для полегшення процесу спілкування із паралізованими хворими. Система не повинна потребувати дорогого обладнання для її встановлення, а повинна запускатися на будь-якому середньостатистичному персональному комп'ютері або ноутбучі із підключеною веб-камерою.

Досягнення поставленої мети передбачає вирішення таких *завдань*:

1. Ознайомлення із оптимальними алгоритмами відстежування погляду та розпізнавання образів і методами їх реалізації.
2. Розробка системи предиктивного набору тексту для полегшення спілкування з паралізованими пацієнтами.
3. Реалізація програмного додатку, що надає пацієнту можливість набирати текстові повідомлення на віртуальній клавіатурі за допомогою руху його очей.

Об'єктом дослідження є система предиктивного введення тексту.

Предметом дослідження є використання систем предиктивного введення тексту для допомоги у спілкуванні із паралізованими пацієнтами.

Методи дослідження. У роботі застосовано алгоритми комп'ютерного зору та розпізнавання образів з використанням методу гістограм орієнтованих градієнтів для визначення координат ключових точок обличчя пацієнта на вхідному зображенні з веб-камери персонального комп'ютера, та для подальшого визначення напрямку погляду пацієнта та моменту моргання очима. Ланцюги Маркова та метод n-грам для навчання системи

предиктивного введення тексту та обробки вхідного тексту від пацієнта, з метою генерації можливих варіантів продовження речення під час набору. Мова програмування Python та сторонні бібліотеки використовувалися для основної частини реалізації програмного додатку та його графічного інтерфейсу

Практичне значення одержаних результатів

Запропонований алгоритм та реалізований програмний засіб здатний суттєво покращити рівень життя великої кількості паралізованих пацієнтів, які знаходяться в інформаційній ізоляції.

Особистий внесок здобувача

В магістерській дисертації були використані ідеї та розробки, які були розроблені автором в попередній (бакалаврській) роботі [1]. Розроблено систему предиктивного введення тексту та інтегровано із системою відстежування погляду (результатом бакалаврської роботи). Були використані ідеї щодо найкращого методу та алгоритму передбачення тексту для даного програмного додатку, що були опубліковані у роботах спільно з науковим керівником.

Публікації

За результатами виконаної роботи були опубліковані наступні наукові статті:

1. **Kovalenko M.V.** Predictive text typing system for communication with paralyzed patients / Kovalenko M.V., Solomin A.V. // Modern engineering and innovative technologies. – Germany, 2021.– Issue №17, Part 3, p. 19-22.

2. **Kovalenko M.V.** Improving the virtual keyboard for communication with paralyzed patients // Kovalenko M.V., Solomin A.V. // Technique and technology of the future ‘2021’. – Karlsruhe, Germany, 2021, p. 14-16.

Структура дисертації

Магістерська дисертація за темою «Сценарії використання технології відстежування обличчя на основі методу гістограм орієнтованих градієнтів» виконана студентом кафедри біомедичної кібернетики ФБМІ Коваленком

Максимом Владиславовичем зі спеціальності 122 *«Комп'ютерні науки»* за освітньо-професійною програмою *«Комп'ютерні технології в біології та медицині»* та складається зі: вступу; 4 розділів («Аналіз джерел з предметної області», «Проектування складових частин системи», «Розробка комунікаційної системи», «Реалізація програмного додатку з функцією предиктивного набору тексту»), розділу зі стартап проєкту, висновків до кожного з цих розділів; загальних висновків; списку використаних джерел, який налічує 70 джерело та додатки. Загальний обсяг роботи 100 сторінок. В дисертації представлено 26 рисунків і 12 таблиць.

РОЗДІЛ 1

АНАЛІЗ ДЖЕРЕЛ З ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Розуміння проблематики

1.1.1 Що таке параліч?

Параліч – це втрата м'язової функції частини тіла. Він може бути локалізованим або генералізованим, частковим або повним, тимчасовим або постійним. Параліч може вразити будь-яку частину тіла в будь-який момент життя людини.

План лікування та перспективи стану залежатимуть від основної причини паралічу, а також від наявних симптомів. Технологічні інновації та терапевтичні втручання можуть допомогти зберегти незалежність і якість життя пацієнтів.

1.1.2 Які існують види паралічу?

Лікарі можуть класифікувати параліч різними способами:

Розташування

Локалізований параліч вражає лише одну частину тіла, наприклад, обличчя або руку.

Генералізований параліч – це група станів, які вражають кілька частин тіла. До типів належать:

- моноплегія, яка вражає тільки одну руку або ногу
- геміплегія, яка вражає одну руку та одну ногу з одного боку тіла
- параплегія, яка вражає обидві ваші ноги
- квадриплегія або тетраплегія, яка вражає обидві руки та обидві ноги

Тяжкість

У людини з частковим паралічем є певний контроль над м'язами в уражених частинах тіла. Якщо у людини повний параліч, вона не зможе контролювати м'язи в уражених областях.

Тривалість

Параліч може бути тимчасовим. Наприклад, параліч Белла - це стан, який може викликати тимчасовий параліч обличчя у людини. Удари також можуть тимчасово паралізувати одну із сторін тіла. З часом і лікуванням можна відновити відчуття та координацію часткового або повністю .

В інших випадках параліч може бути постійним.

Мляві або спастичні

Млявий параліч призводить до того, що м'язи скорочуються і стають в'ялими. Це призводить до м'язової слабкості. Спастичний параліч включає напружені і тверді м'язи. Це може призвести до неконтрольованого посмикування м'язів або спазмів [24].

1.1.3 Що викликає параліч?

Може статися так, що людина повністю обмежена у спілкуванні, не в змозі виразити себе ні звичайною мовою, ні мовою жестів. Це пов'язано з втратою рухової активності - частковою (гіпокінезія) або повною (акінезія) [4]. У першому випадку рухливість втрачається через різні захворювання нервової системи та посттравматичні стани головного та спинного мозку, а також інсульти. У другому випадку втрата активності є наслідком складних психічних розладів і паралічів [44].

Параліч означає порушення рухливості або повну його втрату в частинах тіла, модифікацію та прогресуючу втрату рухових функцій через пошкодження нервової системи. Причини порушень руху різноманітні: дегенеративні, спадкові або вроджені вади розвитку центральної нервової системи; травми при народженні (ураження плечового сплетення, церебральний параліч); інфекційні захворювання - поліомієліт, менінгіт, туберкульоз, сифіліс, вірусний енцефаліт [35].

За даними Фонду Крістофера і Дани Рівів, інсульт є основною причиною паралічу в Сполучених Штатах. Він відповідальний за майже 30 відсотків випадків. Пошкодження спинного мозку становлять приблизно 23 відсотки випадків. Розсіяний склероз викликає приблизно 17 відсотків випадків [53].

Інсульт посідає 2-ге місце у переліку головних причин смерті і є провідною причиною інвалідності в Україні. Це захворювання може спричинити параліч кінцівок, втрату зору, порушення ходи, мови та психічних функцій. Крім того, інсульт різко збільшує ризик розвитку деменції. Особливу тривогу викликає зростання частоти інсульту серед людей працездатного віку: близько третини всіх інсультів виникає у віці до 60 років, і лише невелика частка з тих, хто пережив інсульт, повертається до роботи [2].

Параліч може бути викликаний причинами органічного характеру - порушенням обміну речовин, інфекціями та інтоксикаціями, раком, ураженням судин, неправильним харчуванням, токсичними причинами - дефіцитом вітаміну В1, отруєнням важкими металами, алкогольним невритом тощо. Акінезія може бути наслідком важких травм і переломів при ураженні рухових центрів та їх шляхів. Зниження рухової активності також може бути спричинене розсіяним склерозом або захворюваннями психогенного характеру або будь-яким іншим ураженням центральної нервової системи [35].

Поліомієліт, геморагічний та ішемічний інсульт, синдром Гієна-Барре та інші ураження змушують людину втрачати здатність рухатися та адекватно реагувати на зовнішні подразники. Пацієнт втрачає рухову активність і мовлення, але залишається повністю свідомим. У цьому випадку ми маємо справу з синдромом ізоляції чи диференціювання [43].

1.2 Технологія розпізнавання обличчя

1.2.1 Історія розвитку технології

Автоматичне розпізнавання обличчя було вперше запроваджено в 1960-х роках. Вуді Бледсо, Хелен Чан Вольф і Чарльз Біссон працювали над використанням комп'ютера для розпізнавання людських облич. Їхній ранній проект розпізнавання обличчя отримав назву «людина-машина», тому що координати рис обличчя на фотографії повинні були бути встановлені людиною, перш ніж вони могли бути використані комп'ютером для розпізнавання. На графічному планшеті людина повинна була точно

визначити координати таких рис обличчя, як центр зіниць, внутрішній і зовнішній куточок очей, а також лінію росту волосся. Координати були використані для обчислення 20 відстаней, включаючи ширину рота та очей. Таким чином людина може обробляти близько 40 зображень на годину і створити базу даних обчислених відстаней. Потім комп'ютер автоматично порівнює відстані для кожної фотографії, обчислює різницю між відстанями і повертає закриті записи як можливу відповідність [42].

У 1970 році Такео Канаде публічно продемонстрував систему зіставлення обличчя, яка знаходила анатомічні особливості, такі як підборіддя, і розраховувала співвідношення відстаней між рисами обличчя без втручання людини. Пізніші тести показали, що система не завжди може надійно ідентифікувати риси обличчя. Тим не менш, інтерес до цієї теми зріс, і в 1977 році Канаде опублікувала першу детальну книгу про технологію розпізнавання облич [61].

У 1993 році Агентство перспективних оборонних дослідницьких проектів (DARPA) і Дослідницька лабораторія армії (ARL) заснували програму технології розпізнавання облич FERET для розробки «можливостей автоматичного розпізнавання обличчя», які можна було б використовувати в продуктивному реальному середовищі «для забезпечення безпеки, розвідки та працівників правоохоронних органів при виконанні покладених на них обов'язків». Системи розпізнавання облич, які були випробувані в дослідницьких лабораторіях, були оцінені, і тести FERET показали, що, хоча ефективність існуючих автоматизованих систем розпізнавання обличчя різна, кілька існуючих методів можна було б ефективно використовувати для розпізнавання облич на нерухомих зображеннях, зроблених у контрольованому середовищі [18].

Тести FERET породили три американські компанії, які продали автоматизовані системи розпізнавання обличчя. Vision Corporation і Miros Inc були засновані в 1994 році дослідниками, які використовували результати тестів FERET як точки продажу. Viisage Technology була заснована

підрядником із захисту ідентифікаційних карт у 1996 році для комерційного використання прав на алгоритм розпізнавання обличчя, розроблений Алексом Пентлендом з Массачусетського технологічного інституту [19].

Збільшення чисельності в'язнів США в 1990-х роках спонукало штати США створити пов'язані й автоматизовані системи ідентифікації, які включали цифрові біометричні бази даних, у деяких випадках це включало розпізнавання обличчя. У 1999 році штат Мінесота включив систему розпізнавання обличчя FaceIT від Visionics в систему бронювання знімків, яка дозволила поліції, суддям і працівникам суду відстежувати злочинців у всьому штаті [17].

1.2.2 Методи розпізнавання обличчя

Хоча люди і можуть розпізнавати обличчя без особливих зусиль, [34] розпізнавання обличчя - це складна проблема розпізнавання образів у комп'ютерному обчисленні. Системи розпізнавання обличчя намагаються ідентифікувати людське обличчя, що є тривимірним за зміною зовнішнього вигляду, освітленням та виразом обличчя, заснованим на його двовимірному зображенні.

Для виконання цього обчислювального завдання, системи розпізнавання обличчя виконують чотири кроки. Спочатку розпізнавання обличчя використовується для сегментації обличчя від фону зображення. На другому етапі зображення сегментованого обличчя вирівнюється для урахування позиції обличчя, розміру зображення та фотографічних властивостей, таких як освітлення та відтінки сірого. Метою процесу вирівнювання на третьому кроці є забезпечення точної локалізації рис обличчя. Особливості, такі як очі, ніс і рот, визначаються і вимірюються на зображенні, щоб представити обличчя у вигляді координат. На четвертому етапі тільки що встановлений вектор ключових точок обличчя співвідноситься з обличчями із бази даних [34]. На рис.1.1 зображено приклад використання технології на групі людей.

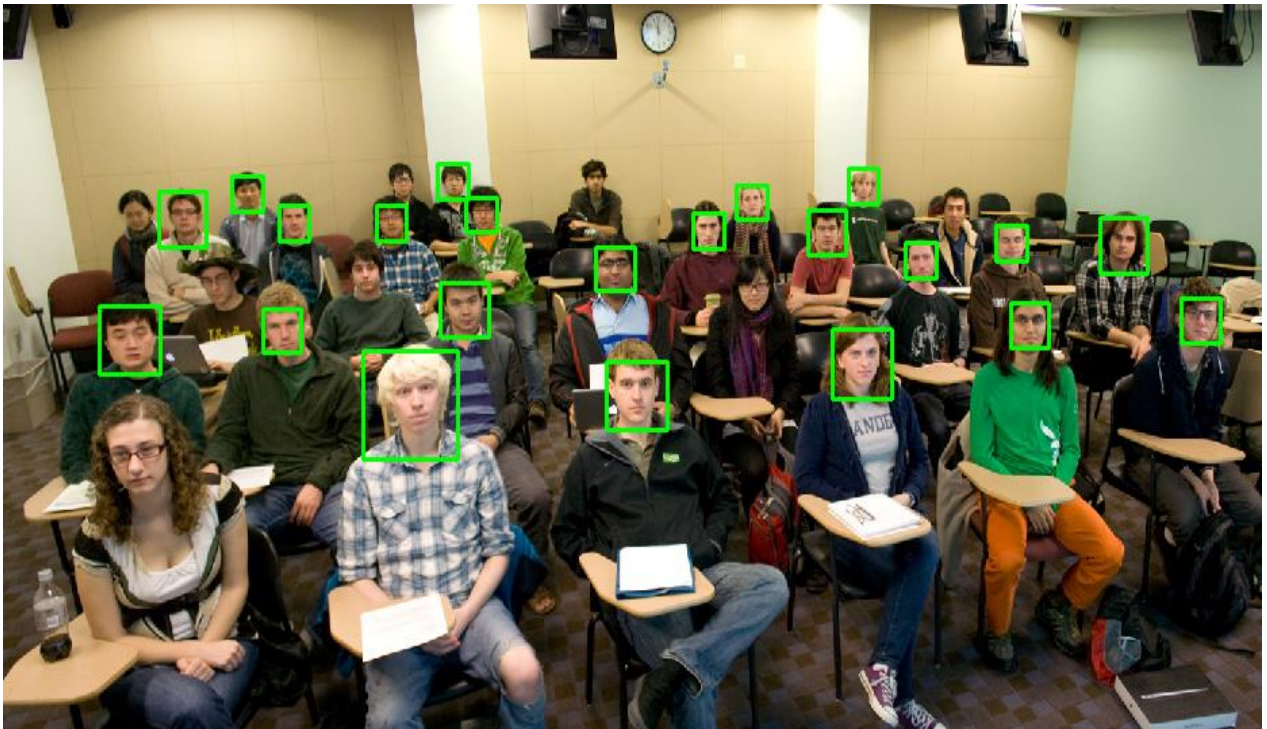


Рисунок 1.1. Приклад використання технології розпізнавання обличчя на групі студентів [62]

1.2.3 Традиційні методи

Деякі алгоритми розпізнавання обличчя визначають риси обличчя шляхом вилучення орієнтирів або функцій, від зображення обличчя суб'єкта. Наприклад, алгоритм може аналізувати відносну позицію, розмір та / або форму очей, носа, та щелепи [9]. Ці функції потім використовуються для пошуку інших зображень із відповідними функціями [12].

Інші алгоритми нормалізують галерею зображень обличчя, а потім стискають дані обличчя, зберігаючи лише дані, які корисні для розпізнавання обличчя. Тестове зображення порівнюється з даними обличчя [52]. Одна з найперших успішних систем базується на методах відповідності шаблонів [45], що застосовується для набору важливих ознак обличчя, що забезпечують своєрідне «стиснене» зображення обличчя. Приклад таких зображень показано на рис.1.2.

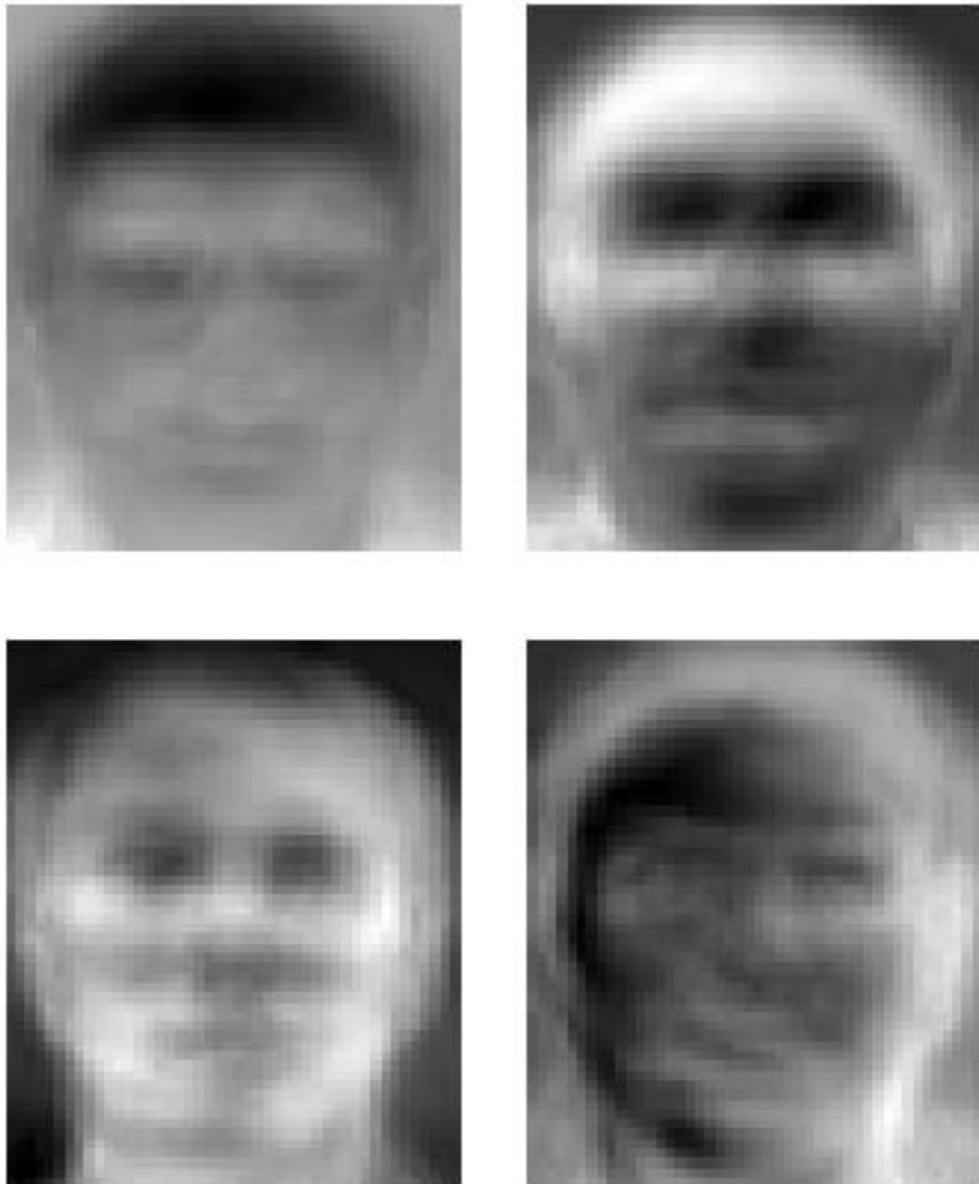


Рисунок 1.2. Узагальнені (стиснуті) зображення обличчя створені на основі великих масивів зображень [63]

Алгоритми розпізнавання можна розділити на два основних підходи: геометричний, який дивиться на відмінні ознаки, або фото-метричний, що є статистичним підходом, що розбиває зображення на значення та порівнює значення з шаблонами для усунення відхилень. Деякі класифікують ці алгоритми на дві широкі категорії: цілісні та функціональні моделі.

1.2.4 Переваги та недоліки технології

У 2006 році ефективність тогочасних алгоритмів розпізнавання обличчя оцінювалася у Face Recognition Grand Challenge (FRGC). Зображення з високою роздільною здатністю, 3-D скани обличчя та зображення сітківки ока були використані для тестування. Результати показали, що нові алгоритми в 10 разів більш точні, ніж алгоритми розпізнавання обличчя 2002 року та в 100 разів більш точні, ніж у 1995 році. Деякі з алгоритмів змогли перевершити учасників людей у та змогли однозначно ідентифікувати ідентичних близнюків [57].

У порівнянні з іншими біометричними методами, розпізнавання обличчя може бути не найбільш надійним та ефективним. Фактори, такі як освітлення, вираз обличчя, позиція та візуальний шум під час визначення обличчя, можуть вплинути на продуктивність систем розпізнавання обличчя.

Серед усіх біометричних систем, розпізнавання обличчя має найвищий процент помилково позитивний результатів та відхилень [55].

1.3 Технологія відстежування погляду

Відстежування погляду - це процес визначення точки погляду (куди дивиться людина), або руху очей відносно голови. Трекер очей - це прилад для вимірювання позицій та руху очей. Такі трекери використовуються в дослідженнях зорової системи, в психології, психолінгвістиці, маркетингу, у дизайні продуктів, а також як пристрій для взаємодії людини з комп'ютером. Трекери очей все частіше використовуються для реабілітаційних та допоміжних приладів (пов'язаних, наприклад, з керуванням інвалідними кріслами, роботизованими кінцівками та протезами).

1.3.1 Типи трекерів

Трекери відстежують рух очей багатьма способами, але в основному вони підпадають під одну з трьох категорій:

1. Відстежування руху предмета (як правило, спеціальної контактної лінзи), прикріпленого до ока;
2. Оптичне відстежування без прямого контакту з оком;
3. Вимірювання електричних потенціалів за допомогою електродів, розміщених навколо очей.

1.3.1.1. Спеціальні контактні лінзи

Використання першого способу передбачає кріплення до ока спеціальної контактної лінзи із вбудованим дзеркальним елементом або датчиком магнітного поля.

Рух прикріпленого пристрою заміряється з припущенням, що він не зміщується значно під час руху ока. Вимірювання з щільно прилягаючими контактними лінзами надають надзвичайно чутливі записи руху.

Це основний метод для дослідників, що вивчають динаміку та фізіологічну природу руху очей. Цей метод дозволяє вимірювати рух очей у будь-якому напрямку [33].

1.3.1.2. Оптичне відстежування

Другим методом є оптичне відстежування руху очей. Для цього у око направляють інфрачервоне світло, воно відбивається від очей і фіксується відеокамерою або іншим спеціально розробленим оптичним датчиком. Потім отримані дані аналізуються, щоб отримати інформацію про рух очей виходячи зі змін у зображеннях. На рис.1.3. зображено приклад, що використовує оптичний метод відстежування погляду.

Оптичні методи, особливо ті, що ґрунтуються на обробці відеосигналу, широко використовуються для відстеження погляду і вважаються неінвазивними та недорогими.

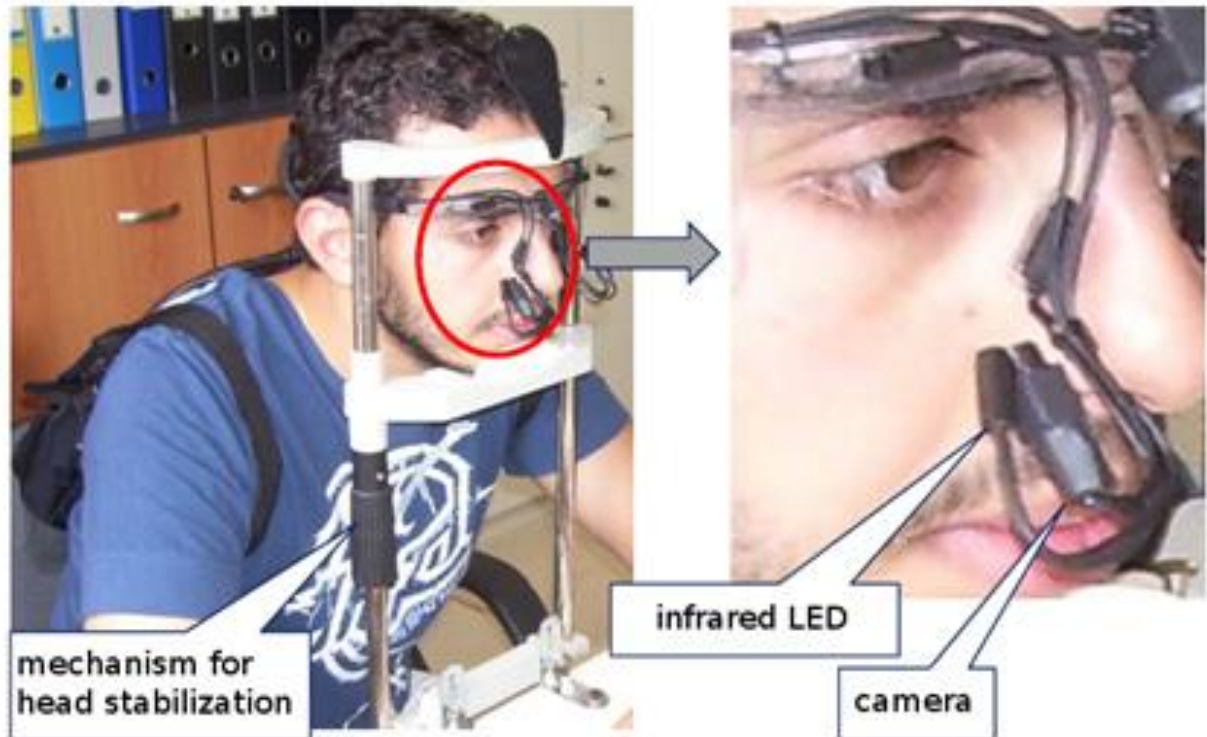


Рисунок 1.3. Приклад пристрою, що використовує оптичний метод відстежування погляду [64].

1.3.1.3. Вимірювання електричного потенціалу

У третьому способі використовуються електричні потенціали, вимірювані електродами, розміщеними навколо очей. Очі - це джерело стійкого електричного потенційного поля, яке можна виявити навіть в повній темряві та якщо очі закриті. Електричний сигнал, який може бути отриманий за допомогою двох пар контактних електродів, розміщених на шкірі навколо одного ока, називається Електроокулограма (EOG) [31].

Якщо очі рухаються від центрального положення до периферії, сітківка наближається до одного електрода, тоді як рогівка наближається до протилежного. Ця зміна орієнтації диполя і, відповідно, поля електричного потенціалу призводить до зміни вимірюваного сигналу EOG, і навпаки. Шляхом аналізу цих змін можна відстежити рух очей.

У даній роботі було запропоновано використання оптичного методу відстежування погляду із застосуванням гістограм орієнтованих градієнтів та алгоритму пошуку примітивів, заснованого на перетвореннях Хафа.

1.3.2. Застосування технології

Широкий спектр галузей використовує методи стеження за очима. Вони стають у нагоді у психології, маркетингу, вивченні людської поведінки, взаємодії людини з комп'ютером. Розроблені програми що збирають дані про активність людських очей під час читання, заняттях, сприйнятті реклами, тощо.

Останнім часом великі автомобільні концерни все частіше анонують вбудовані датчики та камери для відстежування погляду водіїв та їх стану, для випадків коли людина починає засинати за рулем, щоб ввімкнути звуковий сигнал, розбудити її та автоматично і безпечно зупинити автівку [41]. Приклад такого трекера зображено на рис.1.4.

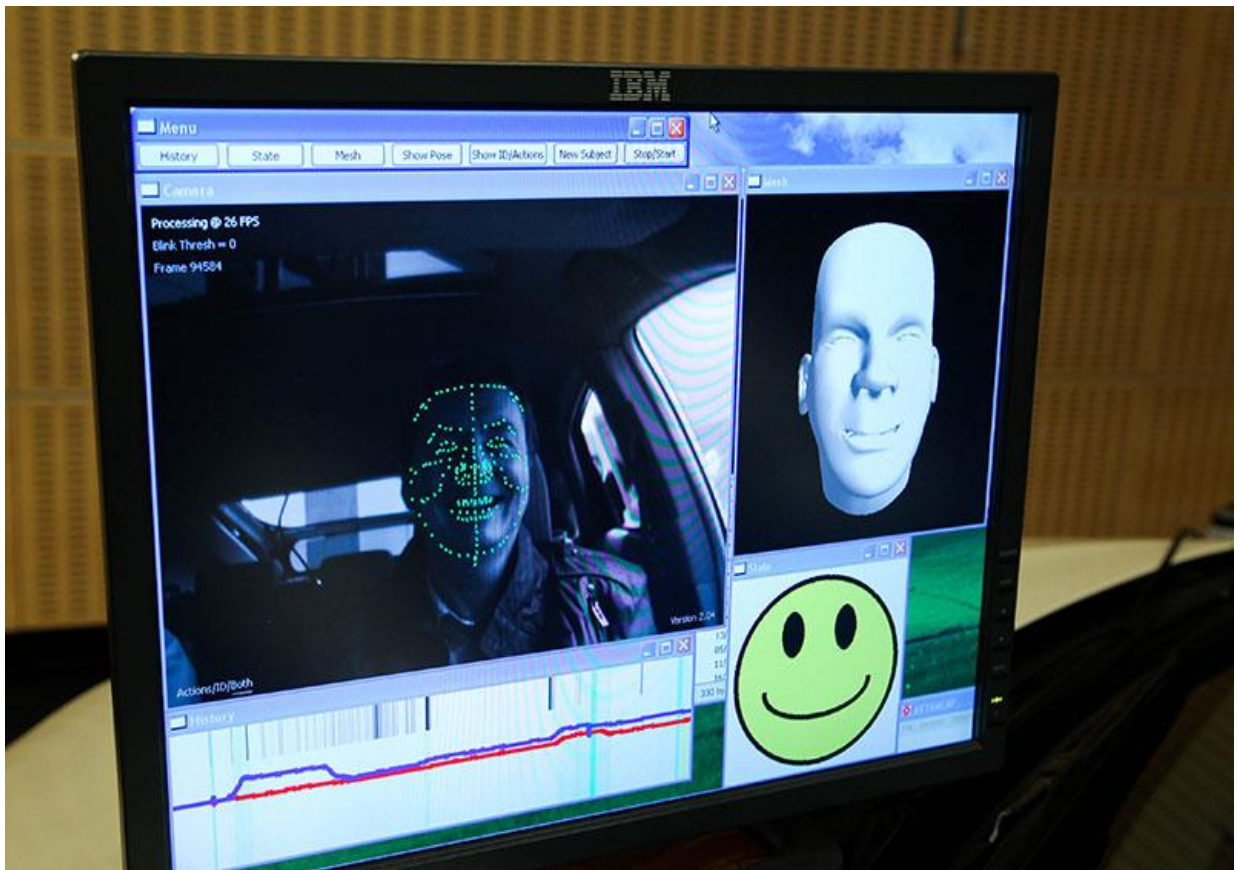


Рисунок 1.4. Відстежування положення обличчя водія автомобіля [65].

Технологія використовується й у медичних дослідженнях та лікуванні хворих. Зокрема у реабілітаційних програмах як інструмент для створення великої кількості комп'ютерних програм із вправами для пацієнтів та для відстеження прогресу реабілітації.

1.4 Прогнозування тексту

Прогнозування тексту - система прискореного введення тексту в цифрові пристрої, при якій програмне забезпечення пристрою в процесі набору пропонує варіанти закінчення слів і фраз, ґрунтуючись на наявних в його словнику, а також може пропонувати виправляти поширені помилки.

Прогнозування тексту може дозволити вводити ціле слово за допомогою однієї клавіші. Предиктивний текст робить ефективним використання меншої кількості ключів для введення запису в текстове повідомлення, електронну пошту, адресну книгу, календар тощо. [49]

Найбільш широко використовуваними прогностичними текстовими системами є T9, ІТАР, ezitext, і буквені / упорядковані. Існує багато способів побудови пристрою, який передбачає текст, але всі інтелектуальні текстові системи мають початкові лінгвістичні параметри, які пропонують прогнози, що визначаються пріоритетними, з метою адаптуватися до кожного користувача.

Навчувана система адаптується, за допомогою пам'яті пристрою, та зворотного зв'язку користувача, що призводить до коригування ключів. Більшість прогностичних текстових систем створюють базу даних для кожного користувача для полегшення цього процесу. [54]

Вибір інтелектуальної текстової системи що є найкращою для використання, передбачає відповідність стилю інтерфейсу користувача програмного додатку та визначення рівню здатності користувачів керувати інтелектуальним текстовим програмним забезпеченням.

Існують різні рівні ризику у використанні інтелектуальних текстових систем, порівняно з іншими системами. Предиктивні текстові системи потребують часу, щоб навчитися добре прогнозувати текст користувача, і тому, як правило, система мають завчасно встановлені бази даних сгенеровані із великою кількості анонімних баз даних користувачів аналогічних систем.

1.4.1 Зародження технології

Коротка служба повідомлень (SMS) дозволяє користувачеві мобільного телефону надсилати короткі текстові повідомлення. Найпоширенішою системою введення тексту SMS є система "Multi-Tap". При використанні системи "Multi-Tap", клавіші натискаються кілька разів, щоб отримати доступ до списку всіх букв на цій клавіші. Наприклад, натискання клавіші "2" один раз відображає "А", двічі відображає "В", а тричі відображає "С". Щоб набрати дві послідовні букви, які знаходяться на тій же клавіші, користувач повинен або призупинитись, або натиснути кнопку "Наступна".

Користувач може вводити текст, натискаючи клавіші на буквено-цифровій клавіатурі, не дивлячись на дисплей електронного обладнання. Таким чином, системі "Multi-Tap" легко навчитись, і вона може використовуватися без будь-якого візуального зворотного зв'язку. Однак цей метод не є дуже ефективним, оскільки потенційно вимагає багатьох натискань клавіш для введення єдиної літери [20]. Приклад стандартної клавіатури для обміну текстовими повідомленнями зображено на рис.1.5.

При ідеальному інтелектуальному наборі тексту. всі слова, що використовуються користувачем знаходяться у словнику, пунктуація ігнорується, а орфографічні та пунктуаційні помилки відсутні. Ідеальний словник включатиме всі сленгові слова, усі можливі іменники, аббревіатури, URL-адреси, іноземні слова та інші унікальні для конкретного користувача слова.



Рисунок 1.5. Стандартна клавіатура ITU-T E.161, що використовується для обміну текстовими повідомленнями [66].

Ці ідеальні умови мають змогу забезпечити мінімальну кількість необхідних натискань клавіш для вводу будь-якого слова чи частини інформації. Користувач натискає кнопки, що відповідають певним літерам бажаного слова, і, якщо слово існує у інтелектуальному текстовому словнику, або правильно визначається системою, воно з'явиться на екрані.

1.4.2 Використання сучасних алгоритмів та нейронних мереж

Для того щоб зробити максимально корисні прогнози, текстовий предиктор потребує якомога більше інформації про мову та стиль користувача. Найчастіше ця інформація отримується за допомогою машинного навчання та нейронних мереж.

Традиційно для розробки програм підказки / прогнозування тексту використовують дві моделі: «N-грамна модель» [50] та «Довга короткочасна пам'ять (LSTM - Long short-term memory)» [25]. У кожній з них є переваги та недоліки.

Висновки до розділу 1

В розділі розглянуто проблематику спілкування паралізованих хворих із оточуючим світом. Описано три основні типи трекерів, які використовуються у технології відстежування погляду та визначено тип, до якого відноситься розроблюємий програмний додаток, наведено основні відомості щодо областей застосування цієї технології у сучасному світі. Розглянуто та проаналізовано сучасні технології розпізнавання обличчя, відстежування погляду та предиктивного набору тексту, та сценарії їх використання, які будуть у подальшому задіяні для створення програмного додатку.

РОЗДІЛ 2

ПРОЕКТУВАННЯ СКЛАДОВИХ ЧАСТИН СИСТЕМИ

В даному розділі головною задачею є вибір оптимальних інструментів та методів для успішного проектування системи, що забезпечить максимальну гнучкість та швидкість у процесі розробки та зменшить кількість потенціальних перешкод та помилок на етапах її реалізації. Для цього буде порівняно можливі способи реалізації проекту, які описані в наступних підрозділах

2.1. Теоретична основа системи

Прогнозування тексту є одним із найбільш широко використовуваних методів для підвищення швидкості спілкування за допомогою електронних пристроїв. Системи прогнозування дуже часто використовуються людьми з обмеженими можливостями (наприклад, люди з порушеннями функцій моторики та мовлення). За останні роки широко розповсюдилися нові способи застосування для цієї технології, наприклад, написання коротких текстових повідомлень за допомогою мобільних телефонів [51].

В науковій літературі можна знайти величезну кількість різноманітних методів і прийомів для прогнозування тексту. Їх неоднорідність ускладнює їх розуміння і порівняння, для того щоб вибрати найбільш підходящий метод для конкретної реалізації програмного додатку. У цьому розділі представлено огляд різних методів прогнозування тексту з метою надати систематичний погляд на цю область. Розглядаються програми прогнозування та пов'язані з ними функції, такі як розмір блоку, структура словника, методи передбачення, інтерфейс користувача тощо. Крім того, порівнюються параметри прогнозних вимірювань та опубліковані результати. Розглядається велика кількість факторів, які можуть вплинути на результати прогнозування, включаючи прийняття системи користувачами, і їх вплив на результативність [26].

2.2 Поведінка та вимоги систем прогнозування тексту

2.2.1 Обмеження людей із вадами руху

Люди з фізичними та/або когнітивними вадами, люди з вадами сприйняття, можуть мати знижені комунікативні здібності через труднощі під час використання звичайних способів міжособистісного спілкування. Протягом останніх кількох років кілька комп'ютеризованих систем, як правило, заснованих на теоріях доповнюючої та альтернативної комунікації, були розроблені для допомоги людям з обмеженими можливостями в цілях особистого спілкування. Використовуючи ці системи, люди з серйозними порушеннями мови та моторики можуть подолати багато комунікаційних обмежень. Наприклад, люди з поганим контролем верхніх кінцівок можуть вводити текст за допомогою інтерфейсів на основі вибору елементів за допомогою однієї клавіші з набору елементів, послідовно сканованих системою [46]. На рис.2.1 зображено приклад спеціального контроллера.



Рисунок 2.1. Паралізований пацієнт, користується Інтернетом за допомогою перемикача у формі соски [67]

Системи сканування рухів та вибору вимагають лише обробки залишкових довільно контрольованих рухів у будь-якій частині тіла. Тим не менш, ці комунікаційні пристрої надзвичайно повільні, пропускаючи лише кілька слів за хвилину. Це може призвести до неприємних ситуацій, коли люди з обмеженими можливостями не можуть брати участь у звичайних розмовах і, отже, можуть стати соціально виключеними [39].

2.2.2 Вимоги до систем передбачення тексту та їх користувачів

До появи комп'ютерних комунікаторів для альтернативного спілкування використовувалися дошки, що містять символи, склади та слова. Співрозмовник дивиться на послідовність пунктів, на які вказує користувач дошки зв'язку, і складає речення. Часто користувачеві не потрібно було закінчувати слово, оскільки співрозмовник міг передбачити решту букв. Для цього співрозмовники використовували свої знання про структуру мови (морфологічний та синтаксичний рівні), контекст бесіди (семантичний та прагматичний рівні). Ця здатність прогнозування призвела до значного збільшення швидкості спілкування та відчуття інтеграції між обома мовцями. Отже, проблема для розробників програмних додатків полягає в тому, чи здатний комп'ютер відтворити таку поведінку, пропонуючи розумні прогнози на основі доступної інформації щодо мови та контексту [48].

Як зазначалося раніше, предиктори часто включаються в засоби комунікації, щоб підвищити швидкість спілкування користувача. Таким чином, ці предиктори можуть бути корисними для людей з серйозними руховими та мовними вадами, такими як церебральний параліч, геміплексія тощо, за умови, що вони вміють їх використовувати. Користувачі систем прогнозування повинні мати певну фонетичну та правописну здатність. Їм також потрібно розрізняти слова, які схожі в списках передбачення, і мають бути в змозі гнучко переміщати свою увагу вперед і назад між пристроєм введення і списком передбачення. Навіть якщо методи прогнозування в першу чергу були розроблені для задоволення потреб людей з обмеженими

можливостями, вони можуть принести користь будь-кому, якщо вони належним чином інтегровані в інтерфейс користувача програми [7].

2.2.3 Поведінка систем прогнозування тексту

Методики прогнозування використовують контекстну інформацію, щоб передбачити, що людина збирається написати [14]. Якщо система вміє правильно передбачувати, кількість натискань клавіш, необхідних для написання речення, зменшується, і, крім збільшення швидкості спілкування, зменшуються фізичні зусилля, необхідні для складання повідомлень. На рис.2.2 зображено інтерфейс пошукової системи Google із предиктивним текстом.

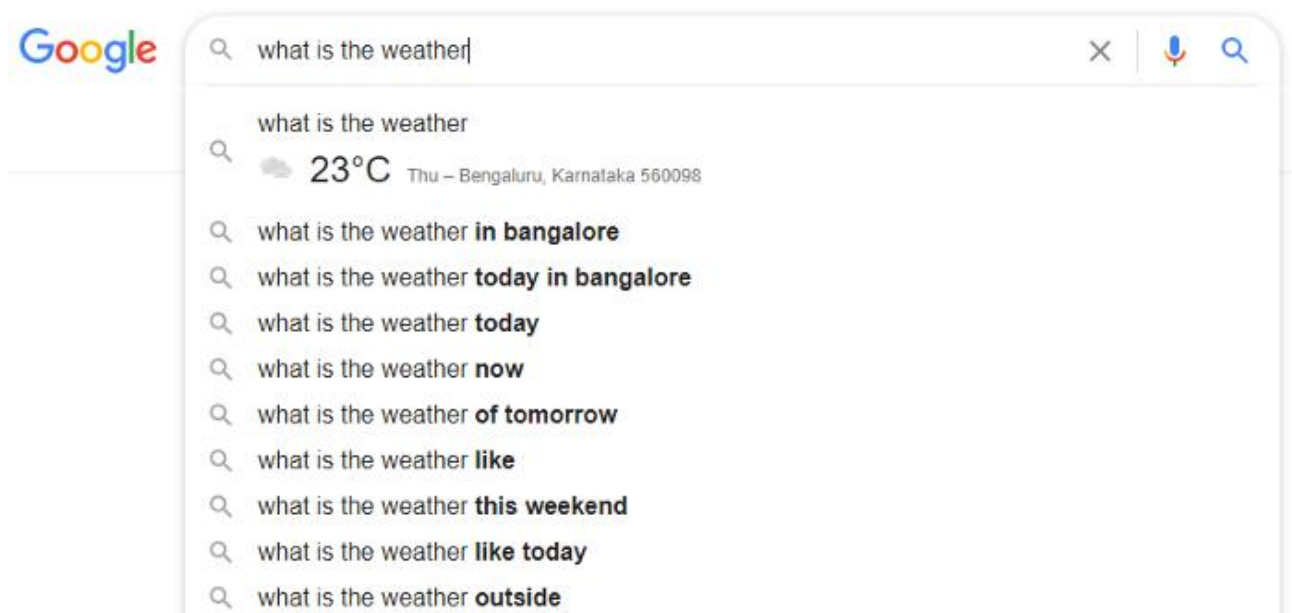


Рисунок 2.2. Приклад використання системи прогнозування тексту компанією Google [68]

Крім того, програмне забезпечення для передбачення також може виправляти орфографічні помилки, змінювати порядок слів у реченнях і в цілому покращувати якість створених повідомлень [60]. Системи передбачення можна розглядати як інтелектуальних помічників, які допомагають користувачам складати тексти. Вони фіксують введений користувачами текст, щоб зробити свої припущення та видавати результати,

які можуть бути включені в програми, що використовуються для складання текстів. Як було сказано раніше, вони намагаються наслідувати поведінку користувачів. Отже, найдосконаліші з них мають свої особливості навчання, здатні робити висновки, пристосовуватися та діяти самостійно. Більше того, у певних випадках вони можуть контактувати з користувачами, в основному для адаптації особистого словникового запасу. Часто у системах, що використовують методи прогнозування, виникають деякі проблеми з інтерфейсом користувача, які впливають на ефективність системи прогнозування [29]. Ці проблеми також будуть розглянуті далі.

2.3 Аналіз систем прогнозування

У контексті систем допоміжної комунікації, предиктор — це система, яка намагається передбачити наступний блок символів (букв, складів, слів, речень тощо), які користувач хоче висловити. Загалом, передбачення базується на раніше введених блоках. Основна мета предиктора – скоротити зусилля та час на опрацювання повідомлення [5]. Щоб зменшити зусилля, необхідно зменшити кількість натискань клавіш, необхідних для складання повідомлення, а для скорочення необхідного часу кількість символів, введених у текст за допомогою одного передбачення, має бути більшою за кількість символів, написаних за допомогою почергового натискання на клавіші.

Існують дві основні концепції, які необхідно враховувати при аналізі комунікаційних систем: чіткість та легкість [13]. Система введення має високу чіткість, якщо завжди є хоча б один варіант, що відповідає потребам користувача, який можна вибрати з усіх запропонованих пропозицій. Тобто символ(и), які користувач хоче написати, завжди є серед тих, які пропонує система. Легкість, яка безпосередньо пов'язана зі швидкістю зв'язку, — це кількість символів, створених за одиницю часу за допомогою системи

Чіткість і легкість - протилежні поняття, так як збільшення одного зазвичай призводить до зменшення іншого [47].

Система, яка пропонує користувачеві лише алфавіт, є прикладом дуже чіткої, але дуже важкої системи, оскільки при її використанні завжди можна зробити вибір, але вихідна швидкість зв'язку низька, особливо коли користувач має фізичні вади.

Дуже легка, але нечітка система — це система, яка одразу пропонує користувачу кілька поширених слів або речень. Якщо вибрано одну з наданих опцій, швидкість зв'язку буде більшою, ніж при складанні слів по буквах; однак потрібної опції може не бути у запропонованому наборі [28].

Мета комунікаційного пристрою полягає в тому, щоб максимізувати та збалансувати як чіткість, так і легкість. При спробі отримати високий рівень чіткості користувачам пропонуються всі можливі символи. У разі максимізації легкості, окрім символів алфавіту, слід запропонувати різні варіанти, довші за один символ [40]. Якщо ці параметри вибираються часто, легкість мовлення збільшується, а чіткість залишається на високому рівні, оскільки щоразу можна вибирати принаймні одну з показаних опцій.

2.3.1 Розмір текстового блоку

Вибір розміру передбачуваних блоків є дуже важливим фактором, оскільки швидкість натискання клавіш і економія часу напряму залежать від розміру передбачуваного блоку. У цій роботі розглядаються лише блоки, що складаються з алфавітних символів (інші символи та піктограми не враховуються). Програма зосереджена на передбаченні тексту. Для чистого тексту мінімальним розміром блоку є символ. У кожній мові кожен символ має різну частоту появи. Першим підходом можна запропонувати брати усі літери, відсортовані за частотою, без урахування попереднього набраного тексту (тобто з нульовою пам'яттю) [59]. Таким чином, пропозиції завжди однакові, незалежно від попередньо набраних символів. Очевидно, що цей метод не передбачає жодного збереження натиснутих клавіш, оскільки вимагає однакової кількості натискань клавіш, щоб прийняти пропозиції та вивести ці символи на екран у разі успіху. Отже, при цьому методі усі символи для

швидшого доступу розподіляються залежно від частоти їх використання в тій чи іншій мові [32].

Оскільки передбачення лише одного символу дає дуже погані результати, потрібно обробляти більші блоки. Для використання доступної статистичної інформації, що надходить із попереднього тексту, часто використовуються N-грами.

N-грама – це послідовність з N елементів. Це може бути послідовність звуків, складів, слів або букв. На практиці частіше зустрічається N-грами як ряд слів, стійкі словосполучення називають колокацією. Послідовність з одного елемента зветься уніграм, з двох послідовних елементів називають біграм, послідовність з трьох елементів – триграма. Чотири і більше елементів позначаються як N-грами [4]. Приклад розбиття на N-грами можна спостерігати на рис 2.3.

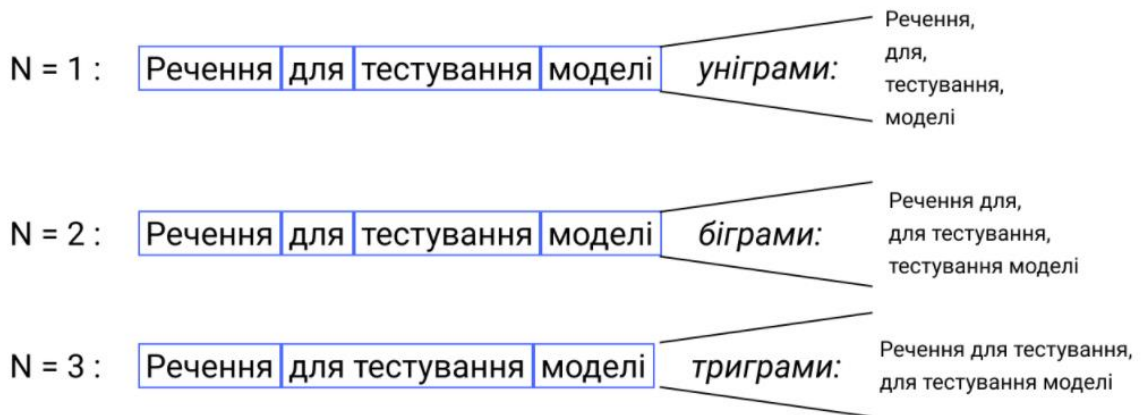


Рисунок 2.3. Візуальна схема розбиття речення на N-грами [69]

N-грами використовується в основному для передбачення на основі імовірнісних моделей. N-грамна модель розраховує ймовірність останнього слова N-грами, якщо відомі всі попередні, причому вважається, що поява кожного наступного слова залежить тільки від попередніх слів.

Математичним інструментом для здійснення розрахунків таких ймовірностей найчастіше використовують засоби, пов'язані з «ланцюгами Маркова» [30].

Як і для будь-якої іншої моделі штучного інтелекту, потрібно навчати модель на великому масиві даних. Після навчання система, як правило, досить добре передбачає появу наступного слова.

У такої системи є і обмеження, оскільки точність передбачення залежить від того, скільки разів конкретне слово з'являлося у наборі вхідних тренувальних даних, тобто передбачення здійснюються лише на основі частоти використання елементів N-грам.

Якщо ми маємо таблицю, що містить N-грамні частоти, то можна передбачити N-й символ після блоку з N-1 символів [21].

Інформація, пов'язана з N-грамами, використовується для відображення матриць динамічного сканування, де найбільш вірогідні (по відношенню до останніх вибраних символів) розміщені біля верхньої лівої клітинки в матриці.

Традиційно максимальна довжина N-грамів становить чотири (тетраграми), в основному через обмеження обчислень і зберігання [10].

Іншими блоками, які слід враховувати, є склади та морфеми. Оскільки вони є «природними» частинами слів, вони мають цікаві мовні властивості, які дуже корисні для цілей передбачення. Хоча кількість символів у складі змінна (скажімо, S символів на склад у середньому), кількість складів, які використовуються в даній мові, менша за кількість нграм, яку можна обчислити. Таким чином, кількість складів для збереження менше, отже, збільшується ймовірність їх передбачення.

Навіть якщо склади та N-грами дуже передбачувані, користувачі відчують певні труднощі, з тим щоб ідентифікувати та інтегрувати їх у текст. Були запропоновані системи, які намагаються вгадати комбінацію окремих символів, N-грам і складів. Хоча економія, отримана з меншими блоками, є низькою, вони зберігаються в передбачуваному наборі, щоб підтримувати узгодженість системи [37].

З іншого боку, набір морфем у мові складається з коренів слів і можливих афіксів (префіксів, інфіксів і суфіксів). У цьому випадку необхідно, щоб і система, і користувач змогли чітко розрізнити, що таке морфема. Ці блоки дуже цікаві у мовах з великим показником словозмін (як у німецькій, шведській, баскській тощо), оскільки ці мови створюють велику кількість різноманітних форм [22], що поєднують різноманітні морфеми, і було б майже неможливо створити лексикон, що включає всі ці можливі форми.

Системи передбачення слів намагаються вгадати слово, яке намагається написати користувач, враховуючи раніше написані символи. Ця стратегія зазвичай дає задовільний коефіцієнт влучань (оскільки слова часто дуже передбачувані) і важливу економію натискань клавіш [6]. Крім того, когнітивні зусилля користувача низькі, оскільки слова легко розрізнити, а їх розділення чітке.

Інший метод — спробувати блоки довше одного слова. У цьому випадку система намагається передбачити наступні слова, беручи до уваги останні введені слова. У разі правильного передбачення, натискання клавіш і економія часу буде дуже високою; однак коефіцієнт попадання для більш ніж одного слова дуже низький, оскільки частоти послідовностей слів надзвичайно низькі, за винятком кількох поширених речень, як-от у коментарях, привітаннях та прощаннях [8].

Підводячи підсумок, серед блоків усіх можливих розмірів, слово зазвичай вибирається завдяки його оптимальному співвідношенню між коефіцієнтом правильного передбачення, економією натискань клавіш і когнітивними витратами з боку користувачів. Крім того, додавання морфологічної, синтаксичної та семантичної інформації може покращити коефіцієнт звернення та економію натискань клавіш. Тому слово є найбільш вживаним прогностичним блоком [38].

2.3.2 Підхід довготривалої короткочасної пам'яті (LSTM)

Більш просунутий підхід, із застосуванням нейронної моделі мови, полягає у використанні довготривалої короткочасної пам'яті (LSTM). Модель

LSTM використовує глибоке навчання з мережею штучних «клітин», які керують пам'яттю, що робить їх більш придатними для передбачення тексту, ніж традиційні нейронні мережі та інші моделі [16].

Основним компонентом LSTM є клітина пам'яті. Вона може зберігати свій стан з плином часу та складається з явної пам'яті (її також називають вектором стану клітини) та блоків воріт. Ворота регулюють потік інформації до/з пам'яті.

Модуль LSTM має 3 типи воріт, які називають забувальними, вхідними та вихідними (з англ. forget, input, output gates).

- Вектор стану клітини являє собою пам'ять LSTM, і зазнає змін шляхом забуття старої пам'яті (забувальні ворота) та додавання нової пам'яті (вхідні ворота).

- Забувальні ворота контролюють, яку інформацію видалити з пам'яті, та вирішують скільки минулої інформації слід пам'ятати.

- Вхідні ворота контролюють, яка нова інформація додається до стану клітини з поточного вводу, та визначають скільки одиниць інформації буде додано до поточного стану.

- Вихідні ворота вирішують, що виводити з пам'яті за певних умов, та яка частина поточної клітини потрапляє на вихід.

Ці ворота реалізуються із застосуванням логістичної функції для обчислення значень між 0 та 1. Для часткового дозволу або заборони передачі інформації до або з цієї пам'яті застосовується множення на це значення. Наприклад, вхідні ворота (англ. input gate) керують мірою передачі нового значення до пам'яті. Забувальні ворота (англ. forget gate) керують мірою зберігання значення в пам'яті. А вихідні ворота (англ. output gate) керують мірою передачі значення з пам'яті для обчислення активування виходу блоку

Стандартні рекурентні нейронні мережі та інші мовні моделі стають менш точними, коли розрив між контекстом і словом, яке передбачається, збільшується. Ось коли LSTM вживається для вирішення проблеми

довгострокової залежності, оскільки він має комірки пам'яті, які запам'ятовують попередній контекст.

Щоб зрозуміти довготривалу короткочасну пам'ять (LSTM), необхідно зрозуміти рекурентні нейронні мережі (RNN). RNN - це тип нейронної мережі, де результати з попереднього кроку подаються як вхідні дані для поточного кроку. Іншими словами, RNN-це узагальнення нейронної мережі, яка має внутрішню пам'ять [27]. RNN призначені для розпізнавання послідовних характеристик даних і використовують шаблони для прогнозування наступного ймовірного сценарію.

Передбачається, що всі вхідні та вихідні дані незалежні один від одного у звичайній нейронній мережі. Однак це не зовсім підходить для передбачення наступного слова у реченні, оскільки було б набагато цінніше знати, які слова стоять перед ним. Рекурентні мережі використовують послідовну інформацію, вони називаються рекурентними, оскільки виконують однаковий процес для кожного елемента послідовності, при цьому вихідні дані залежать від попередніх обчислень, що фіксуються у «пам'ять», яка зберігає інформацію про те, що вже було обраховано [58].

Проте у RNN є проблеми з короткочасною пам'яттю. Якщо послідовність досить довга, їм важко зберігати інформацію від ранніх кроків до більш пізніх. Якщо ви намагаєтесь обробити абзац тексту, щоб зробити прогнози, рекурентні мережі можуть «забувати» важливу інформацію що була присутня у початку абзацу. Це спричиняє потребу в довготривалій короткочасній пам'яті (LSTM), яка є особливим видом RNN, здатною вивчати довгострокові залежності. LSTM мають змогу запам'ятовувати інформацію протягом тривалого періоду часу.

2.4 Поведінка та вимоги систем відстежування погляду

У дослідженнях пов'язаних з відстежуванням очей та погляду в цілому використовується широкий спектр заходів для спостереження за рухами очей

під час виконання різних типів завдань. Ці заходи включають вимірювання на основі фіксації, вимірювання за допомогою зіниць, вимірювання на основі саккад (швидких, суворо узгоджені рухів очей, що відбуваються одночасно і в одному напрямку), та вимірювання на основі переміщень. Далі їх можна класифікувати на вимірювання руху, вимірювання положення, кількосні вимірювання та вимірювання відстані.

У більшості програмного забезпечення для відстеження очей найпопулярнішими показниками виступають тривалість фіксації, кількість фіксацій, час до першої фіксації, тривалість першої фіксації тощо. Більшість програмного забезпечення для відстеження очей також пропонує візуалізацію запису, наприклад, теплові карти та графіки погляду. Однак деякі показники не розраховуються автоматично програмним забезпеченням. У таких ситуаціях необхідна ручна ідентифікація у так званих необроблених даних. Наприклад, доступ до інформації про розміри зіниць найчастіше вимагає ручної ідентифікації в необроблених даних, і ця ручна робота потенційно може зайняти досить багато часу.

2.4.1 Фіксації

Фіксація — це тип руху очей, який часто визначають як період часу, протягом якого око є відносно стабільним; мета фіксацій – візуально зафіксувати об'єкт інтересу. Тривалість фіксації та їх підрахунок часто застосовуються для індексації когнітивних зусиль. Більш тривалі фіксації та більша їх кількість вказують на більш докладну обробку інформації, а коротші фіксації та менша їх кількість вказують на менш напружену обробку інформації. Більш напружена обробка часто пов'язана зі збільшенням труднощів при розпізнаванні образів. Середня тривалість фіксації становить 225 мс під час читання мовчки і 275 мс під час читання вголос. Під час набору тексту, тобто при одночасному читанні та введенні, середня тривалість фіксації становить 400 мс, що значно більше, ніж при «звичайному» читанні. Завдання без читання, такі як візуальний пошук і сприйняття сцени, також дають різну тривалість, а саме тривалість фіксації 275 мс і 330 мс відповідно.

2.4.2 Розмір зіниць

Вимірювання розміру або розширення зіниць часто приймають як показник робочого навантаження на когнітивну систему. Загалом, більш розширені зіниці вказують на більш високе когнітивне навантаження, тобто вони вказують на те, що завдання є відносно складнішим, а менше розширення вказує на нижче когнітивне навантаження, тобто на відносно легше завдання.

Загалом, слід бути обережними при зборі та аналізі даних про розмір зіниці, оскільки цей тип руху очей чутливий не тільки до змін когнітивного навантаження, але й до багатьох інших факторів. Точніше, зіниці розширюються та звужуються як рефлекторна реакція на зміни в інтенсивності освітлення; у відповідь на емоційні події, такі як стрес, біль і страх; у відповідь на прийом ліків і стимуляторів, таких як наркотики та алкоголь; або якщо людина хвора. Хоча важко контролювати емоційний стан учасника під час запису даних стеження за очима, менш проблематично контролювати інші фактори. Підтримка однакової інтенсивності світла в кімнаті, де записуються дані відстеження очей, має вирішальне значення, зокрема для досліджень, які цікавляться зміною розмірів зіниць.

2.4.3 Якість даних при вимірюванні

Як зазначалося вище, якість даних відстежування очей чутлива до різних факторів. Навіть якщо були вжиті всі необхідні запобіжні заходи, дані відстежування все ще можуть бути настільки низької якості, що відображення рухів очей та визначення їх положення буде майже неможливим. Тому ретельна оцінка якості даних при відстежуванні очей є вирішальним кроком у процесі аналізу; однак це крок, яким часто нехтують у процесі дослідження. Нижче представлено та обговорено три методи оцінки якості даних відстеження очей.

Вимірювання фіксацій

Одним із відносно популярних показників якості даних відстежування погляду є обчислення середньої тривалості фіксації. Часто у процесі відстежування фіксації які є занадто короткими, а саме < 200 мілісекунд,

відкидаються та враховуються як похибка. Середня тривалість фіксації під час мовчазного читання зазвичай становить близько 225 мілісекунд. Середній поріг тривалості фіксації в 200 мілісекунд використовують щоб відрізнити прийнятні дані від неприйнятних.

Покладання лише на тривалість фіксації як на показник якості не позбавляє від усіх можливих проблем. Середня тривалість фіксації є відносно грубою мірою, яка ігнорує потенціальну різницю у повноті даних записів відстежування очей. Кількість рухів очей, що було успішно зафіксовано трекером, порівняно з тим, скільки даних про рух не було записано, змінюється між записами в залежності від різних факторів. Наприклад, як було сказано вище, на якість запису може вплинути використання учасником оптичних засобів – окуляри, тощо, що означає, що записані дані відстеження очей можуть бути лише частковими відображеннями рухів очей учасника. Можливо, запис учасника має середню тривалість фіксації >200 мілісекунд, але якщо це середнє значення розраховується на основі лише кількох секунд запису з лише кількома фіксаціями, то воно не відображає загальну якість даних відстежування погляду під час запису.

Час фіксації погляду на екрані

Час погляду на екран використовують як метод для подальшого вимірювання якості даних відстежування. Це простий розрахунок загальної тривалості фіксації у відсотках від загального часу виконання завдання. Оцінка показує, скільки часу користувач провів, дивлячись на екран, що можна інтерпретувати як міру якості даних відстеження очей. Високий показник може свідчити про те, що учасник дивився на монітор протягом значної кількості часу, або що трекер добре фіксував рухи очей користувача. Низька оцінка може свідчити про те, що учасник дивився на монітор лише обмежений проміжок часу, або що трекер погано фіксував рухи очей користувача. При реальному застосуванні програмного додатку ми можемо очікувати, що користувач буде дивитися на монітор протягом досить тривалого часу, щоб мати змогу успішно набирати текст на віртуальній

клавіатурі. Очевидно будуть спостерігатися випадки, коли користувач буде відводити погляд від монітора і дана оцінка якості спостереження дуже рідко буде 100-відсотковою. Як підсумок, ми не можемо бути впевнений, чи низький час погляду на екран є наслідком загальної поганої якості даних відстежування очей, чи користувач відводив погляд від монітора протягом значних періодів часу.

Відношення фіксацій до часу погляду на монітор

Як відповідь на недоліки попередніх методів вимірювання часто використовується третій метод для визначення якості даних при відстежуванні погляду. Метод відношення часу погляду до фіксації заснований на тому, що саккади становлять від 5 до 15 відсотків усіх рухів очей при читанні. В ідеальних умовах запис даних буде відображати це співвідношення, за яким приблизно 85-95 відсотків одиниць погляду на записі можна віднести до фіксацій, а приблизно 5-15 відсотків — до саккад. Відношення розраховується шляхом порівняння загальної кількості одиниць погляду із загальною кількістю одиниць погляду, які були частиною фіксації. Наприклад, у записі, що містить 11 000 зразків одиниць погляду, з яких 10 000 належать до фіксацій, відсоток відношення буде становити 90,9%. На відміну від вимірювання часу погляду на монітор, цей метод не передбачає, що користувачі проводять однакову відносну кількість часу, дивлячись на монітор, і тому є кращою альтернативою вимірюванню якості даних відстежування погляду.

2.4.4 Висновки щодо якості даних

У поєднанні ці три методи пропонують надійний інструмент для визначення, які записи слід включати в аналіз дослідження, а які ні через проблеми з якістю даних. Не менш корисними можуть бути й інші методи відбору ненадійних даних від надійних, наприклад, метод на основі реєстрацій розміру зіниць; ключовим моментом тут є те, що нехтування належним відбором даних може мати серйозний вплив на аналіз даних дослідження і потенційно може спотворювати отримані результати.

Дані відстеження очей потенційно можуть невірно відобразити фактичне використання системи, якщо не дотримуватись обережності під час їх збору. З точки зору вибору трекерів, дистанційні трекари, як правило, краще підходять для досліджень, оскільки вони менш інвазивні, ніж системи, що кріпляться на голові користувача. Що стосується інтерпретації даних відстежування погляду як прояву когнітивної обробки інформації, зв'язок між рухом очима та обробкою інформації є інтуїтивно зрозумілим; однак є проблеми, які можуть робити зв'язок між ними менш очевидним. Вони потенційно можуть ускладнити інтерпретацію даних, і дослідник повинен розглянути, як ці проблеми можуть вплинути на результати дослідження та як можна мінімізувати вплив цих проблем. Під час збору та аналізу даних про рух очей існує безліч факторів, які не обов'язково пов'язані з діяльністю користувача під час використання програмного додатку, і вони потенційно можуть спотворити аналіз записаних даних. Зіниці, наприклад, чутливі до багатьох факторів, включаючи зміни в інтенсивності світла та емоційному стані користувача, і дослідники повинні намагатися контролювати ці потенційно спричиняючі помилки фактори під час збору даних, а також під час їх аналізу. Крім того, хоча показники фіксації та зіниць є популярними індикаторами когнітивної обробки в дослідженнях відстежування погляду, інші показники, популярні в інших наукових дисциплінах, можуть бути корисними для вивчення цих процесів.

2.5 Визначення вимог до системи

Створювана система призначена для:

- 1) полегшення комунікації із паралізованими хворими;
- 2) надання хворим додаткового каналу спілкування.

Система буде мати власний графічний інтерфейс користувача GUI (graphical user interface) і може використовуватись в медичних організаціях для:

- а) встановлення на вже існуючі персональні комп'ютери;
- б) зменшення навантаження на медичний персонал закладу.

Система повинна надавати можливість виконання перерахованих нижче функцій:

- 1) відстежувати обличчя та погляд пацієнта за допомогою веб-камери;
- 2) виводити на монітор віртуальну клавіатуру з варіантами для вибору;
- 3) надавати пацієнту можливість набирати текст за допомогою погляду та передавати ці дані лікарям.

Надійне функціонування системи повинне бути забезпечене шляхом виконання замовником сукупності організаційно-технічних заходів, перелік яких приведено нижче:

- а) встановлення джерел безперебійного живлення для забезпечення роботи технічних засобів;
- б) обладнання пристроїв веб-камерами;
- в) відсутність сторонніх програм, які можуть призвести до непрацездатності системи.

Необхідно уникати потенційних неполадок у роботі системи пов'язаних з некоректною відповіддю програмного забезпечення на дії користувача. Для цього необхідна реалізація засобів захисту та системи обробки помилок на стадії розробки програмного додатку.

Персонал установи де буде використовуватися програмний додаток і який має на пряму взаємодіяти із ним та пацієнтами, повинен пройти курс підготовки для ознайомлення із функціями додатку та навчання повсякденній роботі із ним.

Висновки до розділу 2

В розділі розглянуто основні вимоги як до систем прогнозування тексту, так і для майбутніх користувачів цих систем, проблематику спілкування паралізованих хворих із оточуючим світом. Виділено основні типи

предиктивних систем та різні підходи до передбачення текстового вводу користувача, які будуть у подальшому задіяні для створення програмного додатку. Розглянуто поняття N-грам та Довготривалої короткочасної пам'яті (LSTM). Наведено основні відомості щодо областей застосування цих технологій у сучасному світі. Визначено вимоги до розроблюваної системи.

РОЗДІЛ 3

РОЗРОБКА КОМУНІКАЦІЙНОЇ СИСТЕМИ

Головними завданнями поставлених в рамках цієї роботи є дослідити принципи і засоби прогнозування тексту, та розробити унікальну модель, навчену на певному масиві вхідних даних. Покращення словника моделі будуть зроблені після початкового етапу розробки шляхом включення більшої кількості навчальних даних. У кінцевому вигляді програмний додаток має складатися із віртуальної екранної клавіатури, поля для вводу, вікна із відеосигналом з вебкамери пристрою в реальному часі та області над клавіатурою, де будуть відображатися наступні слова передбачені системою прогнозування тексту як пропозиції для користувачів додатку.

3.1 Обробка природної мови (NLP) з N-грамами і марковськими моделями

В обробці природної мови (Natural Language Processing) одна з найпростіших прогнозних текстових моделей будується з використанням N-грам для визначення наступних слів шляхом присвоєння ймовірностей послідовності слів. N-грама — це група від 1 до N слів, які послідовно «витягуються» з великої кількості тексту [11]. Щоб створити N-грам, слова групуються, як у реченнях документа. Наприклад, у реченні «Це тестова модель природної мови» є п'ять уніграм («Це», «тестова», «модель», «природної», «мови»), чотири біграми («Це тестова», «тестова модель», «модель природної», «природної мови») і три триграми («Це тестова модель», «тестова модель природної», «модель природної мови»). Таким чином, у дуже великих базах з мільйонами рядків тексту модель буде навчатися, використовуючи частину цих N-грамів та їхні ймовірності для наступних слів.

Оскільки кінцева мета прогнозної моделі матиме суворі вимоги до використання пам'яті, ми будемо шукати модель, яка потребує мінімального об'єму пам'яті комп'ютера та працює найбільш ефективно. Модель Маркова передбачає, що ймовірність наступного слова залежить лише від ймовірності попереднього слова [23]. Отже, для моделі з біграмою нам потрібно дивитися лише на одне слово назад, для моделі з триграмою — лише два слова, і так далі. Навіть у величезному вхідному наборі тексту триграми, швидше за все, будуть з'являтися рідко, тому використання біграм буде достатнім. Цей тип моделі гарантує зберігання мінімуму необхідної для роботи інформації.

Щоб оцінити ймовірності N-грам, ми будемо використовувати оцінку максимальної правдоподібності (Maximum Likelihood Estimation - MLE), визначивши кількість N-грамів і нормалізувавши їх так, щоб вони знаходилися у проміжку від 0 до 1. Наприклад, якщо ми подивимося на біграми, тоді розрахунок ймовірності для наступних слів буде проведено за формулою (3.1):

$$P(w_n | w_{n-1}) = \frac{C(w_{n-1}w_n)}{C(w_{n-1})} \quad (3.1)$$

Де $P(w_n | w_{n-1})$ — ймовірність того, що слово W_n буде слідувати за W_{n-1} , і дорівнює загальній кількості біграм, поділений на загальну кількість біграм у вхідному масиві даних.

3.1.1 Попередня обробка тексту

Попередня обробка є важливим завданням і критичним кроком у перед обробкою тексту та пошуком інформації в ньому. В області Text Mining попередня обробка даних використовується для видалення цікавих і нетривіальних знань з неструктурованих текстових даних.

Існує ряд особливостей, які слід враховувати під час попередньої обробки або очищення даних, щоб отримати з них максимум корисної

інформації. Важливо, щоб обробка була виконана в правильному порядку, перш ніж втратити певну важливу інформацію. Наприклад, перевірка на наявність іноземного тексту має відбуватися до того, як символи, відмінні від ASCII, будуть видалені, оскільки це призведе до усунення чужомовних символів [36].

Завдання попередньої обробки складаються з поділу абзаців на окремі речення, переведення усього тексту в нижній регістр, видалення знаків пунктуації, цифр і символів, що не входять до ASCII, усунення нецензурної лексики, видалення зайвих пробілів та виконання перевірки орфографії, щоб виключити слова з повторюваними символами.

3.1.2 Токенізація

Токенізація – це процес розбиття потоку тексту на слова, фрази, символи або інші значущі елементи, які називаються лексемами. Метою лексемізації є дослідження слів у реченні. Список маркерів стає вхідним для подальшої обробки, наприклад аналізу або аналізу тексту. Токенізація корисна як в лінгвістиці (де це форма сегментації тексту), так і в інформатиці, де вона є частиною лексичного аналізу. Текстові дані – це лише блок символів на початку. Усі процеси пошуку інформації вимагають слова із набору даних. Отже, вимогою до програмного додатку є токенізація навчального масиву для передбачення вводу пацієнта. Це може здатися тривіальним завданням, оскільки текст уже зберігається в текстовому форматі на пристрої із встановленим додатком. Проте деякі проблеми все ж залишаються, як-от видалення розділових знаків та інших символів, таких як дужки, дефіси тощо. Основне використання токенизації — визначення значущих ключових слів.

Проблеми токенизації залежать від типу мови. Такі мови, як англійська та французька, називаються розділеними пробілами, оскільки більшість слів відокремлено один від одного пробілами. Такі мови, як китайська та тайська, називаються несегментованими, оскільки слова не мають чітких меж. Токенізація несегментованих мовних речень потребує додаткової лексичної та

морфологічної інформації. На лексемизацію також впливає система письма та типографічна структура слів.

3.1.3 Вилучення стоп слів

Багато слів у текстових масивах повторюються дуже часто, але по суті не мають користі при подальшій обробці, оскільки використовуються для з'єднання слів у реченні. Загальноприйнято розуміти, що стоп-слова не впливають на контекст або зміст текстових документів. Через їх високу частоту появи, їх присутність при аналізі тексту є перешкодою для розуміння змісту документів. Стоп слова – це дуже часто використовувані слова, такі як «і», «е», «це» тощо. Вони не є корисними при обробці текстових масивів. Тому їх необхідно видалити. Однак розробка такого списку стоп-слів є складною і неузгодженою між текстовими джерелами. Цей процес також зменшує кількість текстових даних та покращує продуктивність системи. Такі слова зустрічаються майже в кожному тексті, отже доцільно буде проводити їх вилучення з кожним новим масивом тексту.

3.2 Вимоги до розміру словника

Сучасна українська мова має близько 256 тисяч слів. Постає необхідність визначити скільки слів потрібно словнику прогнозованої моделі розроблюємого програмному додатку, щоб забезпечити достатнє покриття для успіху моделі. Згідно зі статистикою Національної Академії Наук України середній дорослий носій української мови використовує близько 20 000 слів. Перші за частотою вживання 25 слів використовуються в 33% щоденного письма, перші 100 слів зустрічаються в 50% письма дорослих і учнів, перші 1000 слів використовуються у 89% щоденного письма. Але для охоплення приблизно 95% поширених текстів, таких як блоги та новини, потрібно лише 3000 слів [3]. Словник такого розміру повинен відповідати вимогам до обмеження пам'яті зберігання розроблюємого програмного додатку.

Розглядаючи вхідні текстові дані, можна обчислити, скільки унікальних слів використовується для охоплення 50% і 90% тексту, підсумовуючи кількість найбільш часто зустрічаємих слів і розділяючи їх на загальну кількість слів, поки не досягнемо 50 і 90%.

Важливо створити програму для передбачення слів, яка працює швидко та використовує мінімум пам'яті, оскільки кінцевою метою є запуск її на комп'ютері, або мобільному пристрої з швидким часом відгуку для користувача. Якщо попередньо обробити вхідні дані та оптимізувати словник для найбільш часто використовуваних слів, це, відповідно, зменшить розмір необхідного нам словника даних і сховища коду.

3.3 Вимоги до користувацького інтерфейсу

У сучасних програмних додатках програмісти зобов'язані приділяти велику увагу не тільки гарному вигляду, а й максимальній доступності додатку для усіх потенційних користувачів. Оскільки даний програмний додаток розроблюється з акцентом на використання паралізованими пацієнтами, особливу увагу потрібно звернути на легкість використання, можливість гнучкого налаштування, легку читаємість та не перегруженість усіх елементів інтерфейсу.

По статистиці більша частина пацієнтів хворих на параліч тіла це люди похилого віку. Дуже часто вони не є досвідченими користувачами персональних комп'ютерів, отже складні інтерфейси додатку можуть налякати їх та стати перешкодою для спілкування та знайомства із програмою.

Моїм завданням як розробника є зробити інтерфейс додатку максимально дружнім, зрозумілим та інтуїтивним для користувача. Щоби і літні люди, і маленькі діти, які вже вміють читати та писати, могли з легкістю зрозуміти основні принципи роботи із додатком, та освоїти його за максимально коротким проміжком часу.

Інклюзивний дизайн допомагає створювати продукти, які обслуговують якомога більше людей. Хоча доступність є основною метою, інклюзія означає набагато більше. Це дає змогу людям з різними характеристиками використовувати ваш продукт у різних середовищах. Важливо, особливо коли ви створюєте дизайн для мільйонів людей — створювати різні способи участі людей у досвіді.

Під час створення даного програмного додатку метою було дотримання основних принципів інклюзивного дизайну, а саме:

- **Надання порівнянного досвіду**

Переконайтеся, що ваш інтерфейс забезпечує порівнянний досвід для усіх користувачів, щоб вони могли використовувати додаток у спосіб, який відповідає їхнім потребам, не зменшуючи якість його вмісту та функціональність.

Незалежно від обставин, вибору чи контексту усі люди різноманітні. Оскільки вони використовують різні підходи та інструменти для читання та роботи з інтерфейсами, те, що інтерфейс пропонує кожному користувачеві, має бути порівнянним за цінністю, якістю та ефективністю.

- **Пам'ятати про різні обставини**

Люди використовують ваш інтерфейс у різних ситуаціях. Переконайтеся, що ваш інтерфейс забезпечує цінний досвід для людей, незалежно від їхніх обставин. Різні життєві ситуації можуть мати різний вплив на користувача. Для тих, кому взаємодія із додатком від самого початку складна, наприклад, для людей з обмеженими можливостями, цей вплив може ускладнити використання.

- **Будьте послідовними**

Використовуйте знайомі конвенції та застосовуйте їх послідовно.

Знайомі інтерфейси запозичуються з усталених шаблонів. Вони повинні використовуватися послідовно в інтерфейсі, щоб посилити їх значення та

призначення. Це має застосовуватися до функціональності, поведінки, редакції та презентації. Ви повинні показувати ті самі речі однаково, а користувачі повинні мати можливість робити ті самі речі у такий самий спосіб.

- **Надайте користувачу контроль**

Переконайтеся, що люди контролюють додаток. Люди повинні мати доступ до вмісту та взаємодіяти з ним своїм улюбленим способом.

Не пригнічуйте та не вимикайте можливість змінювати стандартні налаштування інтерфейсу та платформи, такі як орієнтація, розмір шрифту, масштабування та контраст. Крім того, уникайте змін вмісту, які не були ініційовані користувачем, якщо немає способу керувати ним.

- **Пропонуйте вибір**

Подумайте про надання людям різних способів виконання завдань, особливо складних або нестандартних.

Часто існує кілька способів виконання завдання. Ви не можете припустити, який спосіб може надавати комусь перевагу. Надаючи альтернативи для виконання завдань, ви пропонуєте людям вибір, який підходить їм і їхнім обставинам на даний момент.

- **Розставте пріоритети вмісту**

Допоможіть користувачам зосередитися на основних завданнях, функціях та інформації, розставляючи їх пріоритети в вмісті та макеті.

Інтерфейси може бути важко зрозуміти, коли основні функції не чітко представлені та не визначені пріоритетами. Сайт або програма можуть надавати багато інформації та функцій, але люди повинні мати можливість зосередитися на чомусь одному. Визначте основне призначення інтерфейсу, а потім вміст і функції, необхідні для досягнення цієї мети.

- **Додайте цінність**

Майте на увазі цінність функцій і те, як вони покращують роботу різних користувачів.

Функції повинні додавати цінність користувацькому досвіду, забезпечуючи ефективні та різноманітні способи пошуку вмісту та взаємодії з ним. Подумайте про функції пристрою, такі як API голосу, геолокації, камери та вібрації, а також те, як інтеграція з підключеними пристроями чи другим екраном може забезпечити вибір.

Ці принципи інклюзивного дизайну спрямовані на те, щоб поставити людей на перше місце. Йдеться про проектування для потреб людей з постійними, тимчасовими, ситуативними або змінними вадами.

Вони призначені для того, щоб надати всім, хто бере участь у розробці та розробці веб-сайтів і додатків – дизайнерам, професіоналам із досвіду користувачів, розробникам, власникам продуктів, творцям ідей, новаторам, художникам і мислителям – широкий підхід до інклюзивного дизайну.

3.4 Проектування програмного продукту

Розробку графічного інтерфейсу користувача було вирішено здійснювати за допомогою мови програмування Python.

Для опису процесів які є невід’ємною частиною розроблюємого програмного додатку було розроблено та реалізовано контекстні діаграми додатку. З їх допомогою можна визначити вхідні та вихідні дані, які потрапляють в систему та виходять з неї як результат її роботи. Також описуються додаткові допоміжні сторонні засоби та методи, включаючи сторонні бібліотеки Dlib та OpenCV для застосування алгоритмів комп’ютерного зору та відстежування погляду.

Даними що подаються на вхід до програмного додатку є сигнал з вебкамери комп’ютера встановленого біля пацієнта, який буде ним користуватися. На виході ми отримуємо повідомлення від пацієнта яке буде оброблюватися лікарем/оператором програмного додатку.

Контекстну діаграму розроблюємого додатку зображено на рис.3.1.



Рисунок 3.1. Контекстна діаграма програмного додатку

3.4.1. Декомпозиція 1-ого рівня

Кожен з блоків цієї контекстної діаграми можна відобразити на іншій діаграмі, що логічно стане складовою частиною загальної діаграми. Цей прийом називається декомпозиція, тобто створення нових підзадач із оригінальної задачі задля спрощення процесу розробки та кращої його візуалізації.

Процес роботи даного програмного додатку можна відобразити як 4 окремі етапи (рис.3.2):

- Визначення положення обличчя та очей пацієнта на відеосигналі;
- Відстежування напряму погляду та моргання очима;
- Додавання обраного користувачем символу чи фрази до повідомлення;

- Додавання нового набору спрогнозованих слів до віртуальної клавіатури користувача;

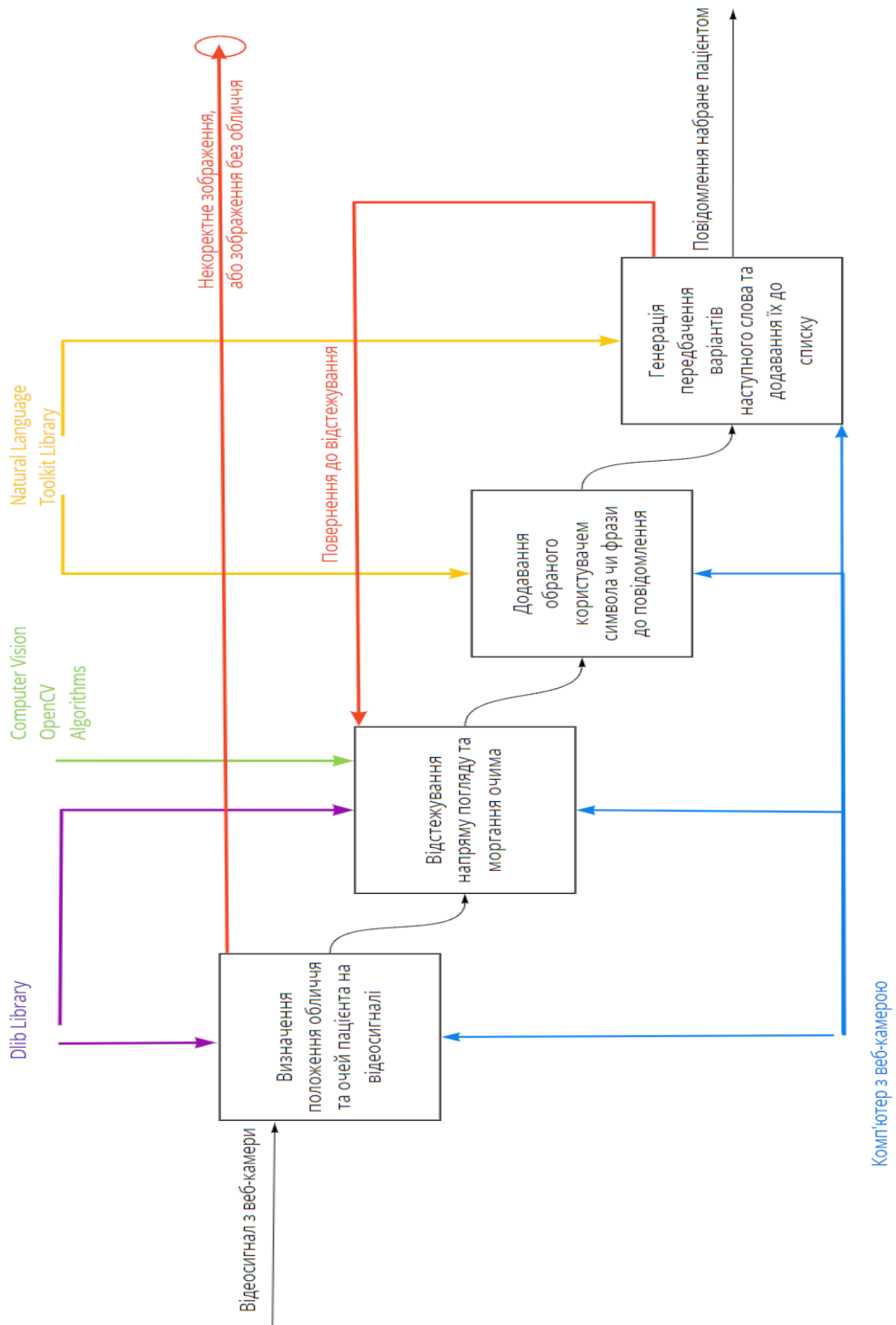


Рисунок 3.2. Декомпозиція першого рівня

3.4.2. Декомпозиція 2-ого рівня

Діаграми декомпозиції другого рівня представлені для ключових етапів розробки даного програмного додатку, а саме «Визначення положення обличчя та очей пацієнта на відеосигналі» (рис.3.3), «Відстежування напрямку погляду та моргання очима» (рис.3.4) та «Прогнозування наступного набору слів» (рис.3.5).

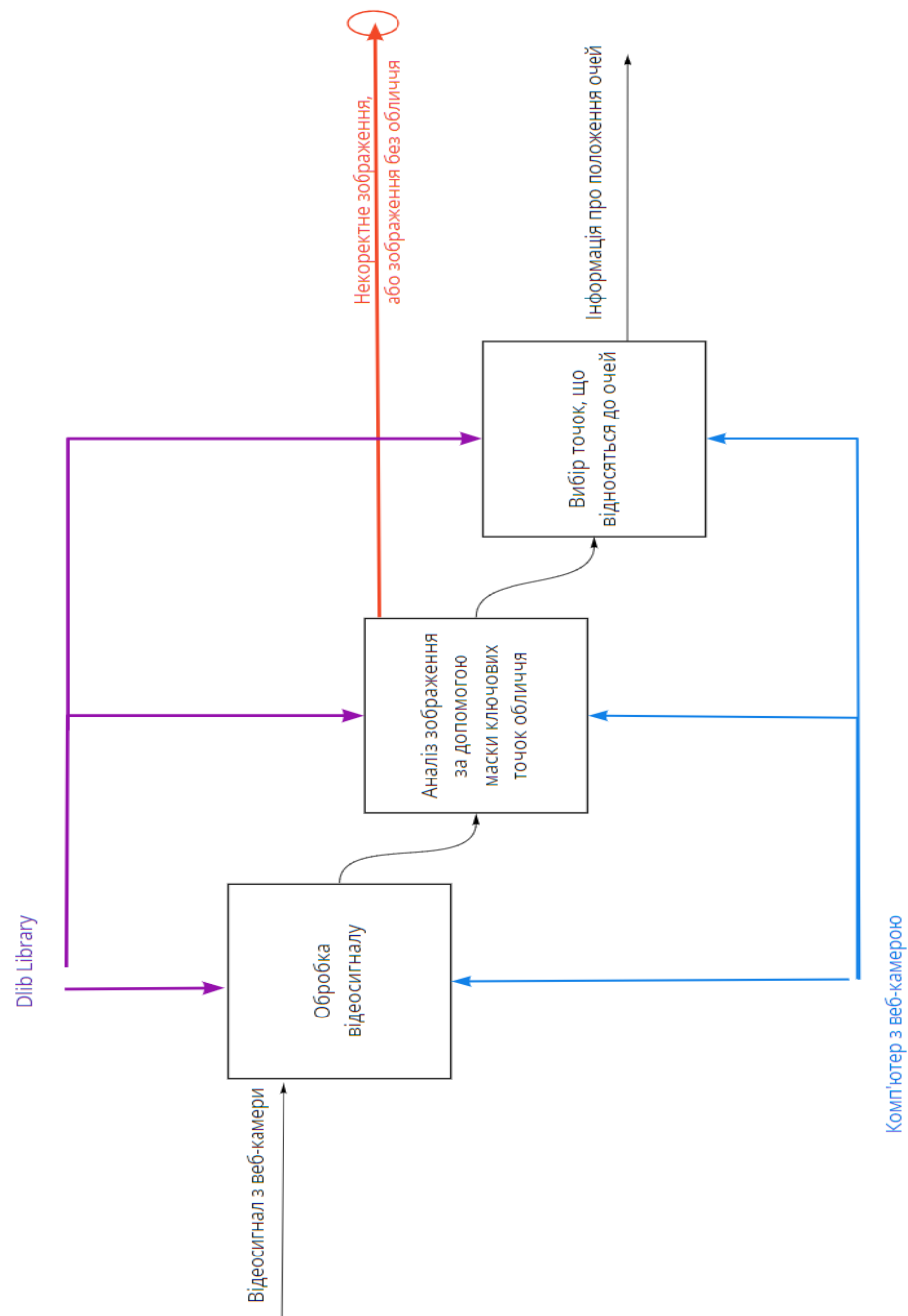


Рисунок 3.3. Діаграма декомпозиції другого рівня для етапу «Визначення положення обличчя та очей пацієнта на відеосигналі»

Діаграма представлена на рис.3.3 описує етап визначення положення обличчя та очей пацієнта на відеосигналі.

Цей етап складається з трьох підетапів:

- обробка сигналу з веб-камери;
- аналіз зображення за використовуючи маску ключових точок обличчя;
- вибір точок, що належать до очей.

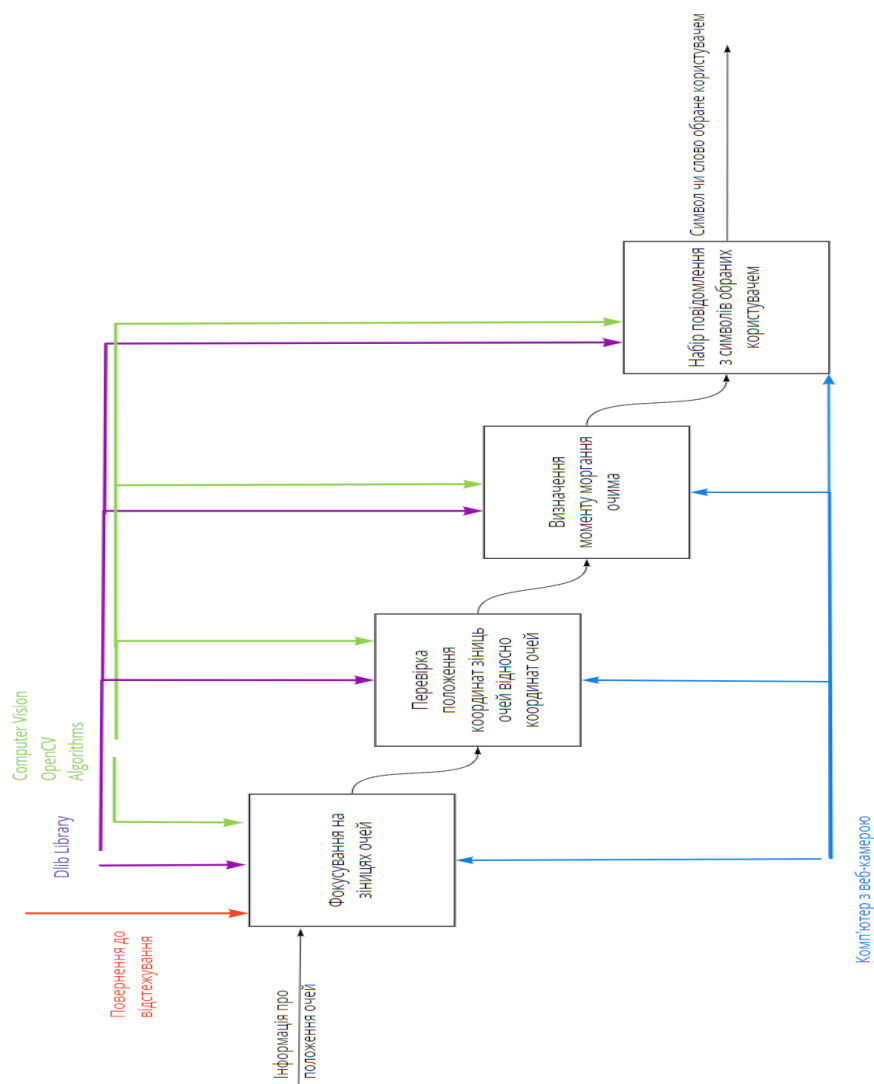


Рисунок 3.4. Діаграма декомпозиції другого рівня для етапу «Відстежування напрямку погляду та моргання очима»

Діаграма представлена на рис.3.4 описує етап відстежування напрямку погляду та моргання очима, що складається з таких етапів:

- фокусування на зіницях очей;
- перевірка положення координат зіниць відносно координат очей;
- визначення моменту моргання очима;

визначення функції, яка була показана пацієнту під час моргання

Діаграма представлена на рис.3.5 описує етап генерації та передбачення варіантів наступного слова та додавання їх до списку слів.

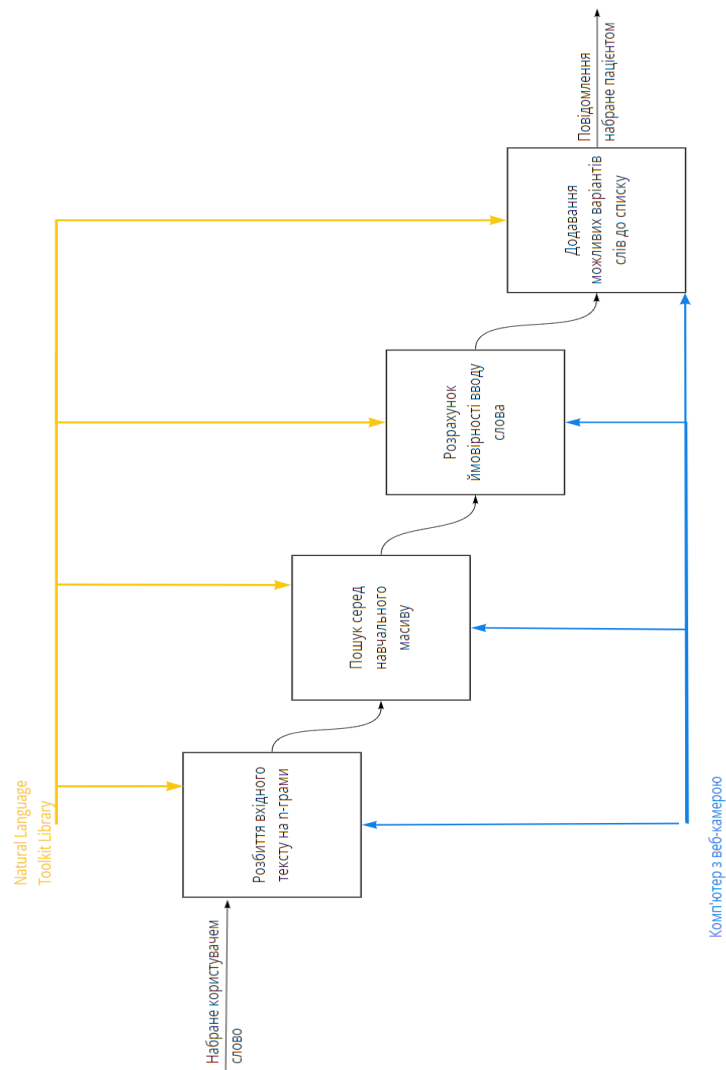


Рисунок 3.5. Діаграма декомпозиції другого рівня для етапу «Відстежування напрямку погляду та моргання очима»

Цей етап складається з чотирьох підетапів:

- Розбиття вхідного тексту на n-грами;
- Пошук серед навчального масиву;

- Розрахунок ймовірності вводу слова;
- Додавання можливих варіантів слів до списку.

3.4.3. Діаграма дерева вузлів

Діаграма дерева вузлів допомагає краще відобразити ієрархію процесів, що відбуваються у даній системі. Усі процеси та підпроцеси даної програми зображено на рис.3.6.

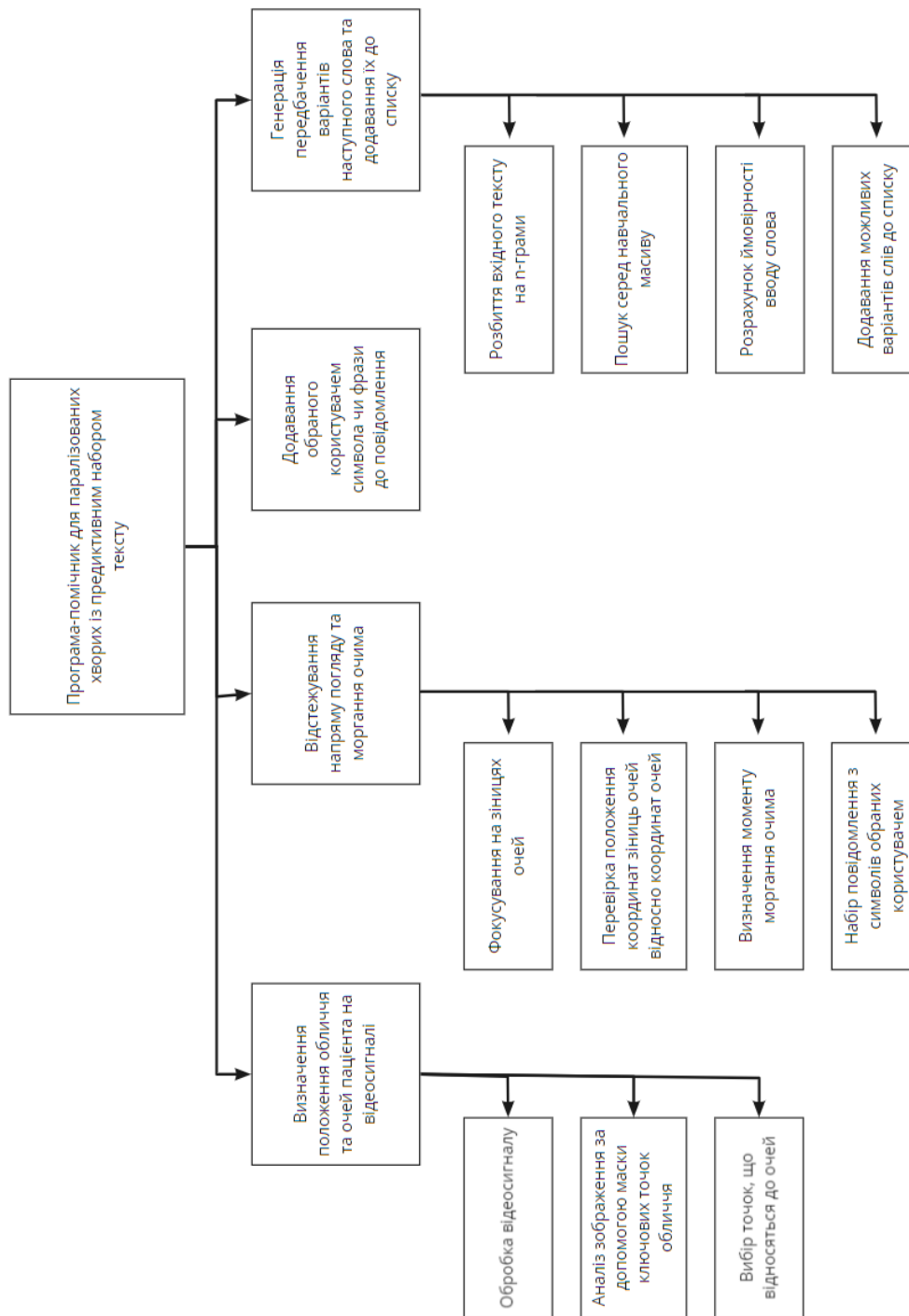


Рисунок 3.6. Діаграма дерева вузлів

3.4.4. Діаграма послідовності

На рис.3.7 наведено діаграму послідовності. У верхній частині діаграми наведено перелік об'єктів (логічні сутності), які взаємодіють між собою у процесі виконання сценарію. Часова шкала на даній діаграмі спрямована згори донизу.

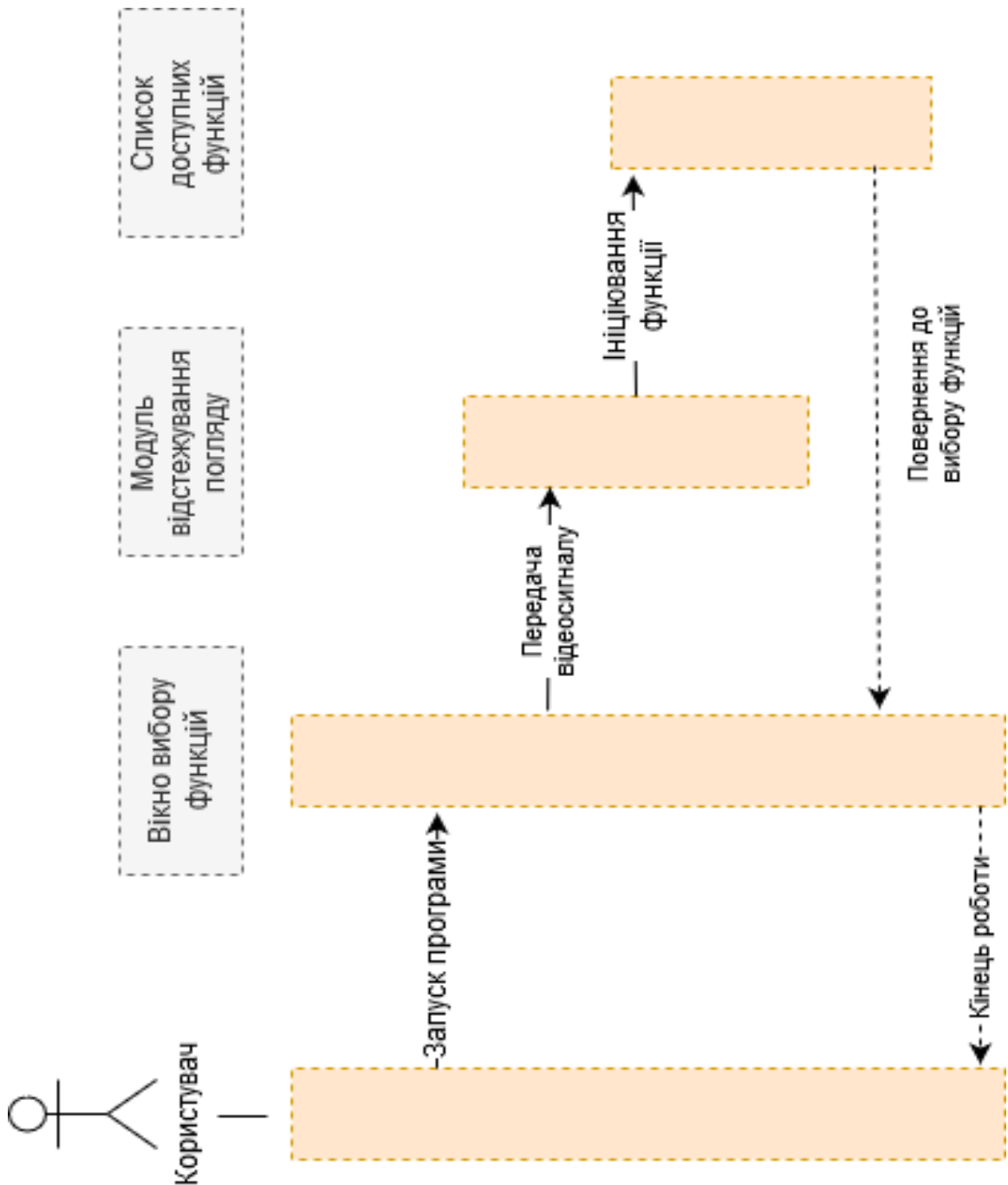


Рисунок 3.7. Діаграма послідовності

3.4.4. Діаграма станів

Діаграма станів використовуються для моделювання динамічного характеру системи. Вона визначає різні можливі стани об'єкта під час його життєвого циклу та події, що можуть мати на них вплив. Головною метою діаграми станів є моделювання життєвого циклу від створення до закінчення. Діаграма станів програмного додатку зображена на рис.3.8.

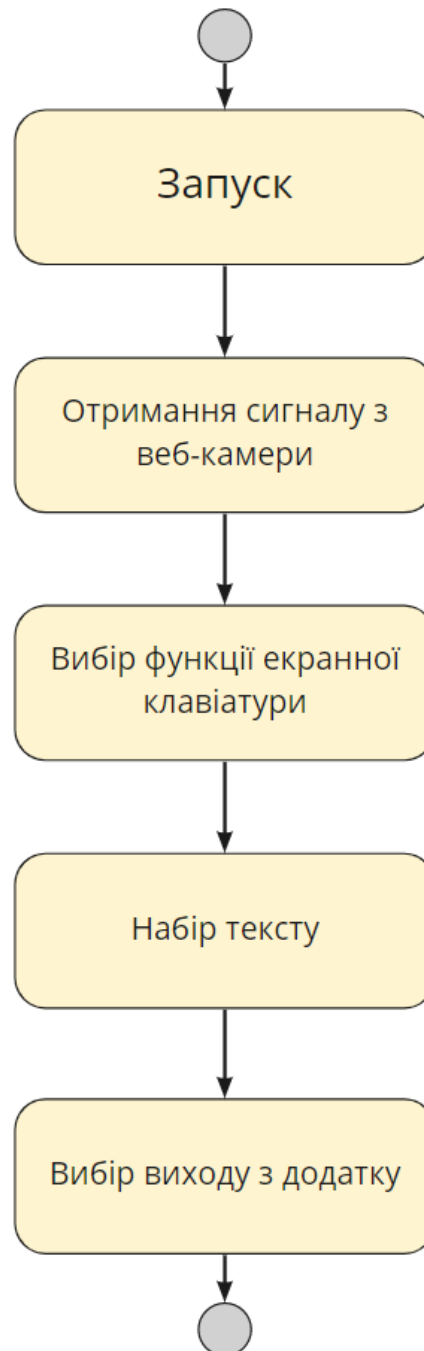


Рисунок 3.8. Діаграма станів

Висновки до розділу 3

В розділі 3 докладно розглянуто вимоги до розроблюємого програмного додатку. До них відноситься використовувана модель, вимоги до нормалізації вхідних даних, вимоги до розміру зберігаємого в пам'яті пристрою словника. Ці вимоги безпосередньо впливають на усі подальші етапи проектування та розробки додатку. Було прийнято рішення що словник додатку буде містити в собі близько 3000 унікальних слів, це має забезпечити покриття близько 95% слів необхідних у повсякденном житті. Було розроблено контекстну діаграму, описано діаграму декомпозиції першого та другого рівня, діаграму дерева вузлів, діаграму послідовності та діаграму станів.

РОЗДІЛ 4

РЕАЛІЗАЦІЯ ПРОГРАМНОГО ДОДАТКУ З ФУНКЦІЄЮ ПРЕДИКТИВНОГО НАБОРУ ТЕКСТУ

4.1 Теоретичне обґрунтування роботи програми

Головною метою роботи є створення програмного додатку, який може значно полегшити процес спілкування із оточенням для паралізованих пацієнтів та сприяти створенню нових каналів спілкування із ними. Акцент при розробці ставиться на використання технології відстежування погляду та предиктивного тексту.

Ці задачі відносяться до класу систем розпізнавання образів та систем передбачення набору тексту. При розпізнаванні обличчя воно повинне розпізнаватися на будь-якому фоні, незалежно від факторів, що можуть впливати на якість зображення, допоки обличчя можна цілком розгледіти на вхідному зображенні. Система предиктивного набору тексту, у свою чергу, повинна завжди намагатися спрогнозувати та надати на вибір користувача найбільш оптимальні та відповідні до поточного контексту слова або фрази для подальшого їх вибору. Можливість навчання цих систем може сприяти покращенню їх роботи при використанні додатку.

Базовим алгоритмом, що буде використовуватися для визначення обличчя пацієнта під час роботи програми є метод гістограм орієнтованих градієнтів або HOG. Для вирішення задачі розпізнавання образів він використовує спеціальні дескриптори, та вважається одним із найперспективніших методів для задач такого типу [56].

Принцип його роботи, це ітеративне розділення зображення на клітини для подальшого визначення напрямів градієнтів яскравості, які далі будуть використані як вектори ознак, що несуть у собі данні про можливу форму та тип об'єкту на зображенні.

Перевага цього способу в тому, що напрями градієнтів, а точніше місцеположення локальних тінювих зон на зображенні, напряму пов'язані із особливостями рельєфу поверхні спостерігаємого об'єкту (в даному випадку із особливостями будови людського обличчя), які дуже добре визначають його індивідуальність і надають нам змогу правильно класифікувати його при подальшій роботі із зображенням [1].

Окрім цього, погана якість вхідного зображення, шуми, або погане освітлення, не сильно впливають на градієнти, а отже й на якість їх визначення та подальшу роботу. Таким чином цей метод якомога краще задовільняє вимоги до системи розпізнавання обличчя, поставленні при розробці програмного додатку. На рис.4.1. зображено приклад роботи алгоритму.

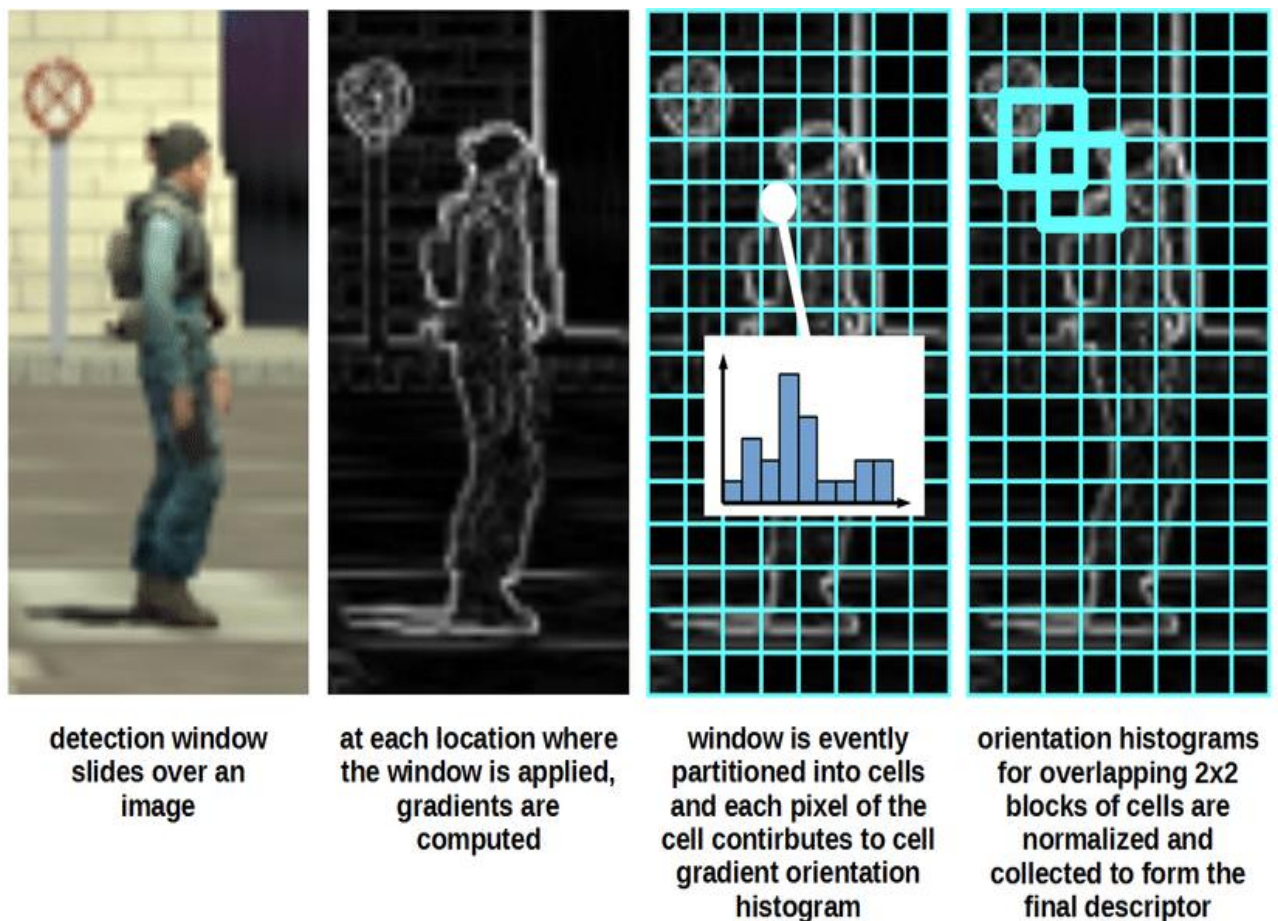


Рисунок 4.1. Загальний принцип дії методу гістограм орієнтованих градієнтів [70].

Наразі цей метод застосовується для різноманітних задач потребуючих розпізнавання обличчя, або будь яких інших об'єктів. А у зв'язку із останніми подіями у світі він все частіше використовується для визначення факту носіння масок на обличчі людьми у публічних місцях, задля запобігання поширення коронавірусної інфекції.

У даній роботі він виступає в ролі базового алгоритму для співставлення обличчя пацієнта із усередненим шаблонним зображенням, в якому чітко відокремлені ключові точки – орієнтири по яким відбувається відстежування отримання інформації про положення обличчя.

У якості програмного засобу при створенні додатку було застосовано реалізацію бібліотеки Dlib на мові програмування Python. Її алгоритми використовують метод гістограм орієнтованих градієнтів, а отже можуть використовуватися для вирішення задачі визначення обличчя на зображенні з веб-камери. Отримання зображень з веб-камери у реальному часу та їх попередня обробка здійснюється за допомогою бібліотеки CV2 для мови Python.

Для правильної роботи системи предиктивного набору тексту використовується метод n-gram, який допомагає виокремити значимі смислові одиниці із тексту що вводить пацієнт, та знайти ймовірність вибору пацієнтом наступного слова із масиву усіх попередніх слів відомих програмі. Це можуть бути як слова введені пацієнтом під час цього сеансу, так і слова та речення із текстового файлу, куди вся інформація введена пацієнтом потрапляє після завершення роботи додатку.

Основною бібліотекою для роботи із текстовими даними на мові програмування Python, яка включає в себе методи для роботи з n-грамами було обрано NLTK - Natural Language Toolkit. З її допомогою відбувається попередня обробка даних – нормалізація, токенизація, приведення тексту до одного регістру, вилучення стоп слів. А також подальший його аналіз, розбиття тексту на n-грами, визначення ймовірностей вибору користувача того чи іншого слова наступним під час процесу набору тексту. Цей інструмент

забезпечує максимальну оптимальність сгенерованих передбачень під час роботи програми.

4.2 Етапи роботи додатку

На першому етапі зображення передаємі з веб-камери пристрою із простору кольорів RGB конвертуються у чорно-білий простір - GRAY. Таке перетворення позитивно впливає на ефективність подальших алгоритмів обробки. Приклад такого перетворення продемонстровано на рис.4.2.



Рисунок 4.2. Перетворення зображення з простору кольорів RGB у GRAY

Наступними кроками є послідовна передача чорно-білих зображень на вхід функцій `get_frontal_face_detector` та `shape_predictor` бібліотеки Dlib. Використання цих двох функцій у поєднанні дає нам змогу отримати дані про

поточні координати обличчя та зокрема про координати ключових точок – орієнтирів обличчя (рис.4.3.) [15].

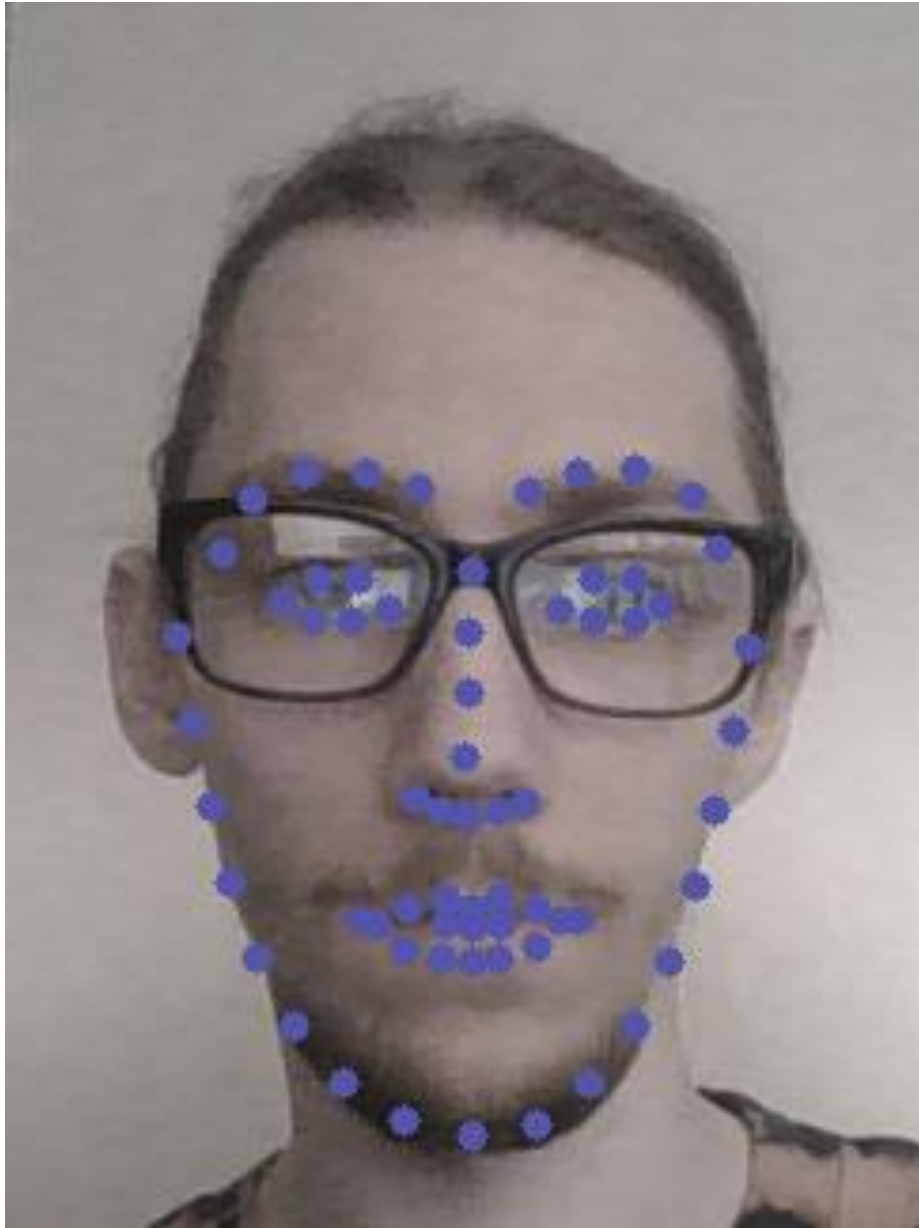


Рисунок 4.3. Визначення ключових точок обличчя на зображенні

Алгоритм визначення цих точок заснований на основі методу HOG, технологій нейронних мереж, ансамблю вирішальних дерев регресії. Алгоритм попередньо навчався (тренувався) на навчальних наборах помічених вручну ключових точок обличчя на попередньо зібраних в базу

даних зображеннях. Як результат, визначення положення ключових точок відбувається достатньо достовірно.

За орієнтири зазвичай приймаються: рот, права брова та ліва брова, праве та ліве око, ніс, щелепа. Використаний в даній роботі алгоритм визначає локалізацію 68-и ключових точок (рис.4.4).

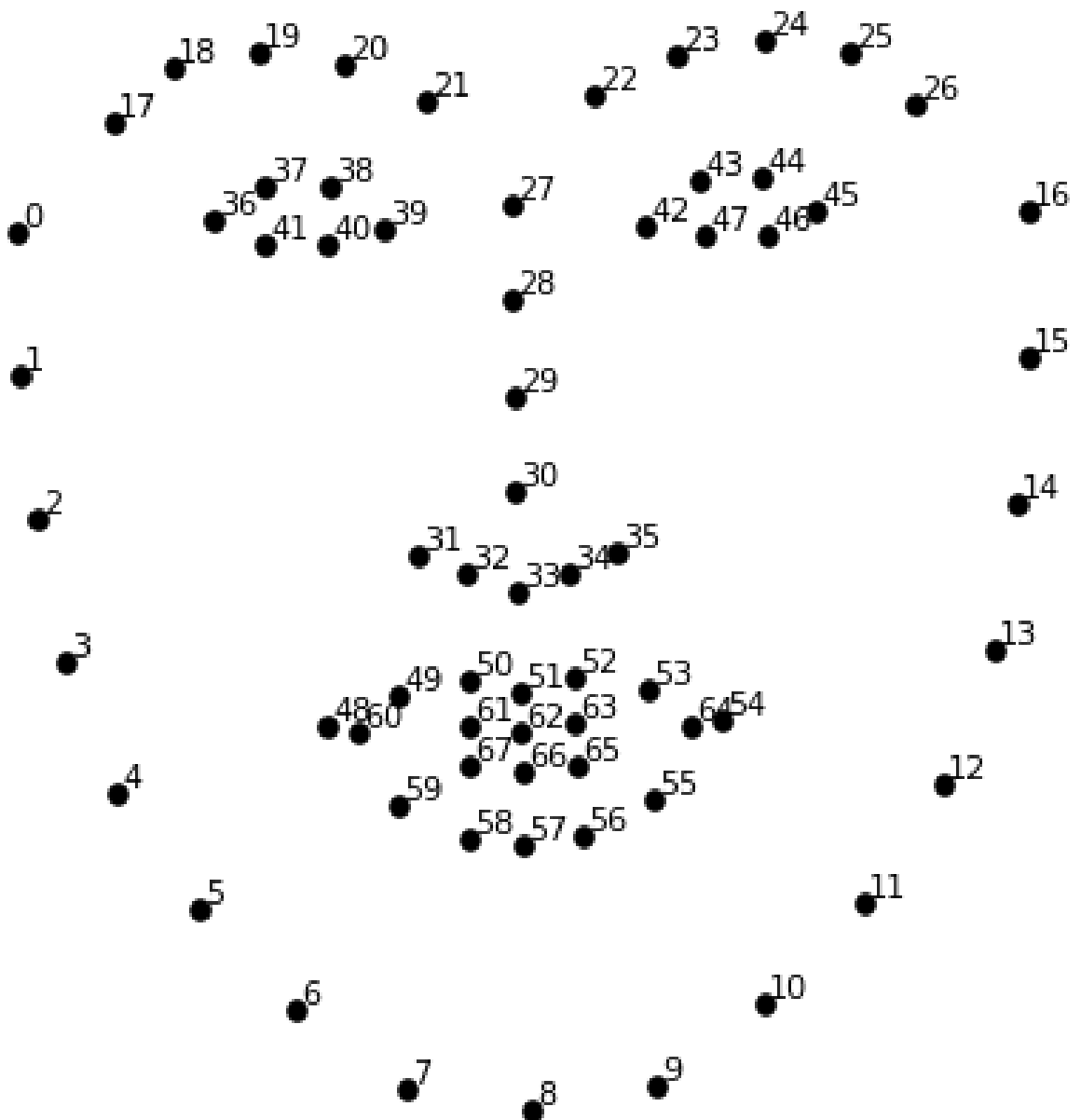


Рисунок 4.4. Локалізація 68-и ключових точок – орієнтирів обличчя

У результаті роботи алгоритму ми отримуємо масив пронумерованих міток із відповідними координатами їх розташування. Далі вже досить просто визначити положення частин обличчя та відстані між орієнтирами.

Наприклад, з рис. 4.4 видно, що важливими є групи точок 36-41, та 42-47, які відповідають за праве та ліве око.

Маючи координати точок з цих груп, ми можемо вирахувати відстані між ними. Головна ідея полягає у тому, що обраховуючи відстані між крайніми лівими та крайніми правими точками очей (точки 36:39 та 42:45), можна помітити, що вони залишаються майже незмінними незалежно від того, чи відкриті в людини очі чи ні. А відстані від середньої верхньої точки ока до середньої нижньої точки (у нашому випадку це середнє арифметичне між координатами точок 37-38:41-40 та 43-44:47-46) змінюються, коли користувач закриває очі.

Можна провести з'єднуючі лінії між відповідними ключовими точками на очах (рис.4.5).



Рисунок 4.5. Візуалізація відстаней між крайніми ключовими точками очей.

Порівнюючи співвідношення між довжинами горизонтальних та вертикальних ліній, можна визначити, відкриті чи закриті очі у людини на зображенні у кожен момент часу. Коли очі відкриті це співвідношення

становить 3-4, збільшення цього числа буде свідчити про те, що очі пацієнта зараз закриті [1].

Вертикально «поділивши» кожне око на ліву та праву частину ми можемо визначити напрям погляду порівнявши відношення між чорними та білими пікселями у кожній половині ока.

4.2 Взаємодія з програмним додатком

Віртуальна клавіатура поділена на дві частини для зручності використання. Кожна приблизно відповідає лівій та правій частині справжньої комп'ютерної клавіатури. Тому на головному екрані користувач бачить два пункти на вибір – «ліво» та «право». Щоб вибрати ту чи іншу половину клавіатури пацієнту просто треба направити свій погляд в ту чи іншу сторону. Стартовий екран програми зображений на рис.4.6.

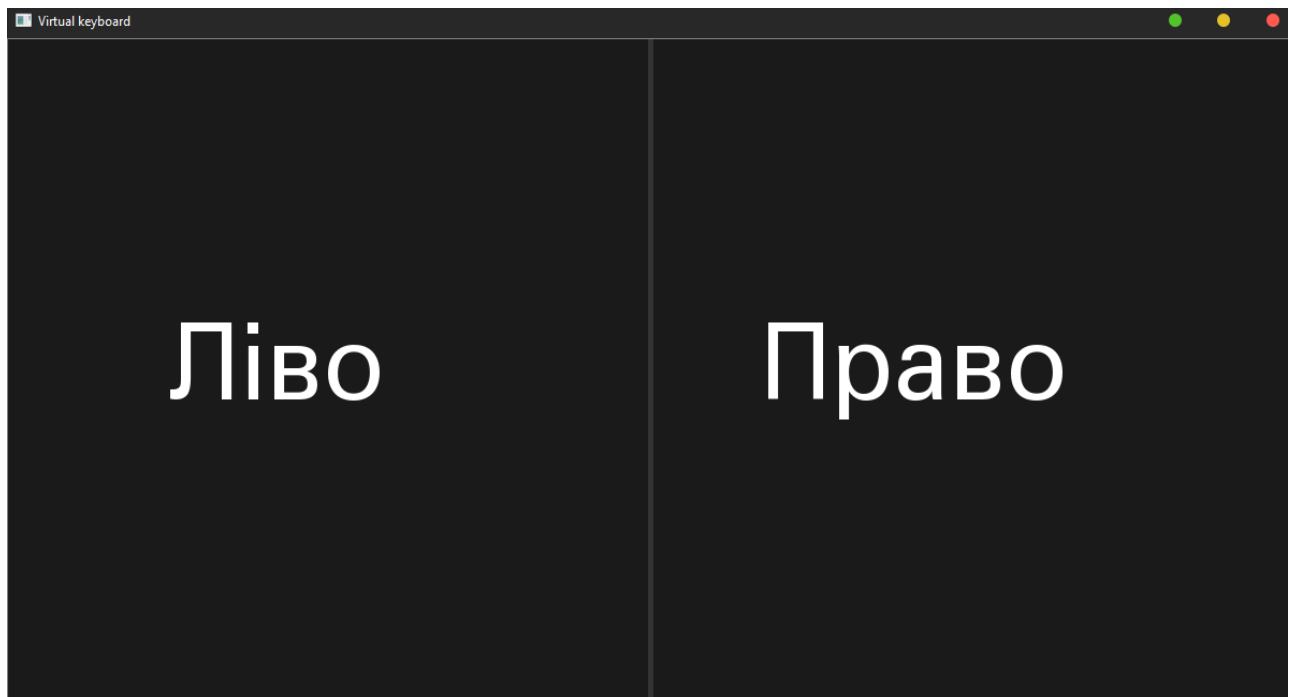


Рисунок 4.6. Головний екран програми.

Після вибору бажаної частини клавіатури перед користувачем з'являється віртуальна клавіатура, що повторює за своїм виглядом ліву чи праву половину реальної клавіатури. Значок стрілки виконує функцію кнопки назад. При виборі слова «Вихід» сеанс користування програмним додатком буде завершено.

Нижній рядок клавіатури завжди містить в собі слова сгенеровані за допомогою функції передбачення тексту. Ці слова генеруються після введення користувачем кожного нового слова. Відстежування моменту коли користувач закінчив писати поточне слово відбувається дуже легко, після кожної введеної літери програма перевіряє чи не було це пробілом, оскільки в українській мові саме пробіл свідчить про закінчення одного слова і початок наступного.

Як тільки це сталося, програма нормалізує введений користувачем текст за вимогами описаними у розділі 3. Після розбиття на n-грами, алгоритм аналізує введений текст, та намагається знайти найращі варіанти для наступного слова використовуючи навчальний масив даних та текстовий файл із повною історією вводу для данного користувача додатку (рис.4.7).

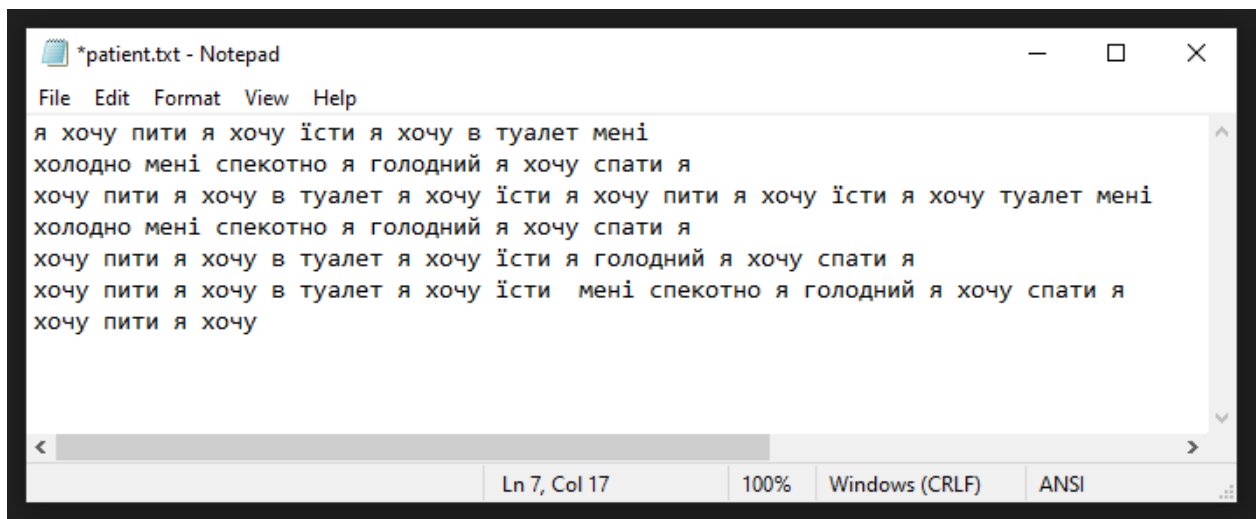


Рисунок 4.7. Тестовий текстовий файл із історією набору тексту користувачем

Після обчислення найвірогідніших варіантів слів, 5 з них додаються до віртуальної клавіатури, для того щоб користувачу було максимально зручно та легко ввести наступне слово не проходячи процес його набору літера за літерою.

Переміщення між літерами та словами відбувається автоматично. Кожна з них по черзі підсвічується на декілька секунд. Якщо пацієнт хоче обрати цю літеру або слово, йому необхідно просто заплющити очі на три секунди. Як тільки програма фіксує закриття очей на заданий проміжок часу – літера додається до попередньо набраних пацієнтом символів. При необхідності проміжок часу необхідний для фіксування додатком закриття очей можна збільшити, чи зменшити, і таким чином налаштувати його під кожного конкретного пацієнта. Такий інтервал необхідний щоб виключити можливість помилкового вибору через природню потребу пацієнта час від часу моргати. На рис.4.8 зображено одну з половин віртуальної клавіатури програмного додатку.

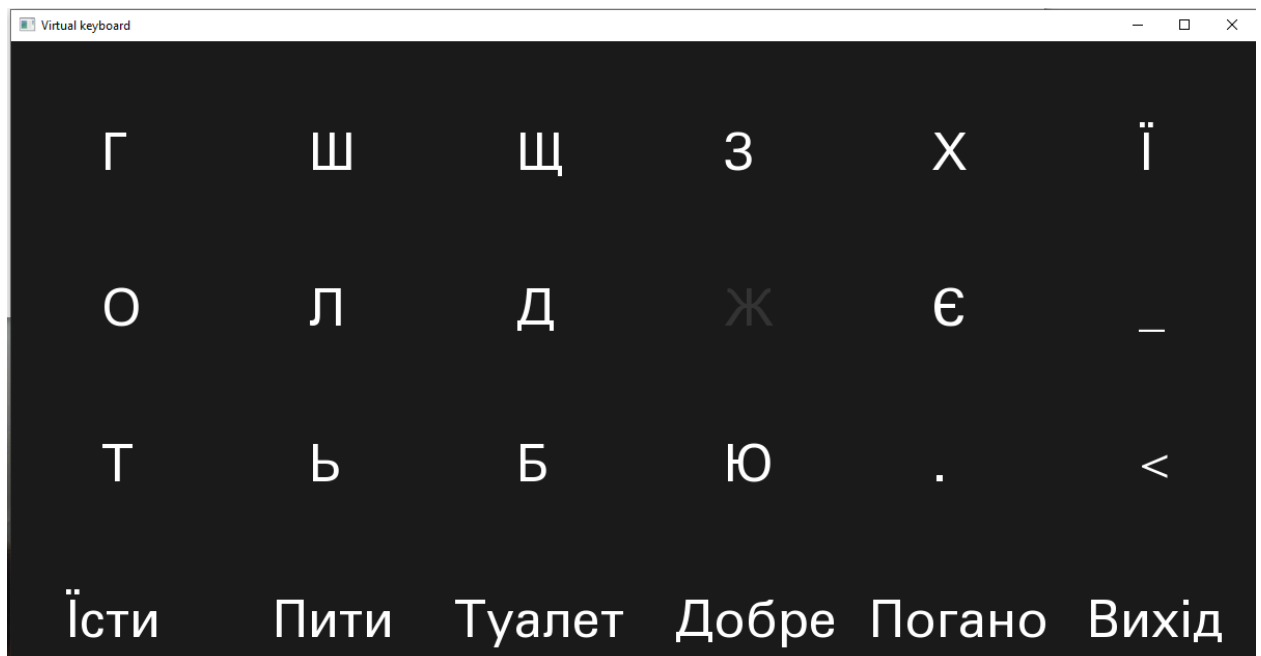


Рисунок 4.8. Права половина віртуальної клавіатури

Після цього вибір пацієнта з'являється у окремому вікні та додається до попередньо набраного тексту. Таким чином оточення пацієнта може бачити те що він написав. Вікно з набраним текстом зображено на рис.4.9.



Рисунок 4.9. Приклад фрази надрукованої за допомогою погляду

Для того щоб пацієнт міг бачити себе та розуміти, чи правильно програма визначає момент коли він моргає очима, а лікарям було простіше правильно налаштувати положення камери та монітора перед пацієнтом, присутнє додаткове вікно програми. Воно показує відеосигнал з вебкамери та положення очей в конкретний момент часу. Коли індикатори навколо очей червоного кольору – очі відкриті. Якщо індикатори стають зеленими – пацієнт заплющив очі.

Також в нижній частині вікна є індикатор-полоса, яка заповнюється коли програма фіксує закриття очей. Коли вона повністю зафарбовується відбувається вибір. Це зроблено для додаткової зручності пацієнта. Таким чином йому буде набагато простіше розуміти коли програма коректно считує закриття очей і при необхідності відкорегувати ці параметри у настройках програмного додатку. Додаткове вікно програми зображено на рис.4.9.

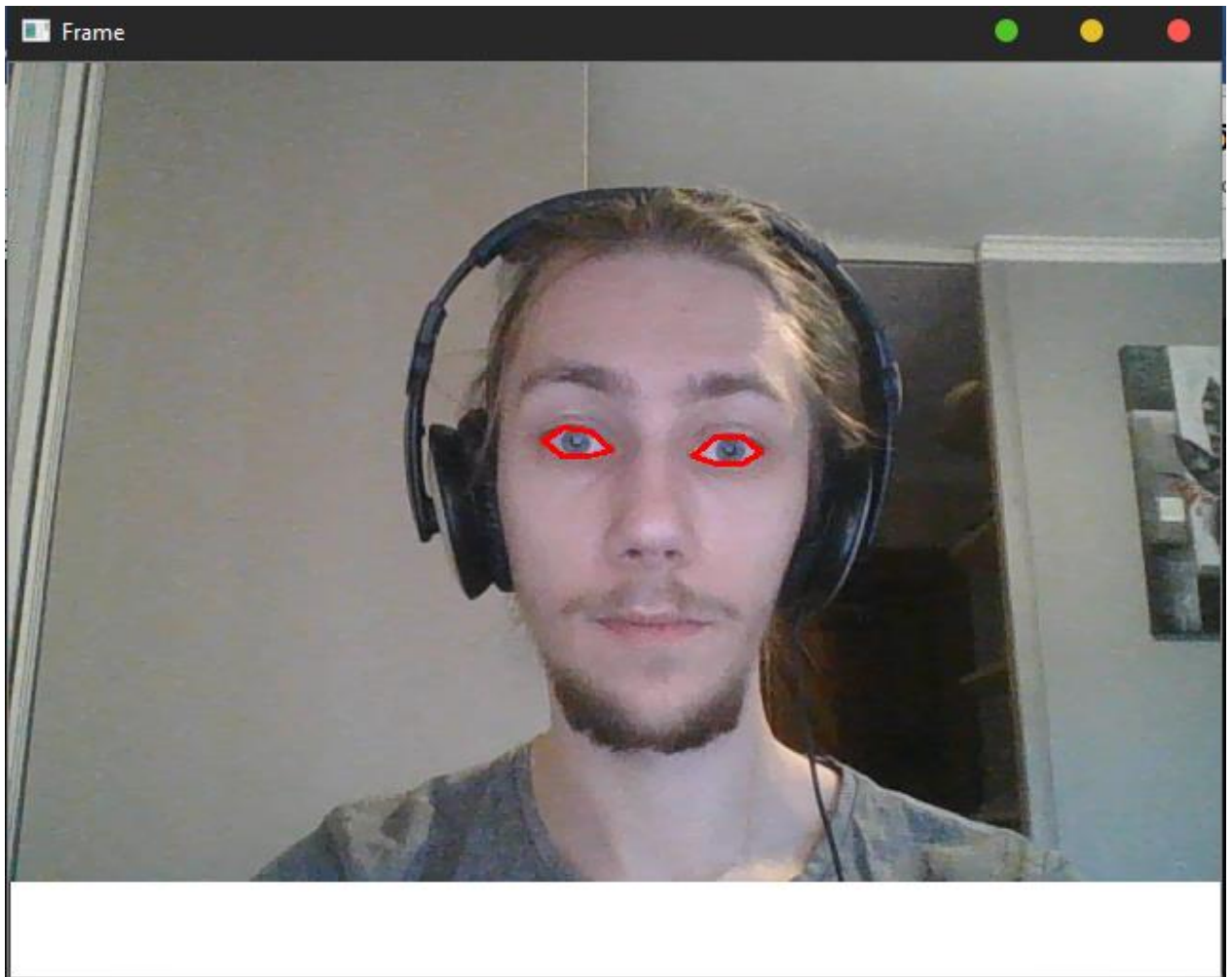


Рисунок 4.9. Додаткове вікно з визначеним положенням очей та полосою індикатором

Висновки до розділу 4

В роботі запропоновано реалізацію варіанту створення каналу спілкування для пацієнтів з втратою мови, обтяженою обмеженням рухливості. Застосовано комп'ютерні алгоритми локалізації ключових точок – орієнтирів на обличчі в режимі реального часу. Створено систему предиктивного набору тексту, що дозволяє пацієнту набагато швидше та зручніше користуватися програмним додатком та спілкуватися із своїм оточенням. Розроблено програму, що дозволяє паралізованим пацієнтам спілкуватися з оточення шляхом набору тексту на віртуальній клавіатурі за допомогою погляду та передбачує можливі наступні варіанти вводу. Додаток

розроблено згідно принципів інклюзивності з метою максимальної доступності додатку для усіх потенційних користувачів.

РОЗДІЛ 5

СТАРТАП АНАЛІЗ ПРОЕКТУ

Стартап-проект за темою магістерської дисертації є продовженням комплексного диплому бакалавра [1], який було під керівництвом доцента кафедри Біомедичної Інженерії Соломіна Андрія В'ячеславовича (теперішнього наукового керівника).

Метою даного розділу є проведення маркетингового аналізу стартап-проекту для визначення потенційної можливості його ринкового впровадження та можливих напрямів реалізації цього впровадження.

5.1 Резюме проекту

Головною ідеєю є розробка універсального програмного додатку-помічника для хворих із різними ступенями паралічу тіла, з використанням сучасних алгоритмів розпізнавання образів та відстежування погляду. Додаток може бути встановлено на будь який персональний комп'ютер або ноутбук з веб-камерою, з її допомогою додаток буде постійно відстежувати погляд пацієнта. Завдяки цьому пацієнти зможуть спілкуватися із оточуючими через набір тексту на віртуальній екранній клавіатурі вибираючи потрібні літери очима.

Додаток буде реалізовано на мові програмування Python із використанням сучасних бібліотек машинного навчання та комп'ютерного зору.

Основними споживачами продукту у майбутньому стануть державні та приватні лікарні, які зможуть створювати сучасні палати для паралізованих хворих, де буде встановлено даний програмно-апаратний комплекс. Планом на перший рік після реалізації програмної частини проекту буде встановити його на 100 ліжко-місць у палатах по всій Україні.

Найвагомішими перевагами серед конкурентів є відсутність прив'язки до чіткої технологічної бази, додаток може бути встановлено на вже існуючі в лікарнях комп'ютери. А також потенційне створення мобільного додатку на базі розроблених алгоритмів.

Приблизна вартість встановлення додатку та навчання персоналу в розрахунку на 1 лікарняне ліжко-місце - \$100. З планів на перший рік прибуток буде становити \$10000, беззбитковість настане приблизно через 2 роки, коли кількість встановлених програмно-апаратних комплексів буде покривати затрати на оренду приміщення офісу, зарплати працівникам та первинну розробку додатку.

5.2 Організація проекту

Проект реалізується на базі Київського Політехнічного Інституту імені Ігоря Сікорського. Автором проекту є студент бго курсу факультету біомедичної інженерії.

Дата створення проекту – листопад 2021 року.

Адреса – м. Київ, вул. Металістів 4.

5.3 Канва бізнес-моделі проекту

Ключові партнери:

Ключовими партнерами є компанії постачальники електроніки для реалізації програмно-апаратних комплексів із використанням розробляемого продукту

Ключові види діяльності:

- Розробка програмного додатку
- Встановлення контактів із майбутніми споживачами

Ключові ресурси:

Людські ресурси є основними для реалізації основної програмної частини проекту

Ціннісні пропозиції:

Продукт вирішує проблему комунікації із паралізованими хворими у лікарнях для медичного персоналу та родичів пацієнтів.

Продукт не має прямих аналогів в Україні, що робить його надзвичайно цінним для державного сектору охорони здоров'я

Взаємовідносини зі споживачами:

Компанія збирається консультувати споживачів (лікарні) в установці програмно-апаратних комплексів із використанням продукту, а також проводити навчання персоналу по взаємодії із додатком.

Споживчі сегменти:

Державні та приватні лікарні, родичи паралізованих хворих

Канали збуту:

Каналами збуту є участі в державних тендерах, а також прямі продажі приватним лікарням.

Структура витрат:

Найбільші витрати будуть пов'язані із розробкою програмного додатку, вони підуть на заробітні плати програмістів та дизайнерів. Другим пунктом йде реклама та маркетинг. Третім аренда приміщення для офісу компанії. Далі інші витрати.

Потоки надходження доходів:

Доходи генеруються в залежності від умов на базі яких буде виконуватися постачання лікарень розроблюємим програмним додатком. Орієнтована ціна на встановлення 1 одиниці додатку із подальшим навчанням персоналу лікарні становить \$100.

5.4 Ключові види діяльності проекту**Вид проекту за характером інновації**

За характером інновації проект є дослідно-технологічною роботою. Для реалізації основного алгоритму була проведена велика кількість науково-

дослідницької роботи, а подальше впровадження проекту потребує створення повноцінного апаратно-програмного комплексу для функціонування.

Спрямованість проекту

- Випуск продукції, конкурентоспроможної на вітчизняному ринку.
- Оновлення вітчизняної технологічної бази.
- Поліпшення ефективності діагностики та лікування хворих.

У першу чергу проект спрямований на поліпшення умов лікування пацієнтів хворих на параліч тіла. Одночасно з цим він націлений на співпрацю з лікарнями на всеукраїнському рівні. Можливість встановлення цих програмно-апаратних комплексів в свою чергу посприє оновленню вітчизняної технологічної бази.

Висновок щодо науково-технічного рівня ідеї

Проект перевищує існуючі в Україні аналогічні розробки за окремими показниками. За багатьма параметрами ідея є набагато кращою за існуючі в Україні аналоги. Зараз більшість рішень, що використовуються у вітчизняних лікарнях є закордонними аналогами, більша частина з яких не надає такої гнучкості і можливості встановлення додатку на будь-який персональний комп'ютер з веб-камерою, а є комплексними рішеннями, транспортування і встановлення яких є дуже економічно не вигідними. Основні бізнес-процеси проекту представлено в табл.5.1.

Таблиця 5.1

Основні бізнес-процеси проекту

<i>Група процесів</i>	<i>Бізнес-процес</i>	<i>Ступінь опрацювання бізнес-процесу</i>	
		<i>є реалізованим</i>	<i>буде реалізованим</i>
Розробка продукції	розробка та конструювання продукції	Програмне рішення на стадії MVP	Покращення алгоритму та інтерфейсу
Вимоги споживачів	дослідження розвитку ринку	Проведено дослідження українського ринку	

Продовж. табл. 5.1

<i>Група процесів</i>	<i>Бізнес-процес</i>	<i>Ступінь опрацювання бізнес-процесу</i>	
		<i>є реалізованим</i>	<i>буде реалізованим</i>
Вимоги споживачів	організація маркетингу і продажів	Підготовка презентацій та сайту проекту	
Виконання замовлень	забезпечення і матеріально-технічний збут		Співпраця із постачальниками апаратної частини проекту
Обслуговування споживача	консультація персоналу лікарень		Проведення навчальних курсів по роботі з програмно-апаратним комплексом у лікарнях
	технічна підтримка додатку		Постійне оновлення та вирішення потенційних проблем із додатком

5.5 Ціннісні пропозиції та споживачі

Ціннісна пропозиція – сукупність переваг, які проект може запропонувати споживачу.

Характер формування споживчої цінності проекту:

Покращення задоволення існуючих потреб:

Потреба у повноцінному спілкуванні із паралізованими хворими.

Здешевлення задоволення існуючих потреб:

Здешевлення встановлення програмно-апаратних комплексів для спілкуванні із паралізованими хворими у порівнянні з аналогами.

У табл.5.2 представлено головні ідеї проекту.

Таблиця 5.2

Зміст ідеї проекту

<i>Зміст ідеї</i>	<i>Напрямки застосування</i>	<i>Сегменти споживачів</i>	<i>Цінність для споживачів</i>
Розробка універсального програмного додатку-помічника для хворих із різними ступенями паралічу тіла, з використанням сучасних алгоритмів розпізнавання образів та відстежування погляду.	1. Комунікація із паралізованими хворими	Державні та приватні лікарні, родичи паралізованих хворих	Отримати можливість спілкування з паралізованими хворими без спеціального дорогого обладнання та інвазивних процедур.

Аналіз ідеї проекту проведено у табл.5.3.

Таблиця 5.3

Аналіз ідеї проекту

№	<i>Техніко-економічні характеристики ідеї</i>	<i>Продукція конкурентів</i>			<i>W (слабка сторона)</i>	<i>N (нейтральна сторона)</i>	<i>S (сильна сторона)</i>
		<i>Мій проект</i>	<i>IntelliGaze</i>	<i>Tobii</i>			
1.	Можливість встановлення на будь-який ПК з веб-камерою	+	+	-		+	
2.	Спеціалізація на використанні у державних лікарнях	+	-	-			+
3.	Використання інфрачервоного світла для відстежування погляду в умовах поганого освітлення	-	-	+	+		

Перед виконанням проекту було проведено технологічний аудит ідеї проекту, його результати можна спостерігати у табл.5.4.

Таблиця 5.4

Технологічний аудит ідеї проекту

<i>№</i>	<i>Ідея проекту</i>	<i>Технології її реалізації</i>	<i>Наявність технологій</i>	<i>Доступність технологій</i>
1.	Відстежування погляду за допомогою веб-камери ПК	Бібліотека алгоритмів машинного навчання Dlib для мови програмування Python	Наявна	Доступна
2.	Відстежування погляду за допомогою веб-камери ПК	Бібліотека алгоритмів комп'ютерного зору OpenCV для мови програмування Python	Наявна	Доступна
<i>Обрана технологія реалізації ідеї проекту: мова програмування Python</i>				

Для визначення сильних та слабких сторін, можливостей та загроз проекту було проведено SWOT-аналіз проекту (табл.5.5.)

Таблиця 5.5

SWOT-аналіз проекту

<p><i>Сильні сторони (S):</i></p> <ul style="list-style-type: none"> – <i>Можливість встановлення додатку на будь-який ПК</i> – <i>Підходить як для пацієнтів у лікарнях, так і тих, що перебувають вдома</i> – <i>Невелика вартість апаратного забезпечення</i> – <i>Для розробки і підтримки додатку вистачить невеликої команди</i> 	<p><i>Слабкі сторони (W):</i></p> <ul style="list-style-type: none"> – <i>Неможливість гарантувати необхідні для роботи додатку умови освітлення у всіх лікарнях</i> – <i>Недостатня технічна база у державних лікарнях</i> – <i>Необхідність в навчанні персоналу лікарень</i>
<p><i>Можливості (O):</i></p> <ul style="list-style-type: none"> – <i>Співпраця з МОЗ та встановлення програмно-апаратних комплексів із додатком у державних лікарнях</i> – <i>Перспектива стати стандартом спілкування з паралізованими пацієнтами</i> – <i>Перспектива виходу на міжнародний ринок</i> – <i>Можливість розробки програмних рішень для смартфонів</i> 	<p><i>Загрози (T):</i></p> <ul style="list-style-type: none"> – <i>Програмно-апаратні рішення від конкурентів</i> – <i>Нестача фінансування від держави та інвесторів</i> – <i>Відмова у співпраці з МОЗ та лікарнями</i> – <i>Складнощі із постачальниками апаратної частини проекту</i>

У табл.5.6. проаналізовано взаємовідносини із потенційними споживачами та можливі канали збуту.

Таблиця 5.6

Взаємовідносини зі споживачами та канали збуту

№	Сегмент споживачів	Особливості поведінки	Вимоги споживачів	Канали збуту	Інші аспекти взаємовідносин
1.	Державні та приватні лікарні, родичи паралізованих хворих	Мають потребу у інструментах комунікації із паралізованими хворими. Взаємодія із державними лікарнями відбувається через державні установи.	- Додаток має задовільняти потребу у спілкуванні з пацієнтами - Компанія повинна мати підтверження свого рівня кваліфікації у питаннях медицини та комп'ютерної розробки	-Державні замовлення -Продаж апаратно-програмно о комплексу приватним лікарням	

5.6 Обґрунтування ресурсів та витрат проекту

Гарною практикою перед початком роботи над проектом, проводити підрахунок потенційних витрат на проект, визначити ціну майбутнього продукту, обсяги виробництва, обсяги інвестицій та визначити вплив економічних факторів, що можуть на це вплинути.

У табл.5.7 розглянуто визначення ціни товару.

Таблиця 5.7

Визначення ціни

№	Рівень цін на товари-замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на продукцію	Розрахункова ціна продукції
1.	-	€250	Державне фінансування	\$50 - 150	\$100

Визначення обсягу виробництва описано в табл.5.8.

Таблиця 5.8

Визначення обсягу виробництва продукції

Показник	Значення по роках				
	2020	2021	2022	2023	2024
Загальна потреба в продукції	100 од.	500 од.	1000 од.		
Можливі річні обсяги випуску в натуральних показниках					
Ціна одиниці продукції (тис. грн.)	\$100	\$100	\$100		
Річні обсяги випуску в вартісних показниках (тис. грн.)					

У табл.5.9. відбувається розрахунок загальних початкових інвестиційних витрат, від початку досліджень на тему майбутнього програмного додатку і до етапів альфа та бета тестування готового програмного продукту.

Таблиця 5.9

Розрахунок загальних початкових інвестиційних витрат

№	Назва етапу	Терміни виконання	Обсяги фінансування, тис. грн.
1.	Проведення досліджень	01.10.2020 - 01.11.2020	\$500
2.	Опис бізнес процесів	01.11.2020 - 14.11.2020	\$1000
3.	Проектування архітектури системи	15.11.2020 - 15.12.2020	\$2000
4.	Реалізація UI	15.12.2020 - 01.01.2021	\$1500
5.	Реалізація алгоритмів машинного навчання	01.01.2021 - 15.02.2021	\$3000
6.	Alpha тестування	15.02.2021 - 01.04.2021	\$2000
7.	Beta тестування	01.04.2021 - 01.06.2021	\$2500
	<i>Разом</i>		\$12500

У табл.5.10. розраховуються виробничі витрати. У табл.5.11. розраховані загальні витрати. У табл.5.12. проведено оцінку вартості проекту.

Таблиця 5.10

Розрахунок виробничих витрат

№ з/п	Стаття витрат	Сукупні витрати за період, тис. грн.				
		2021	2022	2023	2024	2025
1.	Загальногосподарські витрати	\$5000	\$8000	\$9000		
1. 1.	витрати на оренду та утримання приміщень, обладнання	\$3000	\$4000	\$4000		
1. 2.	комунальні витрати	\$500	\$500	\$500		
1. 3.	витрати на збут, просування та рекламу	\$1500	\$3500	\$4500		
2.	Витрати на матеріальні ресурси (комплектуючі, сировина)	\$2000	\$6000	\$4000		
3.	Витрати на оплату праці	\$10000	\$15000	\$15000		
	<i>Разом:</i>	\$17000	\$29000	\$29000		

Таблиця 5.11

Розрахунок загальних витрат на реалізацію проекту по роках

Показник	Значення по роках					Разом
	2021	2022	2023	2024	2025	
Інвестиційні витрати	\$12500					
Виробничі витрати	\$17000	\$29000	\$29000			
Обсяг загальних витрат, в тому числі за рахунок	\$29500	\$29000	\$29000			
– власних коштів	\$10000	\$19000	\$20000			
– коштів інвестора	\$19500	\$10000	\$9000			

Таблиця 5.12

Грошовий потік так оцінка вартості проекту

№	Показник	Значення по роках			Разом
		2021	2022	2023	
1.	Надходження від проекту (виручка від реалізації продукції, послуг) (D)	\$10000	\$50000	\$100000	\$160000
2.	Загальні витрати	\$29500	\$29000	\$29000	\$87500
3.	Грошовий потік (3 = 1 – 2) (CF)	\$-19500	\$21000	\$71000	\$72500

Висновки до розділу 5

Проект є дуже привабливим для інвестицій з точки зору сучасних медичних компаній та виробників біотехнологічної електроніки, а також для держави. Оскільки основною темою проекту є допомога хворим людям, то проект може вважатися соціально значимим, і може розраховувати не тільки на інвестиції від інвесторів що планують у майбутньому отримати прибуток та акції компанії, а й на небайдужих людей та компанії що сприяють розвитку медицини в Україні.

ЗАГАЛЬНІ ВИСНОВКИ

В результаті виконання даної магістерської дисертації було досягнуто головних її цілей, а саме:

- Проаналізовано технології реалізації та обрано і обгрунтовано оптимальні алгоритми відстежування погляду та розпізнавання образів;
- Розроблено систему предиктивного набору тексту для полегшення спілкування з паралізованими пацієнтами;
- Реалізовано програмний додаток, що надає пацієнту можливість набирати текстові повідомлення на віртуальній клавіатурі за допомогою руху його очей.

Даний програмний додаток розрахований на пацієнтів хворих на параліч тіла, та має на меті допомогти їм та створити новий канал спілкування. Потенціально він може бути впроваджений у будь-який медичний заклад укомплектований персональними комп'ютерами необхідними для його роботи.

Додаток реалізовано на основі мови програмування Python із використанням сучасних бібліотек розпізнавання образів та обробки природньої мови. Він містить у собі систему розпізнавання обличчя та відстежування погляду пацієнтів із використанням методу гістограм орієнтованих градієнтів. А також систему предиктивного набору тексту для полегшення користування програмою засновану на методі n-gram.

В ході роботи було розглянуто та проаналізовано вимоги та обмеження для додатку. При проектуванні було створено контекстну діаграму, та здійснена декомпозиція програмного додатку.

Також було оформлено стартап-проект за темою магістерської дисертації, що винесено в окремий розділ. В ньому коротко наведено основну інформацію щодо стартап-проекту, проаналізовано бізнес процеси, проведено SWOT-аналіз проекту та розраховано інвестиційні витрати.

У подальшій перспективі можливе доопрацювання та удосконалення програмного продукту, розширення можливостей системи предиктивного набору тексту, покращення користувацького досвіду, оптимізація додатку під мобільні пристрої.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Базова література

1. Коваленко М.В. Програма помічник для паралізованих хворих, заснована на технології відстежування погляду. 2020. С. 1–76.
2. Соломін А.В. Комп'ютерна система стеження за очима пацієнта / А.Соломін, Г.Гетун, Г.Корнієнко, М.Коваленко // *Modern Scientific Researches*. – Yornat PE, Minsk, Belarus, 2020. – Issue №11, Part 2. – P. 70–73. DOI: 10.30889/2523-4692.2020-11-02-006.
3. У 2020 році лікування гострого мозкового інсульту є пріоритетом в програмі медичних гарантій [Електронний ресурс] // Міністерство охорони здоров'я України. – 2019. – Режим доступу до ресурсу: <https://www.kmu.gov.ua/news/u-2020-roci-likuvannya-gostrogo-mozkovogo-insultu-ye-prioritetom-v-programi-medichnih-garantij>.
4. Цікаві факти про українську мову [Електронний ресурс] // Мова – ДНК нації. – 2015. – Режим доступу до ресурсу: <https://ukr-mova.in.ua/blog/czikavi-faktu-pro-ukrayinsku-movu>.
5. A. Bhatia, A. Bharadia, K. Sawant and S. Kurhade, "Predictive and corrective text input for desktop editor using n-grams and suffix trees," 2016 International Conference on Advances in Human Machine Interaction (HMI), 2016, pp. 1-4, doi: 10.1109/HMI.2016.7449191.
6. A. Elmogy, B. Mahmoud and M. Saleh, "A Deep Learning Approach for Text Generation," 2019 29th International Conference on Computer Theory and Applications (ICCTA), 2019, pp. 102-106, doi: 10.1109/ICCTA48790.2019.9478833.
7. A. J. KHDR and C. VAROL, "Age and Gender Identification by SMS Text Messages," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018, pp. 1-5, doi: 10.1109/IDAP.2018.8620780.
8. A. Paoloni, S. Ragazzini and G. Ravaioli, "Predictive neural networks in text independent speaker verification: an evaluation on the SIVA database," *Proceeding*

- of Fourth International Conference on Spoken Language Processing. ICSLP '96, 1996, pp. 2423-2426 vol.4, doi: 10.1109/ICSLP.1996.607298.
9. Airport Facial Recognition Passenger Flow Management [Электронный ресурс] // 2015 – Режим доступа до ресурсу:
<http://www.hrsid.com/company/technology/face-recognition>.
 10. A. McCaul and A. Sutherland, "Predictive text entry in immersive environments," IEEE Virtual Reality 2004, 2004, pp. 241-242, doi: 10.1109/VR.2004.1310089.
 11. Biemer, Colan & Hervella, Alejandro & Cooper, Seth. (2021). Gram-Elites: N-Gram Based Quality-Diversity Search. 1-6. 10.1145/3472538.3472599.
 12. Bonsor, K. (September 4, 2001). "How Facial Recognition Systems Work". Retrieved June 2, 2008.
 13. B. Yang and J. Hong, "Performing literature review using text mining, Part II: Expanding domain knowledge with abbreviation identification," 2017 IEEE International Conference on Big Data (Big Data), 2017, pp. 3297-3301, doi: 10.1109/BigData.2017.8258314.
 14. G. Huang, "Attention Guided Multi-Scale Regression for Scene Text Detection," 2021 2nd International Conference on Computing and Data Science (CDS), 2021, pp. 498-502, doi: 10.1109/CDS52072.2021.00092.
 15. G. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91-110, 2004.
 16. G. Sun, J. Liu, W. Mengxue, W. Zhongxin, Z. Jia and G. Xiaowen, "An Ensemble Classification Algorithm for Imbalanced Text Data Streams," 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), 2020, pp. 1073-1076, doi: 10.1109/ICAICA50127.2020.9182576.
 17. Gates, Kelly (2011). Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance. NYU Press. p. 54. ISBN 9780814732090
 18. Gates, Kelly (2011). Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance. NYU Press. pp. 48–49. ISBN 9780814732090.
 19. Gates, Kelly (2011). Our Biometric Future__ : Facial Recognition Technology and the Culture of Surveillance. NYU Press. pp. 49–50. ISBN 9780814732090.

20. Gong, Jun; Peter Tarasewich (2005). Alphabetically constrained keypad designs for text entry on mobile devices (PDF). ACM Press. p. 211.
doi:10.1145/1054972.1055002. ISBN 1-58113-998-5. Retrieved 2011-11-03.
21. H. Komatsu, S. Takabayashi and T. Masui, "Corpus-based predictive text input," Proceedings of the 2005 International Conference on Active Media Technology, 2005. (AMT 2005)., 2005, pp. 75-80, doi: 10.1109/AMT.2005.1505271.
22. H. Li, F. Zhu and J. Qiu, "Towards Document Image Quality Assessment: A Text Line Based Framework and a Synthetic Text Line Image Dataset," 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019, pp. 551-558, doi: 10.1109/ICDAR.2019.00094.
23. Harris, Carl. (2001). Markov chain(s) fundamentals of Markov chain(s) coin toss sequence and Markov chains. 10.1007/1-4020-0611-X_579.
24. Heaven Stubblefield. What are the symptoms of paralysis? [Электронный ресурс] / Heaven Stubblefield // healthline. – 2018. – Режим доступа до ресурсу: <https://www.healthline.com/health/paralysis#symptoms>.
25. Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. – 1997, v.9 (8). – p. 1735–1780. DOI:10.1162/neco.1997.9.8.1735.
26. Hubert Eichner, Chlo Kiddon, and Daniel Ramage. 2018. Federated learning for mobile keyboard prediction.
27. J. G. Cleary and W. J. Teahan, "An open interface for probabilistic models of text," Proceedings DCC'99 Data Compression Conference (Cat. No. PR00096), 1999, pp. 522-, doi: 10.1109/DCC.1999.785679.
28. J. Joestar, D. Brando, "A Chinese predictive text entry method for mobile devices," 2014 International Conference on Machine Learning and Cybernetics, 2014, pp. 309-314, doi: 10.1109/ICMLC.2014.7009134.
29. J. Lovinger and I. Valova, "Scrubbing the Web for Association Rules: An Application in Predictive Text," 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), 2015, pp. 439-442, doi: 10.1109/ICMLA.2015.54.

30. J. Wei, K. Chen, J. He, Z. Huang, Y. Lian and Y. Zhou, "A New Approach for Integrated Recognition and Correction of Texts from Images," 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019, pp. 615-620, doi: 10.1109/ICDAR.2019.00104.
31. Kenneth Maiese. Locked-In Syndrome [Электронный ресурс] / Kenneth Maiese // Merck Manual. – 2019. – Режим доступа до ресурсу: <https://www.merckmanuals.com/home/brain,-spinal-cord,-and-nerve-disorders/coma-and-impaired-consciousness/locked-in-syndrome#>.
32. L. Zhang et al., "Assessment of Career Adaptability: Combining Text Mining and Item Response Theory Method," in IEEE Access, vol. 7, pp. 125893-125908, 2019, doi: 10.1109/ACCESS.2019.2938777.
33. Laureys, Steven & Pellas, Frederic & Eeckhout, Philippe & Ghorbel, Sofiane & Schnakers, Caroline & Perrin, Fabien & Berre, Jacques & Faymonville, Marie-Elisabeth & Pantke, Karl-Heinz & Damas, Francois & Lamy, Maurice & Moonen, Gustave & Goldman, Serge. (2005). The locked-in syndrome: What is it like to be conscious but paralyzed and voiceless?. Progress in brain research. 150. 495-511. 10.1016/S0079-6123(05)50034-7.
34. Li, Stan Z.; Jain, Anil K. (2005). Handbook of Face Recognition. Springer Science & Business Media. p. 2. ISBN 9780387405957.
35. Liji Thomas. What Causes Paralysis? [Электронный ресурс] / Liji Thomas // News-Medical.Net. – 2018. – Режим доступа до ресурсу: <https://www.news-medical.net/health/What-Causes-Paralysis.aspx>
36. Lin, Xiaofan. (2006). Predictive Text Fitting. 4073. 13-23. 10.1007/11795018_2.
37. Lu Gao and P. A. Beling, "Machine quantification of text-based economic reports for use in predictive modeling," SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483), 2003, pp. 3536-3541 vol.4, doi: 10.1109/ICSMC.2003.1244437.
38. M. Habib, M. Faris, R. Qaddoura, A. Alomari and H. Faris, "A Predictive Text System for Medical Recommendations in Telemedicine: A Deep Learning

- Approach in the Arabic Context," in IEEE Access, vol. 9, pp. 85690-85708, 2021, doi: 10.1109/ACCESS.2021.3087593.
39. M. Pandey, K. Chaudhari, R. Kumar, A. Shinde, D. Totla and N. D. Mali, "Assistance for Paralyzed Patient Using Eye Motion Detection," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697455.
 40. Megalingam, Rajesh Kannan & Karath, Malavika & Valluri, Vaishnavi & Shreedhar, Aparna & Katti, Meghana & Gurjar, Sapna & Nair, Aiswarya. (2015). Language for people suffering from complete paralysis based on Eye Gestures. 10.1109/NGCT.2015.7375167.
 41. Mondragon, C. K.; Bleacher, B. (2013). "Eye tracking control of vehicle entertainment systems". Retrieved 3 August 2020.
 42. Nilsson, Nils J. (2009). The Quest for Artificial Intelligence. Cambridge University Press. ISBN 9781139642828.
 43. Paralysis [Электронный ресурс] // Encyclopaedia Britannica – Режим доступа до ресурсу: <https://www.britannica.com/science/paralysis>.
 44. Paralysis: Causes, Symptoms, Diagnosis & Treatment [Электронный ресурс] // Cleveland Clinic. – 2017. – Режим доступа до ресурсу: <https://my.clevelandclinic.org/health/diseases/15345-paralysis>.
 45. R. Brunelli, Template Matching Techniques in Computer Vision: Theory and Practice, Wiley, ISBN 978-0-470-51706-2, 2009 (1(<http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470517069.html>) TM book)
 46. Research on communication with completely paralyzed patients prompts misconduct investigation [Электронный ресурс] // American Association for the Advancement of Science. – 2019. – Режим доступа до ресурсу: <https://www.science.org/content/article/research-communication-completely-paralyzed-patients-prompts-misconduct-investigation>.
 47. S. Akkaradamrongrat, P. Kachamas and S. Sinthupinyo, "Text Generation for Imbalanced Text Classification," 2019 16th International Joint Conference on

- Computer Science and Software Engineering (JCSSE), 2019, pp. 181-186, doi: 10.1109/JCSSE.2019.8864181.
48. S. Jiang, C. Zhai and Q. Mei, "Exploiting Knowledge Graph to Improve Text-based Prediction," 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 1407-1416, doi: 10.1109/BigData.2018.8622123.
 49. Scott MacKenzie (2002). "KSPC (Keystrokes per Character) as a Characteristic of Text Entry Techniques"
 50. Sidorov G., Velasquez F., Stamatatos E., Gelbukh, A., Chanona-Hernández L. Syntactic Dependency-Based N-grams as Classification Features // Advances in Computational Intelligence. Lecture Notes in Computer Science. 7630. – 2013. – p. 1–11; DOI:10.1007/978-3-642-37798-3_1. ISBN 978-3-642-37797-6.
 51. Simon Amez, Stijn Baert, Smartphone use and academic performance: A literature review, International Journal of Educational Research, Volume 103, 2020, 101618, ISSN 0883-0355,
 52. Smith, Kelly. "Face Recognition" (PDF). Retrieved June 4, 2008.
 53. Stats about paralysis [Электронный ресурс] // Christopher & Dana Reeve Foundation. – 2017. – Режим доступа до ресурсу: <https://www.christopherreeve.org/living-with-paralysis/stats-about-paralysis>.
 54. Т9, или Предикативный ввод текста в мобильных телефонах [Электронный ресурс] // iXbt – Режим доступа до ресурсу: <https://www.ixbt.com/mobile/review/prtxtsms.shtml>.
 55. "Top Five Biometrics: Face, Fingerprint, Iris, Palm and Voice". Bayometric. January 23, 2017. Retrieved April 10, 2018.
 56. W. T. Freeman, M. Roth. Orientation histograms for hand gesture recognition. Intl. Workshop on Automatic Face and Gesture-Recognition, IEEE Computer Society, Zurich, Switzerland, pages 296–301, June 1995.
 57. Williams, Mark. "Better Face-Recognition Software". Retrieved June 2, 2008.
 58. Y. Gu and J. Shen, "Short Text Classification Based on Keywords Extension," 2019 Chinese Automation Congress (CAC), 2019, pp. 2616-2621, doi: 10.1109/CAC48633.2019.8996664.

59. Y. Ikegami, Y. Sakurai, E. Damiani, R. Knauf and S. Tsuruta, "Flick: Japanese Input Method Editor Using N-Gram and Recurrent Neural Network Language Model Based Predictive Text Input," 2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2017, pp. 50-55, doi: 10.1109/SITIS.2017.19.
60. Y. Lin and J. Wang, "Research on text classification based on SVM-KNN," 2014 IEEE 5th International Conference on Software Engineering and Service Science, 2014, pp. 842-844, doi: 10.1109/ICSESS.2014.6933697.
61. de Leeuw, Karl; Bergstra, Jan (2007). The History of Information Security: A Comprehensive Handbook. Elsevier. p. 266. ISBN 9780444516084.

Інформаційні ресурси

62. Рисунок 1.1. Приклад використання технології розпізнавання обличчя на групі студентів / Project 4: Face detection with a sliding window [Електронний ресурс] // CS 143: Introduction to Computer Vision. – 2011. – Режим доступу до ресурсу: <http://cs.brown.edu/courses/csci1430/2011/results/proj4/hangsu/>.
63. Рисунок 1.2. Узагальнені (стиснуті) зображення обличчя створені на основі великих масивів зображень / https://www.researchgate.net/figure/First-10-eigenfaces-with-highest-eigenvalues_fig4_264167711 [Face Recognition Based on Improved Fuzzy RBF Neural Network for Smart Device]
64. Рисунок 1.3. Приклад пристрою, що використовує оптичний метод відстежування погляду / Eye-tracking and cartography [Електронний ресурс] // National Technical University of Athens School of Rural & Surveying Engineering Cartography Laboratory – Режим доступу до ресурсу: http://users.ntua.gr/bnakos/Eye_Tracking_Eng.html.
65. Рисунок 1.4. Відстежування положення обличчя водія автомобіля / Eyes on the road – your car is watching you [Електронний ресурс] // Automotive World Ltd.. – 2016. – Режим доступу до ресурсу: <https://1bgnvt3q09toyb96v2ecsygm-wpengine.netdna-ssl.com/wp-content/uploads/2014/09/eye-tracking.jpg>.

66. Рисунок 1.5. Стандартна клавіатура ІТУ-Т Е.161, що використовується для обміну текстовими повідомленнями / An E.161 standard modern telephone keypad, as used for text messaging [Електронний ресурс]. – 2008. – Режим доступу до ресурсу: <https://en.wikipedia.org/wiki/Multi-tap#/media/File:Telephone-keypad.svg>.
67. Рисунок 2.1. Паралізований пацієнт, користується Інтернетом за допомогою перемикача у формі соски / Yan Ching-hong, a paralyzed patient, surfs the internet by using a pacifier-shaped switch [Електронний ресурс]. – 2012. – Режим доступу до ресурсу: <https://www.alamy.com/yan-ching-hong-a-paralyzed-patient-surfs-the-internet-by-using-a-pacifier-shaped-switch-invented-by-luo-ching-hsing-a-professor-at-taiwans-national-cheng-kung-university-in-tainan-southern-taiwan-december-19-2012-the-pacifier-shaped-switch-which-can-connect-to-computers-via-a-usb-port-allows-paralyzed-patients-to-type-on-computers-and-even-play-video-games-that-help-with-their-recovery-reuterspichi-chuang-taiwan-tags-science-technology-health-society-image378749832.html>.
68. Рисунок 2.2. Приклад використання системи прогнозування тексту компанією Google / Bharath K. Next Word Prediction with NLP and Deep Learning [Електронний ресурс] / Bharath K // Towards Data Science. – 2020. – Режим доступу до ресурсу: <https://towardsdatascience.com/next-word-prediction-with-nlp-and-deep-learning-48b9fe0a17bf>.
69. Рисунок 2.3. Візуальна схема розбиття речення на N-грами / Solomin A.V. Predictive text typing system for communication with paralyzed patients / Solomin A.V, Kovalenko M.V. // Modern engineering and innovative technologies. – Germany, 2021.– Issue №17
70. Рисунок 4.1. Загальний принцип дії методу гістограм орієнтованих градієнтів / Marek Kraft. Taming the HoG: The Influence of Classifier Choice on Histogram of Oriented Gradients Person Detector Performance [Електронний ресурс] / Marek Kraft. – 2017. – Режим доступу до ресурсу: https://www.researchgate.net/profile/Marek_Kraft/publication/318134187/figure

re/fig1/AS:578366711111685@1514904777612/Formation-of-the-histogramof-oriented-gradients-descriptor.png