

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ Центр післядипломної освіти  
(повна назва)  
Кафедра \_\_\_\_\_ Програмної інженерії  
(повна назва)

**КВАЛІФІКАЦІЙНА РОБОТА**  
**Пояснювальна записка**

рівень вищої освіти \_\_\_\_\_ другий (магістерський)

**Дослідження та аналіз методів для розробки додатку**

**визначення положення та орієнтації об'єктів у**

**просторі**

(тема)

Виконав:

Випускник 2 курсу, групи ІПЗзДМ-19-1

Рябоволова А.О.

(прізвище, ініціали)

Спеціальність 121 Інженерія програмного  
забезпечення

(код і повна назва спеціальності)

Тип програми Освітньо-наукова

(освітньо-професійна або освітньо-наукова)

Керівник проф. Білоус Н.В.

(посада, прізвище)

Допускається до захисту

Зав. кафедри

\_\_\_\_\_ (підпис)

З.В. Дудар

(прізвище, ініціали)

2021р.

## Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ Центр післядипломної освіти \_\_\_\_\_

(повна назва)

Кафедра \_\_\_\_\_ Програмної інженерії \_\_\_\_\_

(повна назва)

Рівень вищої освіти \_\_\_\_\_ другий (магістерський) \_\_\_\_\_

Спеціальність \_\_\_\_\_ 121 – Інженерія програмного забезпечення \_\_\_\_\_

(код і повна назва)

Тип програми \_\_\_\_\_ освітньо-наукова програма \_\_\_\_\_

(освітньо-професійна або освітньо-наукова)

Освітня програма \_\_\_\_\_ Інженерія програмного забезпечення \_\_\_\_\_

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_

(підпис)

« 26 » березня 2021 р.

**ЗАВДАННЯ****НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студента \_\_\_\_\_ Рябоволової Анастасії Олегівни \_\_\_\_\_

(прізвище, ім'я, по батькові)

1. Тема роботи «Дослідження та аналіз методів для розробки додатку визначення положення та орієнтації об'єктів у просторі» затверджена наказом університету від 26.03.2021 № 34Стз.
2. Термін подання студентом роботи до екзаменаційної комісії «20» травня 2021 р.
3. Вихідні дані до роботи комп'ютерний зір, визначення положення та орієнтації об'єктів
4. Перелік питань, що потрібно опрацювати в роботі мета роботи, аналіз предметної галузі, постановка задачі, можливості сучасних алгоритмів, задачі визначення положення та орієнтації об'єктів у просторі

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, ілюстрацій (слайдів) мета завдання, обґрунтування доцільності розроблення, постановка задачі, методи і алгоритми, структурно-логічна схема взаємодії даних, опис отриманих результатів, інтерфейс програмної системи, демонстраційні матеріали

6 Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата
Спецчастина	проф. Білоус Н.В.		

### КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної галузі	25.02.2021	<i>Виконано</i>
2	Огляд існуючих методів та алгоритмів	05.04.2021	<i>Виконано</i>
3	Дослідження існуючих алгоритмів	06.04.2021	<i>Виконано</i>
4	Підготовка пояснювальної записки	07.04.2021	<i>Виконано</i>
5	Спецчастина	20.04.2021	<i>Виконано</i>
6	Підготовка презентації та доповіді	08.05.2021	<i>Виконано</i>
7	Попередній захист	11.05.2021	<i>Виконано</i>
8	Перевірка на плагіат	12.05.2021	<i>Виконано</i>
9	Нормоконтроль, рецензування	20.05.2021	<i>Виконано</i>
10	Занесення диплома в електронний архів	21.05.2021	<i>Виконано</i>
11	Допуск до захисту у зав. кафедри	21.05.2021	<i>Виконано</i>

Дата видачі завдання \_\_\_\_\_ 2021р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_ проф. Білоус Н. В.  
(підпис) (посада, прізвище, ініціали)

**РЕФЕРАТ / ABSTRACT**

Кваліфікаційна робота магістра містить: 45 с., 15 рис., 0 табл., 19 джер.

ПОЛОЖЕННЯ ОБ'ЄКТА, ОРІЄНТАЦІЇ ОБ'ЄКТА, КОМП'ЮТЕРНИЙ ЗІР, ОЦІНКА ПОЗИ, PnP, PoseCNN, Pix2pose, POSIT, PYTHON, OpenCV, OpenGL.

Об'єктом дослідження магістерської роботи є методи розробки додатку визначення положення та орієнтації об'єктів у просторі.

Метою роботи є визначення та аналіз методів для розробки додатку визначення положення та орієнтації об'єктів у просторі та розробка програмного продукту з використанням результатів досліджень.

Для даної магістерської роботи необхідна підключена камера та комп'ютер з відео картою з підтримкою CUDA розрахунків.

Результатами даної роботи є аналіз методів визначення положення та орієнтації об'єктів. На базі яких реалізовано додаток визначення положення та орієнтації об'єкта, основною особливістю якого є можливість обрати алгоритм та аналізувати використання ресурсів та результатів роботи. Елементами наукової новизни роботи є додаток з реалізацією сучасних рішень та подальшою можливістю використання його для навчання і порівняння у режимі реального часу.

Дана робота проводилась під кураторством учбового наукового центра УНПЦА.

Методи проектування та розробки базуються на технологіях OpenCV та OpenGL з використанням мови програмування Python.

OBJECT POSITION, OBJECT ORIENTATIONS, COMPUTER VISION, POSITION EVALUATION, PnP, PoseCNN, Pix2pose, POSIT, PYTHON, OpenCV, OpenGL.

The object of research of the master's work is the methods of developing an application for determining the position and orientation of objects in space.

The aim of the work is to define and analyse methods for the development of an application for determining the position and orientation of objects in space and the development of a software product using research results.

This master's research work requires a connected camera and a computer with a video card that supports CUDA calculations.

The results of this work are an analysis of methods for determining the position and orientation of objects. Based on which the application of determining the position and orientation of the object is implemented, the main feature of which is the ability to choose the algorithm and analyse the use of resources and results. Elements of scientific novelty of the work are the application with the implementation of modern solutions and the subsequent possibility of using it for learning and comparison in real time.

This work was supervised by the Outsourcing Scientific Training and Production Centre (OSTPC).

Design and development methods are based on OpenCV and OpenGL technologies using the Python programming language.

Я, Рябоволова Анастасія Олегівна, студентка гр. ПЗЗдм-19-1, здобувач вищої освіти на другому (магістерському) рівні кафедри «Програмна інженерія», заявляю: моя кваліфікаційна робота на тему «Дослідження та аналіз методів для розробки додатку визначення положення та орієнтації об'єктів у просторі», що буде представлена в екзаменаційну комісію для публічного захисту, виконана самостійно, в ній не містяться елементи плагіату і вона може бути опублікована в електронному архіві відкритого доступу ElAr KhNURE. Всі запозичення з друкованих та електронних джерел мають відповідні посилання.

Я ознайомлений (а) з діючим положенням «Про протидію академічному плагіату в ХНУРЕ», згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування дисциплінарних заходів.

## ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ	8
ВСТУП	9
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	11
1.1 Загальне поняття комп'ютерного зору	11
1.2 Загальне поняття визначення положення та орієнтації об'єкта	12
1.3 Огляд сучасних задач визначення положення та орієнтації	14
1.4 Аналіз методів визначення положення та орієнтації	15
1.5 Актуальність задачі	17
2 АНАЛІЗ АЛГОРИТМІВ ТА ПОСТАНОВКА ЗАДАЧІ	19
2.1 Алгоритм визначення положення на основі вірогідного підходу	19
2.2 PoseCNN алгоритм	22
2.3 Pix2Pose алгоритм	25
2.4 POSIT алгоритм	27
2.5 Порівняння алгоритмів	29
2.6 Постановка задачі та функціональні вимоги до системи	30
3 ПРОЕКТУВАННЯ ДОДАТКУ	32
3.1 UML моделювання	32
3.1.1 Діаграма прецедентів	32
3.1.2 Діаграма станів	33
3.1.3 Діаграма компонентів	34
3.2 Моделювання схеми бази даних	35
4 ОПИС ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ	36
4.1 Вибір інструментів реалізації	36
4.2 Опис інтерфейсу та функціональності	37
4.3 Апробація результатів роботи та тестування додатку	41
ВИСНОВКИ	43
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	44

	7
ДОДАТОК А	46
ДОДАТОК Б	47
ДОДАТОК В	48
ДОДАТОК Г	49
ДОДАТОК Д	50

## ПЕРЕЛІК СКОРОЧЕНЬ

CNN – convolutional neural network, згорткова нейронна мережа.

CUDA – compute unified device architecture, програмно-апаратна архітектура паралельних обчислень, яка дозволяє збільшити обчислювальну здатність завдяки використанню графічних процесорів фірми Nvidia.

EM – expectation–maximization algorithm, алгоритм очікування-максимізації.

PoseCNN – назва алгоритму оцінки пози з використанням згорткової нейронної мережі.

POSIT – назва алгоритму оцінки пози на базі визначення та зіставлення опорних точок.

Pix2Pose – назва алгоритму оцінки пози з використанням згорткової нейронної мережі.

PnP – Perspective-n-Point, це проблема оцінки відносного положення та орієнтації - між відкаліброваною камерою перспективи та 3D-об'єктом (або між камерою та всією 3D-сценою) із набору  $n$  видимі 3D-точки з відомими  $(X, Y, Z)$  координатами об'єкта (або сцени) та їх 2D-проекції з відомими  $(u, v)$  піксельними координатами.

RGB – red, green, blue або червоний, зелений, синій, адитивна кольорова модель, описуюча здатність кодування кольорів за допомогою трьох основних кольорів.

RGB-D – RGB з четвертим параметром глибини зображення.

## ВСТУП

Вміння розпізнавати положення та орієнтацію об'єктів є дуже важливим фактором для розвитку автоматизованих систем. Автоматизованих систем, які зможуть з великою точністю маніпулювати з об'єктами або аналізувати умови навколишнього середовища.

Також, це є дуже впливовим поштовхом для розвитку медичних рішень допомоги лікарям або покращення результатів обстежень та удосконалення таких напрямів як: використання роботів у сфері послуг та обслуговування, удосконалення безпілотних пристроїв, тощо.

У рамках даної магістерської роботи були проаналізовані попередні роботи на базі кафедри ПІ та розширено сферу розпізнавання пози та орієнтації об'єктів у просторі. Були проаналізовані попередні роботи які пов'язані з обробкою та аналізом зображень.

Метою магістерської є проведення аналізу методів для розробки додатку визначення положення і орієнтації об'єктів у просторі та побудову додатку з використанням сучасних рішень для вирішення задачі оцінки пози об'єктів. Дослідження сучасних реалізацій методів на основі існуючих алгоритмів та інструментів.

Об'єкт дослідження – визначення положення та орієнтації об'єктів у просторі, алгоритми реалізації методів та інструменти побудови додатку для оцінки пози.

Предмет дослідження – методи для розробки додатку визначення положення та орієнтації об'єктів у просторі, сучасні алгоритми реалізації цих методів та інструменти.

У роботі досліджено методи для розробки додатку визначення положення та орієнтації об'єктів у просторі. Розглянуті методи на базі опорних точок та згорткових нейронних мереж. Також, проаналізовані існуючі підходи та рішення у

якості побудованих алгоритмів та бібліотек Для розробки програмного продукту було обрано мову програмування Python.

Елементами наукової новизни даної роботи є втілення існуючих підходів у сучасному додатку з використанням зручного інтерфейсу та подальшою можливістю застосування цього рішення для навчання студентів і експериментів для аналізу алгоритмів у режимі реального часу.

Дана магістерська робота пройшла апробацію на науковій конференції. Результати роботи були представлені на міжнародній науково-практичній конференції «Priority directions of science and technology development» яка відбувалася з 18.04.2021 по 20.04.2021 в Києві, Україна. Результати роботи опубліковані у збірнику тез конференції.

Практична цінність даної роботи полягає у тому, що були проаналізовані сучасні інструменти та побудовано додаток у якому можна використовувати ці методи та алгоритми для тестування різних підходів та на різних об'єктах у режимі реального часу.

## 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

### 1.1 Загальне поняття комп'ютерного зору

Комп'ютерний зір або комп'ютерне бачення – це міждисциплінарна наукова область яка вивчає як комп'ютер може збільшити кількість знань добутих з зображень або відео ряду. У більш ширшому баченні область комп'ютерного зору ставить собі за завдання навчити машину бачити як людина.

Задачі комп'ютерного зору включають у собі методи отримання зображення, його обробку, аналіз отриманих даних та вилучення цифрової або символічної інформації про навколишній світ. Відеодані можуть бути представлені багатьма формами: зображення або відеоряд отриманні з камери, тривимірні дані графічного процесору або згенеровані дані зображення.

Прикладами сучасних систем комп'ютерного зору можна назвати[1]:

- системи керування процесами такі як промислові роботи, автономні транспортні засоби, тощо;
- системи відеоспостереження;
- системи організації інформації, наприклад для індексації баз даних зображень;
- системи моделювання об'єктів або навколишнього середовища (аналіз медичних зображень, тощо)
- системи взаємодії такі як пристрої введення для систем людино-машинної взаємодії.

Комп'ютерний зір також може бути описаний як доповнення (але не обов'язково протилежність) біологічному зору. У біології вивчається зорове сприйняття людини і різноманітних тварин, в підсумку створюються моделі роботи таких систем в термінах фізіологічних процесів[2]. Комп'ютерний зір, з іншого боку, вивчає і описує системи комп'ютерного зору, які виконано апаратно або програмно.

## 1.2 Загальне поняття визначення положення та орієнтації об'єкта

Оцінка пози – техніка комп'ютерного зору яка дає можливість передбачати та відстежувати позицію об'єкта чи людини. Це досягається завдяки знаходженню різних комбінацій пози та орієнтацій даних об'єктів. Також задача оцінки пози може бути визначена як знаходження положення та орієнтацію камери відносно людини чи об'єкта.

В даний час всі підходи до визначення положення і орієнтації тривимірних об'єктів в просторі можна розділити на дві групи: неоптичні і оптичні. До неоптичних технологій відносяться електромагнітні, механічні, інерційні, акустичні та гібридні[3].

В електромагнітних системах використовуються магніти, розміщені на об'єкті. Система розраховує позиції по спотворень магнітного потоку. Недоліками подібних систем є залежність від магнітних та електричних розрядів, мінлива чутливість сенсорів в залежності від положення, обмежений робочий простір.

У механічних системах використовується каркас, який кріпиться до об'єкта. Рух складових частин об'єкта призводить в рух елементи каркаса. В цьому випадку аналізуються кути між структурними елементами каркаса. Однак переміщення всього об'єкта в просторі механічна система фіксувати не здатна.

Інерціальні системи для збору інформації про рух використовують гіроскопи, розташовані на об'єкті. Система визначає не тільки положення сенсора, але також кут його нахилу. Інерційні системи стійкі до перешкод різного роду, але під час руху призводить до накопичення помилки.

Акустичні системи використовують датчики, які формують сигнал. Приймачі, встановлені на деякій відстані, вимірюють час польоту переданого сигналу. Акустичні системи дуже сприйнятливі до перешкод навколишнього середовища[4]. Також можуть бути встановлені обмеження, пов'язані з кількістю одночасно використовуваних пристроїв, розміром робочого простору, тощо.

Гібридні системи для підвищення ефективності можуть використовувати в своєму складі кілька технологій.

Оптичні методи побудовані на використанні маркерних і безмаркерних методах. Маркерні оптичні системи бувають пасивними і активними. Оптичні пасивні маркерні системи використовують датчики, які відбивають падаюче на них світло. В оптичних активних системах світло випускається самими маркерами[4].

Вищеописані технології вимагають використання спеціальних датчиків, засобів сполучення з блоками прийому і обробки даних, обмеженим робочим простором. Основним недоліком розглянутих методів при використанні в спеціалізованих системах є неможливість застосування датчиків при постійному спостереженні безлічі сцен, що істотно обмежує їх застосування. Безмаркерні оптичні методи є найбільш перспективними з точки зору простоти використання і універсальності[4]. Тому важливою є розробка методів і програмних засобів визначення положення і орієнтації об'єктів з використанням безмаркерних оптичних методів.

Питання розпізнавання на основі пасивних оптичних методів в даний час вивчені недостатньо. Використання комп'ютерного зору обмежена наступними факторами спостереження: наявністю шумових і оптичних ефектів, що зумовлено зміною фону, ефектами різкої зміни освітлення, взаємним перекриттям об'єктів в сцені. Крім того, при реалізації систем існують невизначеності, пов'язані з виділенням інформативних ознак об'єкта, так як двовимірне зображення не володіє повнотою інформації про властивості тривимірного об'єкту. Тому для реалізації підходу необхідно використовувати заздалегідь визначену інформацію про структуру сцени та умови отримання зображення.

Безмаркерні оптичні методи визначення положення і орієнтації об'єкта діляться на трекери і детектори. Трекери відстежують стан об'єкта від кадру до кадру. Детектори визначають положення об'єкта в поточному кадрі без урахування стану в попередніх кадрах[4]. Застосування конкретного безмаркерного оптичного методу залежить від умов спостереження сцени. Підходи, які використовують контури, краще застосовувати в тих випадках, коли зображення візуальної

обстановки отримано в градаціях сірого кольору, на знімках можна досить чітко виділити краї об'єкта, присутні ефекти різкої зміни освітлення. Якщо об'єкти мають виражену текстуру, сцена добре освітлена, видимий простір невелике, то краще використовувати методи оптичного потоку.

Дані зображення, за якими визначається поза об'єкта, можуть бути як окремим зображенням, стереопарою зображення, так відеорядом. Предмети, які розглядаються, можуть бути досить загальними, включаючи живу істоту або частини тіла, наприклад, голову чи руки. Однак методи, які використовуються для визначення пози об'єкта, як правило, специфічні для класу об'єктів, і, як правило, не можна очікувати, що вони будуть добре працювати для інших типів об'єктів.

### 1.3 Огляд сучасних задач визначення положення та орієнтації

Процес визначення положення та орієнтації об'єкта можна розділити на два основні етапи. Перший етап включає знаходження та відстеження об'єкта на зображенні. Друга частина роботи проводиться по визначенню ключових точок 2D зображення та зіставлення з 3D простором аналізованого об'єкта. Тому сучасні методи дуже залежні від форми, кольору та текстурі шуканого об'єкта. За цією специфікою можна розділити область оцінки пози в залежності від типу об'єктів які досліджуються.

Найбільш сучасною задачею визначення положення та орієнтації об'єкта можна назвати оцінку пози людини. Для цієї задачі існує найбільша кількість реалізованих бібліотек і написано багато статей та наукових робіт. Зазвичай сучасні підходи розбивають тіло людини на ключові точки та визначають спеціальні точки-сполучення ключових елементів. Наприклад, сучасна бібліотека OpenPose має можливість визначати декілька людських тіл, рук, лиць і ступень (всього 135 ключових точок) на окремих зображеннях в реальному часі.

Більш специфічним але не менш значущим напрямом можна назвати визначення положення та орієнтації голови людини. Цей напрям дуже активно розвивається у наш час через те, що наявна певна кореляція між увагою та положенням голови, тобто напрямом погляду людини[5].

Також існує більш вузько доменна область для визначення емоцій та самопочуття людини через фіксацію камери лица та подальшим аналізом почуття, настрою, тощо. Цей напрям зазвичай працює з опорними точками лица людини, та в залежності від зміщення точок визначаються стан та емоції людини[6].

Не менш значущим та перспективним напрямом у наш час є напрям медичних досліджень та обстежень[7]. Задача знаходження положення та орієнтації об'єктів дуже допомагає при обстеженнях та аналізі результатів обстеження. Так як кожна пухлина чи матеріальний об'єкт можна використати для визначення положення та орієнтації більш точно ніж це можуть зробити лікарі[8]. Саме тому це є дуже актуальним напрямом розвитку області оцінки пози.

Також дуже важливим сучасним напрямом можна назвати робототехніку, бо саме розуміння положення та орієнтацію об'єктів навколо нас – це одне з найбільших відмінностей людей від роботів. Саме тому, чим швидше та точніше оцінка пози буде визначена тим швидше робот чи роботизований пристрій зможуть реагувати на той чи інший об'єкт.

Зараз існує багато під-областей які швидко розвиваються та без яких ми вже не можемо представити наше теперішнє. Виходячи з сьогоденних тенденцій можна зробити висновок, що визначення положення та орієнтації об'єктів будуть удосконалюватися з кожним роком.

#### 1.4 Аналіз методів визначення положення та орієнтації

Усі методи для визначення положення та орієнтації об'єктів можна класифікувати за двома критеріями: методи засновані на навчанні та методи не

засновані на навчанні. Методи на основі машинного навчання зазвичай використовують CNN, регресії або інші методи засновані на глибокому навчанні для підготовки моделі. Підходи, що не використовують глибоке навчання, належать до наступних двох категорій.

Методи, що базуються на 2D-інформації, в основному використовують 2D-інформацію про сцену, наприклад RGB зображення. Методи, що базуються на 3D-інформації, у свою чергу використовують 3D-інформацію про сцену, таку як хмари точок та інформацію або зображення RGB-D. Обидва підходи обчислюють ключові точки або ключові особливості та узгоджують вхідне зображення з найбільш подібним зображенням у наборі даних відповідно до ключових точок або особливостей[9]. Але вони також мають деякі важливі відмінності.

Що стосується підходів на основі 2D-інформації, то їх основною метою є знайти кореляцію між вхідним зображенням та одним із зображень шаблону за допомогою 2D-інформації, що міститься на зображенні. 2D-інформаційні підходи перетворюють оцінку пози в проблему зіставлення зображень. Результати відповідності мають великий вплив на результати оцінки пози.

Методи на основі 2D-інформації можна розділити на підходи на основі реальних зображень та підходи на основі CAD-зображень відповідно до типу використовуваного шаблону[10]. Коли підхід використовує реальні зображення як шаблон, він належить до підходів на основі реальних зображень. Якщо в підході використовуються зображення, генеровані за допомогою сторонніх систем, це можна розглядати як підхід на основі генерованих зображень.

Загалом підходи на основі згенерованих зображень є більш точними, ніж підходи на основі реальних зображень, оскільки зображення, генеровані моделями, містять мало шуму. Однак іноді моделі неможливо отримати легко, тому реальні зображення використовуються як шаблон у таких ситуаціях.

Підходи, засновані на 3D-інформації, також зосереджені на відповідності між вхідними даними та набором даних; однак вони використовують 3D-інформацію про об'єкт, таку як хмари точок та зображення RGB-D. Підходи, засновані на 3D-інформації, можна розділити на дві категорії. Основна ідея підходів на основі

відповідності полягає у безпосередньому узгодженні вхідного зображення та шаблону та прийнятті пози об'єкта відповідного шаблону як результату оцінки пози вхідного зображення.

Місцеві підходи, засновані на дескрипторі, вимірюють позу, використовуючи відповідність між дескриптором вхідних зображень та шаблонами[11]. Підходи на основі відповідності вимагають великого сховища, щоб зберегти достатню кількість шаблонів для забезпечення точності оцінки пози, і чим більше шаблонів у неї є, тим точнішим буде результат оцінки пози.

Основною метою методів, заснованих на навчанні, є підготовка відповідної моделі для визначення пози об'єкта у невідомій ситуації відповідно до даних тренінгу. Для вимірювання пози об'єкта можна використовувати багато типів моделей, наприклад, моделі регресії та моделі CNN.

Існує багато методів класифікації підходів, заснованих на навчанні. Підходи, засновані на ключових точках, застосовують двоступеневу категорію для вимірювання пози об'єкта, що легше реалізувати, ніж інші підходи.

Тим часом метою цілісних підходів є підготовка наскрізної мережі для вимірювання пози об'єкта. Він розглядає зображення в цілому і намагається передбачити розташування та орієнтацію об'єкта за один крок і дискредитувати простір, перетворюючи завдання оцінки пози в завдання класифікації. Однак цілісні підходи є більш складними та трудомісткими, ніж підходи, засновані на ключових точках.

## 1.5 Актуальність задачі

Оцінка пози - типова проблема в галузі обробки зображень, мета якої є порівняння або злиття зображень, отриманих за різних умов. В останні роки багато досліджень зосереджувались на алгоритмах оцінки пози, але до сих пір існує багато

проблем, таких як ефективність, складність та точність для різних цілей та умов, у галузі дослідження алгоритмів та практичних застосувань.

Знання пози предметів є важливою передумовою для багатьох робото технічних захоплень та маніпуляцій.

Оцінка пози об'єкта - це давня проблема і відкрите поле досліджень з перших днів комп'ютерного зору. Завдання оцінки пози є складною через різноманітність об'єктів у реальному світі, шум датчиків, закрашення, оклюзію на сцені та різні умови освітлення, які впливають на зовнішній вигляд об'єктів на зображенні.

Також, значною проблемою є об'єктні симетрії, які часто зустрічаються в техногенному та виробничому середовищі, призводять до неоднозначностей і намагаються бути вирішеними.

В останні роки в дослідженнях оцінки пози об'єкта переважають підходи на основі загорткової нейронної мережі. Ультрасучасні підходи вирішують це як проблему класифікації, коли простір пози дискретизується, а мережа використовується для прогнозування цього простору.

## 2 АНАЛІЗ АЛГОРИТМІВ ТА ПОСТАНОВКА ЗАДАЧІ

### 2.1 Алгоритм визначення положення на основі вірогідного підходу

Вхідними даними для опису структури сцени є відео зображення. При аналізі використовується опис тривимірних об'єктів на основі комплексної інформації про форму елементів, взаємне положення примітивів в просторі, пропорціях, яскравості, характеристиках фону сцени, тощо.

Для реалізації алгоритму необхідно навчання системи, в результаті якого буде сформований набір зображень об'єкта в різних положеннях. Відбувається зйомка об'єкта з різних положень. Передбачається, що під час переміщення камери по колу, в центрі якої розташовується об'єкт. Зйомка проводиться через однакові кутові відстані.

Для розпізнавання об'єкта необхідно виділити ознаки на зображенні. Наприклад, якщо даний об'єкт являє собою паралелепіпед, то на зображенні він представлений видимими гранями. Таким же чином можна провести декомпозицію зображення реального об'єкта. Для цього завдання використовуються глобальні дескриптори зображення - гістограми орієнтованих градієнтів. Безліч дескрипторів гістограми орієнтованих градієнтів, локалізованих в одному регіоні, дозволяють судити про конкретну частину об'єкта та виділити в кожному регіоні центр. Сукупність виділених таким чином точок буде характеризувати стан і орієнтацію об'єкта.

При знаходженні відповідностей розглядаються виділені області на зображенні-ідеалі і на уже згадуваному зображенні. Для знаходження відповідностей пропонується використовувати алгоритм очікування-максимізації (Expectation-maximization)[14]. Алгоритм дозволяє ітеративно обчислювати значення максимальної правдоподібності при наявності прихованих даних.

Основна ідея алгоритму полягає в визначенні моделі для деякого набору даних і багаторазовому обчисленні параметрів. Обчислення максимальної правдоподібності дозволяє оцінити параметри моделі, для яких спостерігаються

дані є найбільш ймовірними показниками. EM-алгоритм ітеративно виконується в два етапи. На першому кроці (очікування) відсутня інформація обчислюється з урахуванням спостережуваних даних і поточних параметрів моделі. На другому кроці (максимізація) функція правдоподібності досягає максимуму за умови, що відсутні дані відомі. На етапі максимізації алгоритм використовує поточні значення відсутніх даних. Для забезпечення збіжності EM-алгоритму повинна збільшуватися ймовірність на кожній ітерації[14].

У даному випадку зображення форми об'єкта представлено множиною точок довільного розміру. Порядок точок не має значення. Отримуємо дві множини точок згідно з формулою (2.1).

$$\begin{aligned} X &= (x_1, x_2 \dots, x_m)^T \in R^{M \times 2}, \\ Y &= (y_1, y_2 \dots, y_m)^T \in R^{N \times 2} \end{aligned} \quad (2.1)$$

де  $x_1, x_2 \dots, x_m$  – множина точок форми об'єкта;

$y_1, y_2 \dots, y_m$  – множина точок форми знайденого об'єкта;

$T \in R^{N \times 2}$  – вектор переносу;

$M$  – область точок начального зображення;

$N$  – область точок дійсного зображення.

Завдання полягає в знаходженні відповідностей між  $X$  і  $Y$ . Припустимо, що  $y_n$  є спостереженням з моделлю суміші згідно з формулою (2.2).

$$\begin{aligned} p(y_n) &= p(y_n|v_f)p(v_f) + p(y_n|v_b)p(v_b), \\ p(y_n|v_f) &= \sum_{m=1}^M p(y_n|m, v_f)p(m) \end{aligned} \quad (2.2)$$

де  $v_b$  – компонент суміші, який представляє «модель фону», яка дозволяє враховувати можливість викидів;

$v_f$  – компонент суміші, який описує об'єкти, що розпізнаються;

$y_1, y_2 \dots, y_m$  – множина точок форми знайденого об'єкта;

$M$  – область точок начального зображення;

$m$  – компонент суміші.

Зміна положення об'єкта описується параметрами: параметр масштабування; матриця повороту; вектор переносу.

Призначимо рівну ймовірність для кожного компонента суміші згідно з формулою (2.3):

$$p(m) = 1/M \quad (2.3)$$

де  $p$  – ймовірність для кожного компонента;

$m$  – компонент суміші;

$M$  – область точок начального зображення;

Це запобігає відповідності одного об'єкта зображення декільком об'єктам іншого зображення. Оцінки максимальної правдоподібності для інших параметрів моделі  $s$ ,  $D$ ,  $t$  і  $p(v)$  знаходяться за допомогою ЕМ-алгоритму.

Перший крок. Обчислити вираз, використовуючи поточні значення параметрів згідно з формулою (2.4).

$$p(m, v_f | y_n) = \frac{p(y_n | m, v_f) p(m) p(v_f)}{p(y_n)} \quad (2.4)$$

де  $p$  – ймовірність для кожного компонента;

$y_1, y_2 \dots, y_m$  – множина точок форми знайденого об'єкта;

$m$  – компонент суміші.

Після першого кроку потрібно оновити параметри згідно з формулою (2.5).

$$p(v_f) = \frac{1}{N} \sum_n p(v_f|v_n) p(v_b) = 1 - p(v_f),$$

$$(s, D, t) = \mathop{\text{arg}} \sum_{m,n} p(m, v_f|y_n) \|y_n - sDx_m - t\|^2 \quad (2.5)$$

де  $s$  – параметр масштабування;

$D$  – матриця повороту;

$t \in R^2$  – вектор;

$m$  – компонент суміші;

$M$  – область точок начального зображення;

$N$  – область точок дійсного зображення;

$p$  – ймовірність для кожного компонента;

$y_1, y_2 \dots, y_m$  – множина точок форми знайденого об'єкта.

Формула описує завдання знаходження відповідностей між двома множинами точок розмірами  $N$  і  $M$ . Ступінь відповідності пари  $(x_m, y_n)$  задається величиною  $p(m, v_f|y_n)$ .

## 2.2 PoseCNN алгоритм

PoseCNN - нова згортова нейронна мережа (CNN) для оцінки пози об'єкта. Ключовою ідеєю PoseCNN є розбиття завдання оцінки пози на різні компоненти, що дозволяє мережі чітко моделювати залежності та незалежності між ознаками об'єкта[15]. Зокрема, PoseCNN виконує три пов'язані завдання. По-перше, він визначає мітку об'єкта для кожного пікселя на вхідному зображенні. По-друге, він оцінює 2D-піксельні координати центру об'єкта, передбачаючи одиничний вектор від кожного пікселя до центру.

Використовуючи семантичні мітки та пікселі зображення, пов'язані з об'єктом, проходить голосування за розташування центру об'єкта на зображенні.

Крім того, мережа також оцінює відстань до центру об'єкта. Припускаючи відомі властивості камери, оцінку центру 2D-об'єкта та його відстані, дозволяє нам відновити його 3D-переклад  $T$ .

Нарешті, 3D-обертання  $R$  оцінюється шляхом регресії згорнутих ознак, витягнутих всередині обмежувальної рамки об'єкта, до подання кватерніону  $R$ . Голосування за 2D-центр з наступною регресією обертання для оцінки  $R$  і  $T$  може бути застосовано до текстурованих та без текстурних об'єктів і бути стійким до оклюзій, оскільки мережа навчена голосувати по об'єктних центрах, навіть коли вони закриті.

Обробка симетричних об'єктів є ще однією проблемою для оцінки пози, оскільки різні орієнтації об'єктів можуть генерувати однакові спостереження. Наприклад, неможливо однозначно оцінити орієнтацію червоної чаші або дерев'яного блока, хоча набори базових даних пози, такі як набір даних OccludedLINEMOD, розглядають спеціальну симетричну оцінку для таких об'єктів (симетрії зазвичай ігнорується під час нейронного навчання)[15]. Однак це може призвести до поганих результатів навчання, оскільки мережа отримує непослідовність сигналу втрат, такі як великі втрати навіть на орієнтації об'єкта хоча оцінка з боку мережі є правильною стосовно симетрії об'єкта.

Натхненні цим спостереженням, розробники алгоритму представили ShapeMatch-Loss - функцію втрат, яка фокусується на узгодженні тривимірної форми об'єкта. Як зазначено, ця функція втрат дає чудову оцінку для об'єктів з симетрією форми.

Отримавши вхідне зображення, завдання оцінки пози 6D об'єкта полягає в оцінці жорсткого перетворення від об'єктної системи координат  $O$  до системи координат камери  $C$ . Припускається що доступна тривимірна модель об'єкта та системи координат об'єкта визначається у тривимірному просторі моделі.

Жорстке перетворення тут складається з перетворення яке містить 3D обертання  $R$  і 3D переклад  $T$ , де  $R$  визначає кути повороту навколо осі  $X$ , осі  $Y$  та осі системи координат об'єкта  $O$ , а  $T$  – координата початку координат  $O$  в системі координат камери  $C$ .



## 2.3 Pix2Pose алгоритм

Pix2Pose прогнозує 3D-координати об'єкта, використовуючи RGB-зображення без текстурованих 3D-моделей для навчання[16]. 3D координати перекритих пікселів неявно оцінюються мережею, щоб бути надійними до оклюзії. Пропонується спеціалізована функція втрат (трансформаторна втрата) для надійної підготовки мережі до симетричних об'єктів. В результаті прогнозування кожен піксель формує 2D-3D відповідність, яка використовується для обчислення поз алгоритмом.

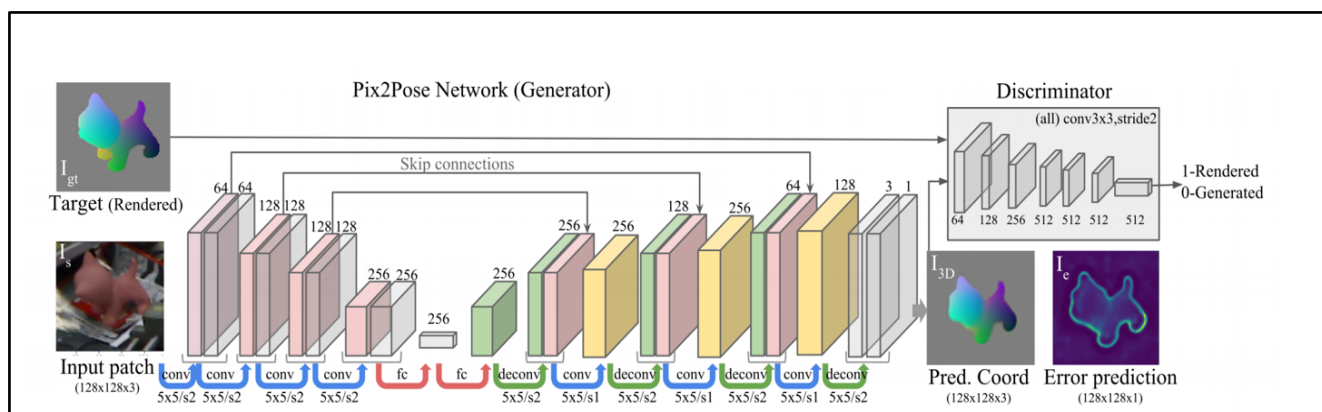


Рисунок 2.2 – Архітектура мережі Pix2Pose

Архітектура мережі Pix2Pose зображена на рисунку 2.2. Вхідні дані мережі – це обрізане зображення  $I_s$ . Використовується обмежувальне поле виявленого класу об'єктів. Вихідні дані мережі - це нормовані тривимірні координати кожного пікселя  $I_{3D}$  в координаті об'єкта та оцінені помилки  $I_e$  кожного передбачення,  $I_{3D}$ ,  $I_e = G(I_s)$ , – де  $G$  позначає мережу Pix2Pose.

Цільовий результат включає координатні передбачення закупорених частин, що робить передбачення більш надійним до часткового загородження. Оскільки координата складається з трьох значень, подібних до значень RGB на зображенні, вихідний  $I_{3D}$  можна розглядати як кольорове зображення. Отже, вихідне зображення легко отримати, відтворюючи кольорову модель координат пози істинного зображення.

Приклад значень тривимірних координат у кольоровому зображенні наведено на рисунку 2.3 Прогнозування похибки  $I_e$  розглядається як показник достовірності кожного пікселя, який безпосередньо використовується для визначення відхилення та зменшення пікселів перед обчисленням пози. Обрізане зображення зменшується до  $128 \times 128$  пікселів із трьома каналами для значень RGB.

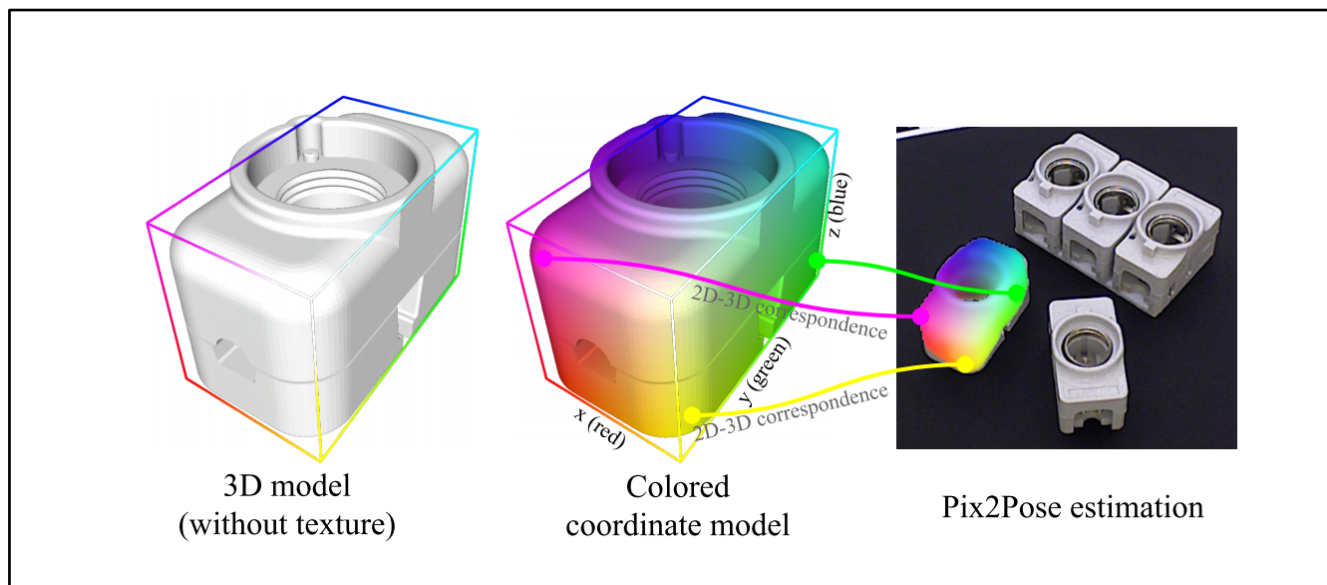


Рисунок 2.3 – Приклад значень тривимірних координат у кольоровому зображенні

Для підтримання деталей карт функцій низького рівня додаються пропуски з'єднань шляхом копіювання половинних каналів виходів з перших трьох шарів у відповідні симетричні шари в декодері, що призводить до більш точної оцінки пікселів навколо геометричних меж.

Розмір фільтра кожного шару згортки та деконволюції зафіксовано до  $5 \times 5$  з кроком 1 або 2, позначеним як  $s1$  або  $s2$  на рис. 1.2 На горловину пляшки наносяться два повністю з'єднані шари між 256 кодами та декодером. Нормалізація партії та функція активації застосовуються до кожного виходу проміжних шарів, крім останнього. В останньому шарі вихід з трьома каналами та активацією створює тривимірне зображення координат  $I_{3D}$ , а інший вихід з одним каналом та активацією сигмоїда оцінює очікувані помилки  $I_e$ .



в координатній системі відліку об'єкта відомі. Зображення точок  $M_i$  називаються  $m_i$ , а їх координати зображень  $x_i$   $y_i$  відомі.

Координати  $X_i$   $Y_i$   $Z_i$  тих самих точок  $M_i$  в системі координат камери невідомі, оскільки поза об'єкта в системі координат камери дорівнює невідомою. Алгоритм показує як безпосередньо знайти матрицю обертання та вектор перекладу об'єкта, не вирішуючи явно координати  $X_i$   $Y_i$   $Z_i$ .

Мета алгоритму - обчислити матрицю обертання та вектор перекладу об'єкта. Матриця обертання  $R$  для об'єкта - це матриця, рядки якої є координатами одиничних векторів  $i, j, k$  системи координат камери, вираженої в об'єктній системі координат ( $M_0u, M_0v, M_0w$ ).

Дійсно, метою матриці обертання є перетворення об'єктних координат векторів, таких як  $M_0M_i$ , в координати, визначені в системі камери точкового добутку  $M_0M_i$ . Між першим рядком матриці та вектором  $M_0M_i$ , правильно забезпечується проекція цього вектора на одиничний вектор  $i$  системи координат камери, тобто координату  $X_1 - X_0, M_0M_i$ , у тому випадку якщо координати  $M_0M_i$  та вектор рядків  $i$  виражені в однаковій системі координат  $M_0M_i$ .

Тому матрицю обертання можна записати згідно з формулою (2.6).

$$R = i_u \ i_v \ i_w \ j_u \ j_v \ j_w \ k_u \ k_v \ k_w \quad (2.6)$$

де  $i, j, k$  – одиничні вектори системи координат камери;

$R$  – матриця обертання для об'єкта.

Для обчислення обертання нам потрібно лише обчислити  $i$  та  $j$  в об'єктній системі координат. Вектор  $k$  потім отримується за перехресним добутком  $ij$ . Вектор перекладу  $T$  - це вектор  $OM_0$  між центром проекції  $O$  і опорною точкою  $M_0$  початком координатної системи об'єкта, тому координати вектора перекладу –  $X_0$   $Y_0$   $Z_0$ . Якщо ця точка  $M_0$  була обрана бути видимою точкою об'єкта, для якої зображення є точкою  $m$ , цей вектор перекладу  $T$  вирівнюється з вектором  $Om_0$  і дорівнює  $\frac{Z_0}{f} Om_0$ . Отже, для обчислення перекладу об'єкта нам потрібно лише

обчислити його  $z$  – координату  $Z_0$ . Таким чином, поза об'єкт повністю визначається, коли знайдено  $i, j$  та  $Z_0$ .

Геометричне тлумачення алгоритму можна описати за допомогою даного опису:

- обчислити матрицю об'єкта  $B$  як псевдообернену матрицю  $A$ ;
- припускається, що масштабовані орфографічні проєкції  $p_i$  та  $M_i$  накладаються на точки зображення;
- обчислюється приблизна поза, припускаючи, що модель проєкції є масштабом;
- змістіть точки об'єкта з знайдених приблизних положень у положення на одній і тій же глибині, але на їхній прямій видимості деформація об'єкта;
- знаходиться зображення цих зміщених точок за масштабованою орфографічною проєкційною моделі;
- якщо ці масштабовані точки орфографічного зображення не збігаються з тими, що були знайдені на попередній ітерації, повертаються до другого кроку, використовуючи ці точки зображення замість вихідних точок;
- в іншому випадку точна поза дорівнює останній приблизній позі.

## 2.5 Порівняння алгоритмів

У результаті дослідницької роботи було проаналізовані сучасні методи для визначення положення та орієнтації об'єктів у просторі.

Було з'ясовано, що серед методів, заснованих на згорткових мережах найкраще та швидше працює Pix2Pose, але недоліком даного методу можна вважати необхідність попереднього визначення 3D моделі об'єкта. PoseCNN метод потребує значно менше попередніх даних про об'єкт, проте працює з точністю приблизно 34%, коли Pix2Pose показує 63% [18]. За показником швидкості Pix2Pose

також випереджає PoseCNN приблизно у два рази. За даними BOP challenge 2020 року, Pix2Pose у середньому потребує 0.45 секунд на одне зображення.

У той час серед аналітичних підходів найкраще себе показав POSIT метод. Це перевірений метод який працює швидко але потребує якісного попереднього аналізу зображення. Найбільш значним недоліком можна назвати невміння вирішувати проблему оклюзії.

Якщо порівнювати Pix2Pose та POSIT – найбільшої точності можна досягти завдяки POSIT, але він потребує якісного попереднього визначення опорних точок. Проте Pix2Pose та всі CNN алгоритми потребують запуску на обладнанні із підтримкою CUDA.

## 2.6 Постановка задачі та функціональні вимоги до системи

Метою роботи є дослідження та аналіз методів для розробки програмної системи визначення положення та орієнтації об'єктів у просторі.

Практичною задачею роботи є розробка додатку для аналізу методів оцінки визначення положення та орієнтації об'єктів у просторі за зображеннями на прикладі додатку розпізнавання пози об'єкта який збирає дані про витрачені додатком ресурси та заміряє час роботи обраного алгоритму та виводить оцінку результату роботи.

У якості вхідних даних додатку було обране зображення яке містить об'єкт аналізу, контрольні точки для визначення сцени, деяка інформація про об'єкт, що залежить від обраного алгоритму.

Вихідними даним додатку будуть координати та кути нахилу об'єкта порівняно с начальною (ідеальною) позицією і метрики для аналізу кожного алгоритму.

Для досягнення мети магістерської роботи були поставлені наступні завдання:

- збір теоретичних відомостей про алгоритми визначення положення та орієнтації об'єктів у просторі;
- реалізація програмного застосунку, який буде реалізувати обрані алгоритми та видаватиме результат, що відповідає вимогам до вихідних даних і аналіз щодо роботи самого алгоритму;
- оцінка якості роботи програмного застосунку;
- оцінка продуктивності роботи кожного алгоритму в плані часу, що потребується для виконання задачі, та ресурсів, що необхідні для комфортного користування;
- підведення підсумків дослідження.

Програмна система повинна виконувати наступні функції:

- загрузка зображення користувачем;
- визначати положення та орієнтацію об'єкта;
- аналізувати ресурси які використовує алгоритм;
- зберігати результат роботи з інформацією о процесі визначення.

Опис функціональних вимог:

- при вході у систему користувач повинен ввести адресу електронної пошти та пароль;
- при вході у систему, користувач повинен бачити інтерфейс доступних алгоритмів та кнопку запуску алгоритму;
- система повинна надавати інформацію про доступні об'єкти, відображати наявні алгоритми;
- система повинна надавати інформацію про кожний запуск розпізнавання з інформацією про використані ресурси.

Отже, поставлена задача провести дослідження методів визначення положення та орієнтації об'єктів у просторі. Розробити програмну систему для визначення положення і орієнтації об'єктів за допомогою обраних алгоритмів та збору метрик і оцінок ефективності роботи кожного алгоритму.

## 3 ПРОЕКТУВАННЯ ДОДАТКУ

### 3.1 UML моделювання

UML – це стандартизована мова моделювання, що складається з інтегрованого набору діаграм, розроблених для допомоги розробникам систем та програмного забезпечення для визначення, візуалізації, побудови та документування артефактів програмних систем, а також для бізнес-моделювання та інші непрограмні системи. UML представляє колекцію найкращих інженерних практик, які довели свою успішність у моделюванні великих та складних систем[19].

#### 3.1.1 Діаграма прецедентів

Це UML діаграма, на якій зображені дії користувача та очікуваний результат. Діаграма прецедентів показана на рисунку 3.1

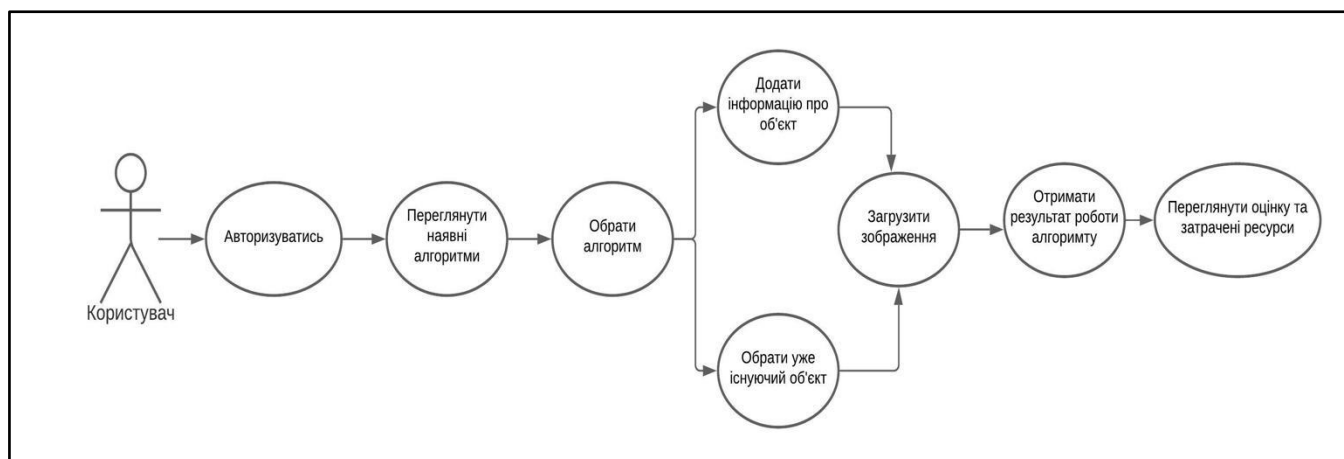


Рисунок 3.1 – UML діаграма прецедентів

Діаграми прецедентів цінні для візуалізації функціональних вимог системи, що перетвориться на вибір проекту та пріоритети розвитку. Вони також

допомагають визначити будь-які внутрішні чи зовнішні фактори, які можуть вплинути на систему і їх слід враховувати.

Вони забезпечують хороший високоякісний аналіз пози системою. Діаграми прецедентів вказують, як система взаємодіє з акторами, не турбуючись про деталі реалізації цієї функціональності.

### 3.1.2 Діаграма станів

Діаграма станів аналізу положення та орієнтації об'єкта показана на рисунку 3.2.

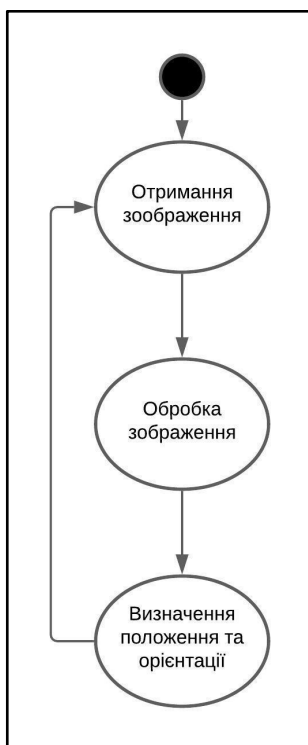


Рисунок 3.2 – UML діаграма станів аналізу положення та орієнтації об'єкта

Діаграма станів – орієнтований граф для кінцевого автомата, в якому: вершини позначають стани, а дуги показують переходи між двома станами.

Аналіз положення та орієнтації об'єкта відбувається наступним чином, спочатку отриманий кадр обробляється та фільтрується, потім фільтровані дані

аналізуються реалізацією обраного алгоритму. Алгоритм повертає зображення з визначеними координатами та кутами нахилу об'єкта.

### 3.1.3 Діаграма компонентів

Системи визначення положення та орієнтації об'єкта є відносно важкими для апаратної частини. Саме тому було вирішено розділити компоненти системи на 4 окремі функціональні блоки. Діаграма компонентів системи зображена на рисунку 3.3

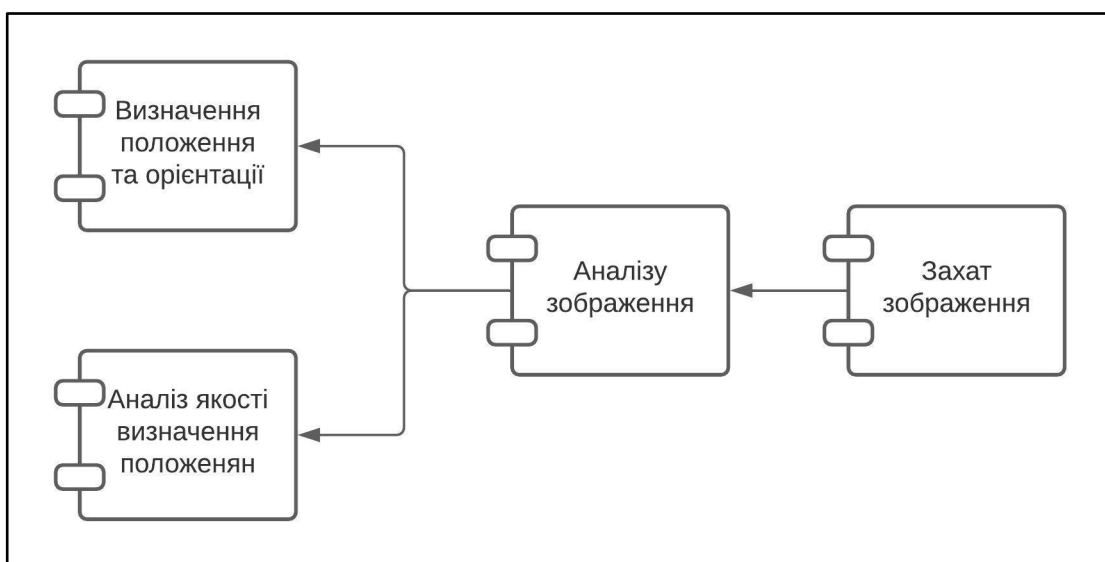


Рисунок 3.3 – UML діаграма компонентів системи

Компонент «Захват зображення» на вхід отримує дані користувача у форматі зображення та є інтерфейсом програмного забезпечення. Він буде реалізовувати наступні функції:

- захват зображення;
- фільтрацію вхідного зображення;
- передачу зображення до блоку аналізу зображення з інформацією про обраний алгоритм.

Компонент аналізу зображення отримує дані від компонента захвату зображення та передає компонентам визначення положення і орієнтації об'єкта та аналізу якості визначення положення та орієнтації. Отримує від них результати аналізу та повертає оброблені дані до компонента захвату зображення, який має зв'язок з базою даних, зберігає та передає до компонента захвату зображення .

### 3.2 Моделювання схеми бази даних

Схема бази даних показана на рисунку 3.4.

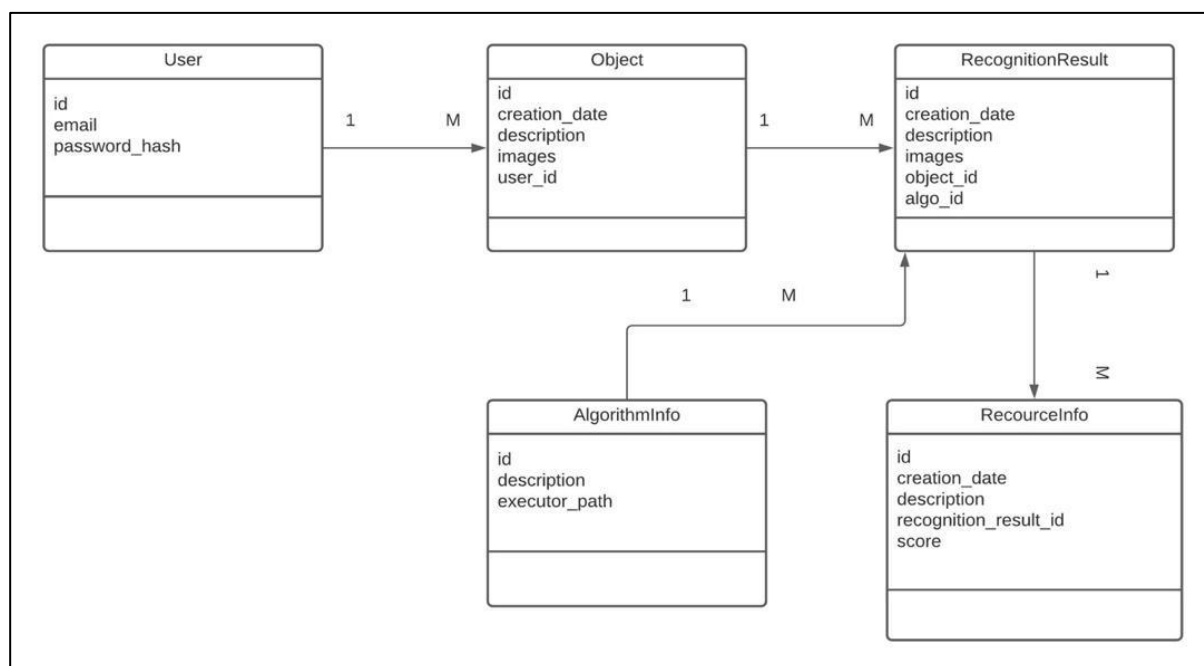


Рисунок 3.4 – Схема бази даних

Планується зберігати в системі дані про користувачів та додані об'єкти з необхідною інформацією про них. Також, буде зберігатися кожен результат роботи алгоритму з описом використаного алгоритму та інформацією про використання ресурсів та оцінкою роботи.

## 4 ОПИС ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

### 4.1 Вибір інструментів реалізації

Для реалізації додатку спочатку необхідно визначити найбільш підходящі інструменти реалізації. Дана система поділяється на 3 модулів розробки і для кожного потрібно визначити свій інструмент та середу розробки.

Операційна система для розробки додатку має бути побудована на unіx-подібній платформі для можливості зручного запуску CUDA-рішень та можливістю використання контейнерів. З найбільш сучасних та відомих систем є MacOS та Ubuntu. MacOS має багато специфіки у комунікації з апаратним забезпеченням та має небезкоштовне програмне забезпечення. Отже, з найбільш доступних та підходящих варіантів було обрано Ubuntu 18.04 як безкоштовна та універсальна операційна система для розробки.

Мову програмування було обрано Python 3 тому, що вона має широку підтримку у data science угруповань та добре документовану бібліотеку для взаємодії з OpenCV.

У якості фреймворку для розробки серверної частини додатку було обрано Flask - легковісний python фреймворк для розробки невеликих серверних рішень. Flask є одним із самих зручних інструментів для побудови легких серверів та має багату підтримку та повну документацію.

Для реалізації клієнтської частини додатку було обрано Vue.js – javascript фреймворк з відкритим кодом для реалізації реактивних односторінкових додатків. Vue.js має широку підтримку та добре написану документацію та добре підходить для побудови невеликих односторінкових рішень.

У якості бази даних для зберігання необхідної інформації для роботи додатку було обрано PostgreSQL як найбільш прогресивна база даних для unіx-подібних систем з вбудованою підтримкою Python.

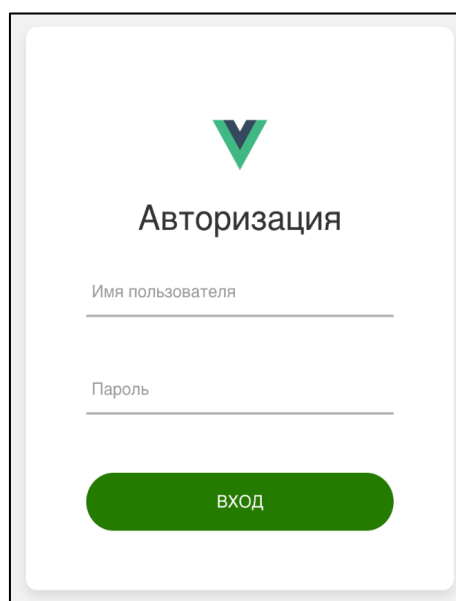
Отже, для реалізації додатку визначення положення та орієнтації об'єктів у просторі на базі операційної системи Ubuntu 18.04 були обрані наступні інструменти:

- Vue.js у якості javascript фреймворку для реалізації інтерфейсу;
- Flask у якості python фреймворку для реалізації API інтерфейсу та основної бізнес логіки додатку;
- PostgreSQL був обраний як база даних для зберігання інформації;
- REST підходи у якості архітектури обміну інформацією у мережі;
- PyCharm у якості середовища розробки.

PyCharm середовище програмування було обране через те, що має додаток для операційної системи Ubuntu та представляє безкоштовну підписку для студентів.

## 4.2 Опис інтерфейсу та функціональності

Для того щоб скористатися системою, користувачу необхідно авторизуватися за допомогою форми авторизації (див. рис. 4.1).



The image shows a login form with a white background and a thin grey border. At the top center is a green logo consisting of a downward-pointing 'V' shape. Below the logo, the word 'Авторизация' is written in a bold, black, sans-serif font. Underneath, there are two input fields. The first is labeled 'Имя пользователя' and the second is labeled 'Пароль'. Both labels are in a small, grey font. At the bottom of the form is a large, rounded green button with the word 'ВХОД' written in white, bold, uppercase letters.

Рисунок 4.1 – Форма авторизації у системі

Авторизація користувача відбувається за логіном та паролем.

Першою сторінкою додатку є сторінка з відео зображенням так з таблицею доступних алгоритмів (див. рис. 4.2).

Algorithms				
№	name	status	description	action
0	Posit with Chessboard	✔	Default algo with chessboard object	STOP
1	Posit with planar	✘	Used with planar objects	SELECT
2	Pix2Pose	✘	Not released yet	SELECT

Рисунок 4.2 – Таблиця наявних алгоритмів та інформації про них.

Користувач може обрати алгоритм за допомогою кнопки з назвою «select». Та одразу побачити роботу алгоритму у формі знизу (див. рис. 4.3).

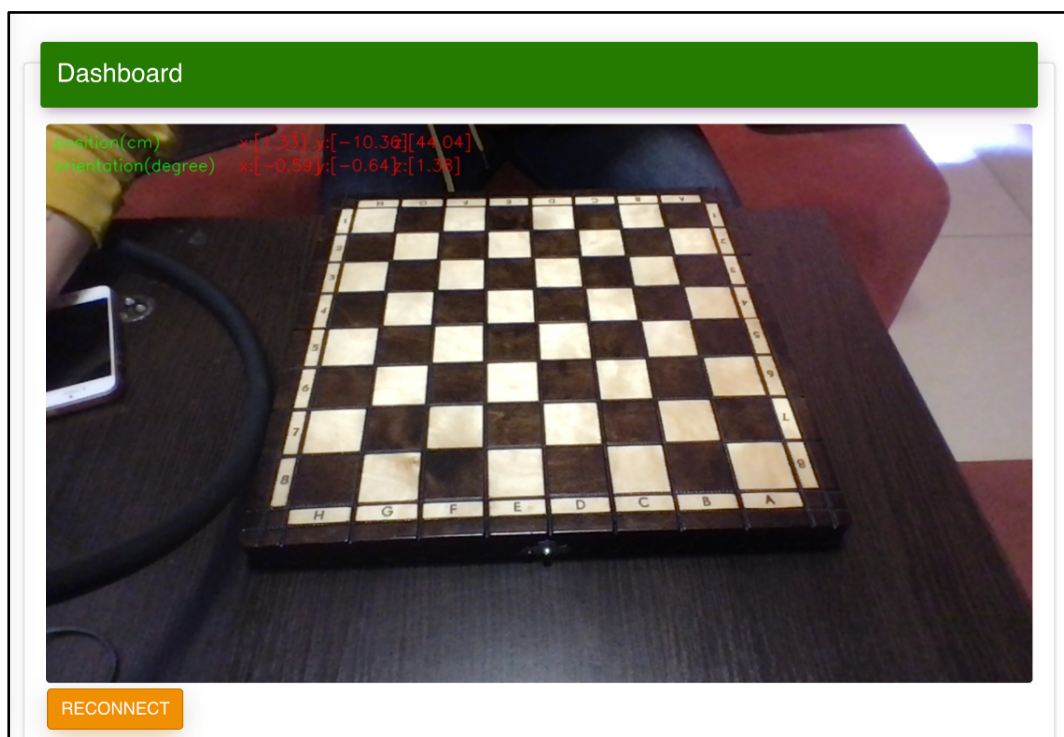


Рисунок 4.3 – Результат роботи обраного алгоритму

У лівому верхньому куті зображення знаходиться інформація про визначену позицію та орієнтацію об'єкта.

Користувач також має можливість переглянути дані щодо калібровки камери у вкладці меню з назвою «calibration». На цій сторінці доступна інформація щодо результату калібровки та показане зображення яке було використано (див. рис. 4.4).

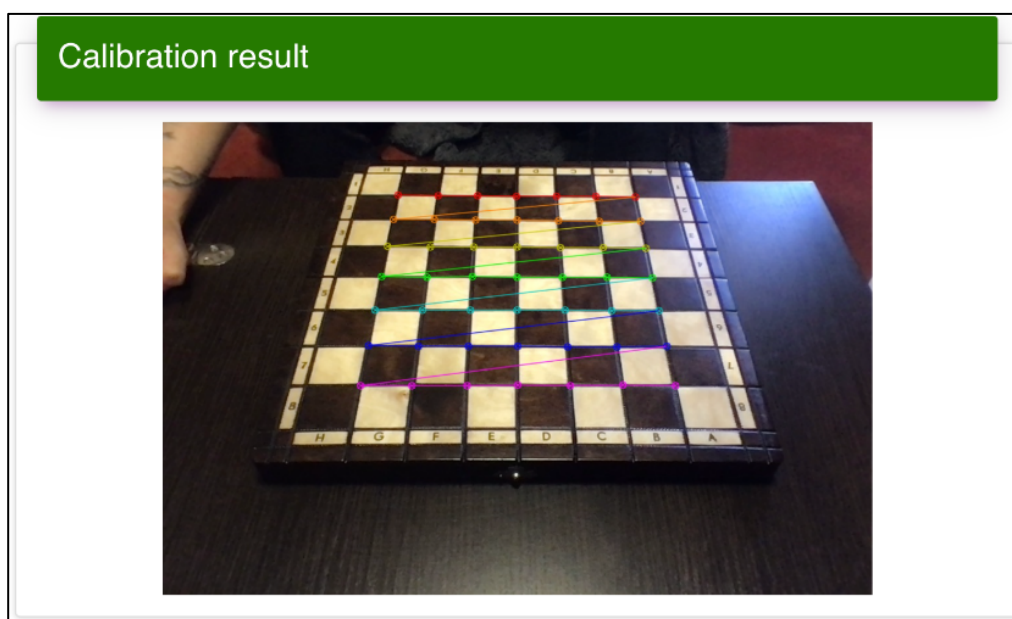


Рисунок 4.4 – Зображення з даними про калібровку камери

Користувач має можливість змінити зображення для калібровки камери. Для цього необхідно загрузити нове зображення за допомогою форми (див. рис. 4.5).

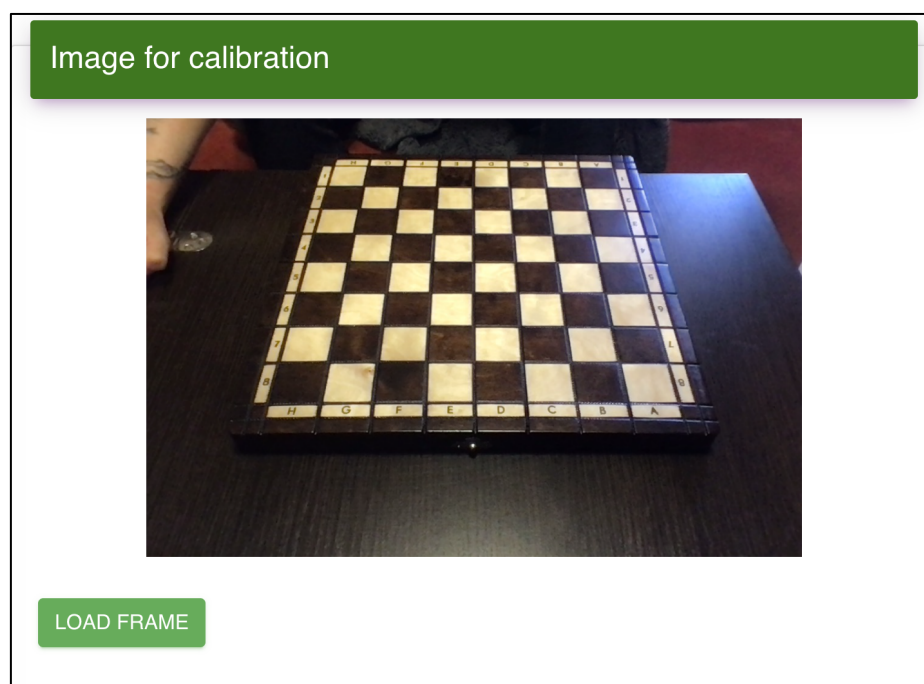


Рисунок 4.5 – Форма загрузки зображення для калібровки

Після зміни зображення додаток запустить процес калібровки з новими даними.

У ході роботи користувач має можливість слідкувати за показниками системи за допомогою карток з постійно оновлюємо інформацією (див. рис. 4.6).

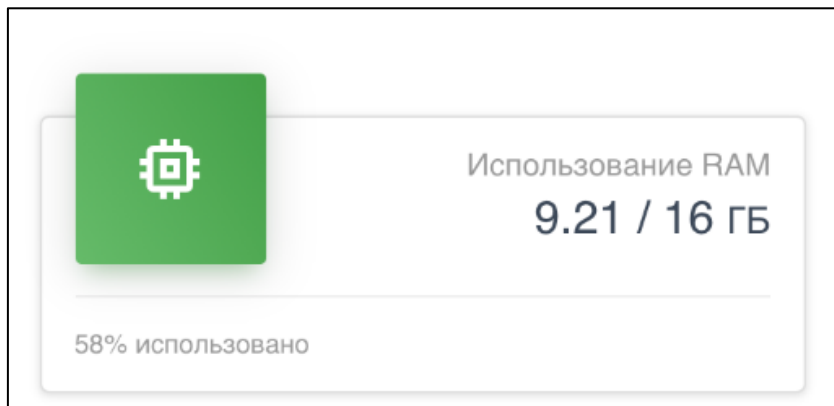


Рисунок 4.6 – Картка з інформацією о системі

Коли ресурси системи доходять до критичного значення колір картки змінюється на червоний задля привернення уваги користувача. У даний час в системі реалізовані картки з інформацією об використанні оперативної пам'яті, процесорних ресурсів та використання жорсткої пам'яті.

Дані о зареєстрованих об'єктів доступні на сторінці настройки (див. рис. 4.7).

Registered objects		
№	name	image
0	box	/models/box/resized_IMG_3875.jpg
1	chessboard	/models/chessboard/image.jpg

Рисунок 4.7 – Таблиця з даними о зареєстрованих об'єктах.

У таблиці представлені назви та шлях до зображення зареєстрованих об'єктів. Назви об'єктів це ті назви які користувач заніс у базу даних при зберіганні нового об'єкту.

### 4.3 Апробація результатів роботи та тестування додатку

У ході магістерської роботи був розроблений додаток для визначення положення та орієнтації об'єкта. У якості даних для тестування додатку були взяті зображення з датасету VOP створеного під керівництвом google та побутові предмети у якості книги, шахової дошки та резинової утки. Ці предмети були обрані через просту прямокутну форму для можливості швидкого моделювання форми об'єктів та через те, що частина моделей для цих предметів вже є у обраному датасеті.

Результати тестування додатку визначення положення та орієнтації показали, що результат роботи алгоритмів залежить від текстур об'єктів, наприклад для книги без яскравої текстури, було дуже проблемно визначити контрольні точки. Також ця проблема була присутня через глянцеvu поверхню книг, яка добре віддзеркалювала світло від поверхні.

Незважаючи на проблемні місця реалізованих алгоритмів, при наявності правильного освітлення і підходящої поверхні та форми об'єктів, реалізований додаток знаходить положення та орієнтацію об'єктів.

POSIT алгоритм працює тільки для компланарних об'єктів, швидкість його роботи напряму залежить від алгоритму визначення ключових точок. При тестуванні було використано масштабно-інваріантне перетворення ознак. При якому обробка одного кадру з ~20 ключовими точками працювала приблизно 0.45с.

Pix2Pose алгоритм показав швидкість роботи для обробки одного кадру приблизно в 4.85с. Результати роботи були оцінені за допомогою метрики середньої відстані, яка показала результат для качки та книги у 22.7% та 45.3% відповідно.

PoseCNN алгоритм показав швидкість роботи для обробки одного кадру приблизно в 1.87с. Результати роботи були оцінені за допомогою метрики

середньої відстані, яка показала результат для качки та книги у 20.2% та 35.1% відповідно.

Також, сам додаток було протестовано за допомогою мануального тестування, у ході якого значних несправностей не було знайдено. Були знайдені декілька недоліків, як наприклад якість шрифту та кольору відображення результатів роботи додатку, які є трішки непомітними на деяких сценах тестування. Всі ці недоліки плануються бути виправленими у наступній версії додатку.

## ВИСНОВКИ

У результаті дослідницької роботи було проаналізовані сучасні методи для визначення положення та орієнтації об'єктів у просторі, було спроектовано додаток з реалізацією існуючих підходів для вирішення задачі оцінки пози об'єктів. Були прочитані та проаналізовані нові дослідження у області визначення положення та орієнтації об'єктів. Також, було проведено дослідження сучасних реалізацій методів на основі існуючих алгоритмів, реалізованих бібліотек та інструментів.

Були розглянуті перспективні напрямки розвитку та області застосування сучасних методів вирішення задачі знаходження пози об'єктів, проаналізовані попередні роботи кафедри ПІ у цьому напрямку. Зазначені подальші можливості для розвитку цього напрямку та удосконалення рішень для інших напрямків досліджень.

Було з'ясовано, що серед методів, заснованих на згорткових мережах найкраще та швидше працює Pix2Pose, але недоліком даного методу можна вважати необхідність попереднього визначення 3D моделі об'єкта.

У той час серед аналітичних підходів найкраще себе показав POSIT метод. Це перевірений метод який працює швидко але потребує якісного попереднього аналізу зображення. Найбільш значним недоліком можна назвати невміння вирішувати проблему оклюзії.

У ході роботи було розроблено додаток визначення положення та орієнтації об'єкту, який відкриває подальші можливості для навчання студентів, проведення дослідів з використанням цього додатку та подальшою можливістю розширення додатку з використанням нових технологій та додаванням нового функціоналу.

Елементами наукової новизни роботи є додаток з реалізацією сучасних рішень та подальшою можливістю використання його для навчання, порівняння існуючих методів у режимі реального часу, проведення досліджень, тощо.

**ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ**

1. Дэвид Форсайт, Жан Понс. Компьютерное зрение. Современный подход. — М. : «Вильямс», 2004. — 928 с.
2. Л. Шапиро, Дж. Стокман. Компьютерное зрение. — М. : Бином. Лаборатория знаний, 2006. — 752 с.
3. Linda G. Shapiro, George C. Stockman. Computer Vision. — Prentice Hall, 2006. — 152 с.
4. Выдрин Д.Ф Платформа Ардуино: преимущества. — Academy, 2017. — 152 с.
5. Rakova, A.O. Bilous, N.V. REFERENCE POINTS METHOD FOR HUMAN HEAD MOVEMENTS TRACKING // Radio Electronics, Computer Science, Control. 2020, С. 121–128.
6. Білоус Н.В. Рассоха О.В. Агекян І.А. Грамм О.В. Дослідження методів для розробки програмної системи розпізнавання емоцій та визначення стану здоров'я людини // Біоніка інтелекту. 2020. С. 65-70.
7. Galyna Shcherbakova, Shi Hao-Su, Victor Krylov, Natalia Bilous, Svitlana Antoshchuk. Estimation of the duration of RR-intervals of electrocardiograms by mean of multi-start optimization based on wavelet transformation. // IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing System. 2017. С 752-755.
8. O Hramm, N Bilous, I Ahegian. Configurable Cell Segmentation Solution Using Hough Circles Transform and Watershed Algorithm. // IEEE 8th International Conference on Advanced. 2019. С.
9. V Krylov, G Shcherbakova, R Pisarenko, N Bilous Signal restoration by means of blind deconvolution based on optimization with wavelet transformation. // IEEE Third International Scientific-Practical Conference. 2016. С. 21-23.
10. T. Pfister. Advancing Human Pose and Gesture Recognition // PhD thesis — University of Oxford, 2015. С. 128-133

11. Guo, Z.; Chai, Z.; Liu, C.; Xiong, Z. A Fast Global Method Combined with Local Features for 6D Object Pose Estimation. // IEEE/ASME International Conference on Advanced Intelligent Mechatronics — Hong Kong, China, 2019. С. 1–6.
12. Ulrich M., Wiedemann C., Steger C. CAD-based recognition of 3D objects in monocular images. // ICRA'09 Proceedings of the 2009 IEEE international conference on Robotics and Automation. 2009. С. 2090–2097.
13. Gordon I., Lowe D.G. What and where: 3D object recognition with accurate pose. // Lecture Notes in Computer Science. 2006. С. 67–82.
14. Захаров А.А., Тужилкин А.Ю., Веденин А.С. АЛГОРИТМ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ТРЕХМЕРНЫХ ОБЪЕКТОВ ПО ВИДЕОИЗОБРАЖЕНИЯМ НА ОСНОВЕ ВЕРОЯТНОСТНОГО ПОДХОДА. // Фундаментальны исследования. № 11-8. 2014. С. 1683-1687.
15. Yu Xiang, Tanner Schmidt, Venkatraman Narayanan, Dieter Fox. PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes. // IEEE International Conference on Computer Vision. 2017. С. 21-25.
16. Kiru Park, Timothy Patten and Markus Vincze. Pix2Pose: Pixel-Wise Coordinate Regression of Objects for 6D Pose Estimation. // ECCV. 2019. URL: [https://openaccess.thecvf.com/content\\_ICCV\\_2019/papers/Park\\_Pix2Pose\\_Pixel-Wise\\_Coordinate\\_Regression\\_of\\_Objects\\_for\\_6D\\_Pose\\_Estimation\\_ICCV\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCV_2019/papers/Park_Pix2Pose_Pixel-Wise_Coordinate_Regression_of_Objects_for_6D_Pose_Estimation_ICCV_2019_paper.pdf) (дата звернення: 20.02.2021).
17. Tommaso Gramegna, LeaVenturino, Grazia Cicirelli, Giovanni Attolico, Arcangelo Distante. Optimization of the POSIT algorithm for indoor autonomous navigation // Robotics and Autonomous Systems 2004. URL: <https://www.sciencedirect.com/science/article/pii/S0921889004000715> (дата звернення: 01.03.2021).
18. Tomas Hodan, Martin Sundermeyer, Bertram Drost, Yann Labbe, Eric Brachmann, Frank Michel, Carsten Rother, Jiri Matas. BOP Challenge 2020 on 6D Object Localization. 2020. С.31-40.
19. Martin Fowler. UML Distilled: A Brief Guide to the Standard Object Modeling Languag. — 3rh Edition, 2004. 125-250 с.