**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ**

ДОПУСТИТИ ДО ЗАХИСТУ

Завідувач кафедри

Литвиненко О.Є.

« » 2020 р.

ДИПЛОМНИЙ ПРОЕКТ

(**ПОЯСНЮВАЛЬНА ЗАПИСКА**)

**ВИПУСКНИКА ОСВІТНЬОГО СТУПЕНЯ "БАКАЛАВР"**

**Тема:** Програмна система обліку даних інтернет-магазину

**Виконавець:**  Левчук В.В.

**Керівник:**  Глазок О.М.

**Нормоконтролер:** Тупота Є.В.

**Київ 2021**

**НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ**

Факультет Кібербезпеки, комп'ютерної та програмної інженерії

Кафедра комп’ютеризованих систем управління

Спеціальність 123 «Комп'ютерна інженерія»

(шифр, найменування)

Освітньо-професійна програма «Системне програмування»

Форма навчання денна

ЗАТВЕРДЖУЮ

Завідувач кафедри

Литвиненко О.Є.

«\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_2021 р.

**ЗАВДАННЯ**

**на виконання дипломної роботи (проєкту)**

                         Левчуку Владиславу Володимировичу

(прізвище, ім'я, по батькові випускника в родовому відмінку)

1. Тема дипломної роботи (проєкту) «Модуль розпізнавання жестів рук для

керування квадрокоптером»

затверджена наказом ректора від «  04  »  лютого  2021 р. № 135/ст.

2. Термін виконання роботи (проєкту): з 17 травня 2021 р. по 20 червня 2021 р.

3. Вихідні дані до роботи (проєкту): постановка задачі на виконання роботи,

мова програмування *C*++, середовище програмування *Visual Studio Code*, бібліотека *MediaPipe*, бібліотека *OpenCV*, одноплатний комп’ютер *Raspberry Pi*

4. Зміст пояснювальної записки:

1) Аналіз засобів, методів та апаратного забезпечення, які застосовують при створенні інтерактивних систем керування жестами;

2) Проектування модуля розпізнавання жестів рук для керування квадрокоптером;

3) Програмна та апаратна реалізація модуля розпізнавання жестів рук для керування квадрокоптером.

5. Перелік обов'язкового графічного (ілюстративного) матеріалу:

1) Програмне забезпечення (Діаграма прецедентів);

2) Програмне забезпечення (Діаграма послідрвності);

3) Програмне забезпечення (Діаграма класів);

4) Апаратне забезпечення (Схеми структурна);

5) Апаратно-програмний модуль (Ескіз поєднання модуля і коптера)

6. Календарний план

|  |  |  |  |
| --- | --- | --- | --- |
| №  пор. | Завдання | Термін  виконання | Відмітка про  виконання |
| 1. | Ознайомитися з літературою за темою дипломного проекту | 17.05.2021-19.05.2021 |  |
| 2. | Провести аналіз та огляд існуючих систем розпізнавання | 19.05.2021-21.05.2021 |  |
| 3. | Ознайомитися із перспективними апаратними платформами, які підходять для реалізації модуля. | 21.05.2021-22.05.2021 |  |
| 4. | Ознайомитися з існуючими бібліотеками для роботи із комп’ютерним зором. | 22.05.2021-24.05.2021 |  |
| 5. | Розробити план реалізації апаратної частини. | 24.05.2021-26.05.2021 |  |
| 6. | Реалізувати апаратну частину модуля і провести його відлагодження. | 26.05.2021-29.05.2021 |  |
| 7. | Розробити план реалізації програмної частини. | 29.05.2021-01.06.2021 |  |
| 8. | Реалізувати програмну частину модуля і провести його відлагодження. | 01.06.2021- 05.06.2021 |  |
| 9. | Написати пояснювальну записку | 05.06.2021-10.06.2021 |  |
| 10. | Підготувати графічний матеріал | 10.06.2021-12.06.2021 |  |
| 11. | Підготувати презентацію та доповідь | 12.06.2021-13.06.2021 |  |

7. Дата видачі завдання: «  17  »       травня        2021 р.

Керівник дипломної роботи (проєкту) Глазок О.М.

(підпис керівника) (П.І.Б.)

Завдання прийняв до виконання Левчук В.В.

(підпис випускника) (П.І.Б.)

**РЕФЕРАТ**

Пояснювальна записка до дипломної роботи “Програмний модуль для вивчення фахової дисципліни за дистанційними технологіями”: 88 с.*,* 31 рис.*, 1* табл*., 16* літературних джерел*,* 1 додаток.

ШТУЧНИЙ ІНТЕЛЕКТ, ГЛИБОКЕ НАВЧАННЯ, ОДНОПЛАТНА ЕОМ, *Raspberry Pi*, *MediaPipe*, *OpenCV*, ЖЕСТ, КВАДРОКОПТЕР, ПЕРЦЕПТРОН, АЛГОРИТМ, КОМП’ЮТЕРНИЙ ЗІР, СИСТЕМА РОЗПІЗНАВАННЯ ЖЕСТІВ.

**Об’єкт дослідження** – розпізнавання жестів рук на основі нейромережевим алгоритмам.

**Предмет дослідження** – модуль розпізнавання жестів рук для керування квадрокоптером.

**Мета дипломного проекту** – створення модуля на основі одноплатної ЕОМ для розпізнавання жестів рук з використанням алгоритмів штучного інтелекту з подальшою інтерпретацією результатів роботи модуля у команди для керування квадрокоптера.

**Метод проєктування** – застосування засобів розпізнавання з поєднанням алгоритмів глибокого навчання та використання апаратних можливостей одноплатних ЕОМ.

**Галузь застосування** – дистанційне жестикуляційне керування динамічних систем.

**Значущість роботи та висновки** – модуль жестикуляційного керування забезпечує просте та інтуїтивне керування рухомими системами, такими як квадрокоптер.

**Прогнози та припущення щодо розвитку об’єкта дослідження** – створення робочого зразка модуля, інтеграція зразка у квадрокоптерта використання його для керування. Подальший розвиток системи передбачає додавання нового функціоналу у модуль, зокрема: розпізнавання та ідентифікація обличчя власника або кола персон, розширення спектру жестів, які підлягають розпізнаванню, додовання можливості розпізнавання звукових команд з подальшим виконання завдань.

|  |  |
| --- | --- |
| ЗМІСТ  ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ, ТЕРМІНІВ……………....7  ВСТУП………………………………………………………………………..……..9  РОЗДІЛ 1 АНАЛІЗ ЗАСОБІВ, МЕТОДІВ ТА АПАРАТНОГО  ЗАБЕЗПЕЧЕННЯ, ЯКІ ЗАСТОСОВУЮТЬ ПРИ СТВОРЕННІ  ІНТЕРАКТИВНИХ СИСТЕМ КЕРУВАННЯ ЖЕСТАМИ…….…....10   * 1. Вступ до розділу……………………………………..…….…....10   2. Огляд загальної схеми роботи інтерактивних систем….…….10   розпізнавання жестів, теорія і приклади……….........………...10   * 1. Теорія розпізнавання образів…………………………...……...10   2. Концепція та теоретичне представлення нейронних мереж....10   3. Застосування одноплатних комп’ютерів у модулях   керування жестами…………………….………….…….……....10   * 1. Висновки до розділу……………….….………….…..………....10   РОЗДІЛ 2 ПРОЕКТУВАННЯ МОДУЛЯ РОЗПІЗНАВАННЯ ЖЕСТІВ РУК  ДЛЯ КЕРУВАННЯ КВАДРОКОПТЕРОМ…….................................10  2.1. Опис апаратних елементів, які застосовувалися під час  проектування...............………………….………….…………...10  2.1.1. Одноплатний комп’ютер…...….………….…………...10  2.1.2. Сервоприводи……………….….………….…………...10  2.1.3. Камера та альт-азимутальна установка………………10  2.1.4. Живлення…………………....….………….…………...10  2.2. Технології, застосовані для обробки зображень, реалізації  процеса виявлення жестів та керування альт-азимутальною  установкою…………………………..….………….…………...10  2.2.1. Бібліотека Mediapipe……….….………….…………...10  2.2.2. Бібліотека OpenCV…………….….……….…………..10 | |
|  | *5* |

|  |  |
| --- | --- |
| 2.2.2.1. Загальні відомості….….….…….…………..10  2.2.2.1. Розмивання Гауса….……..…….…………...10  2.2.2.1. Кольорова модель HSV….………….……...10  2.2.2.1. Порогові операції….………….………….....10  2.2.2.1. Алгоритм виявлення контурів….…..……...10  2.2.2.1. Алгоритм Рамера-Дугласа-Пекера………...10  2.2.2.1. Поєднання роботи алгоритмів…...………...10  2.2.3. Система збірки Bazel………….………….…………...10  2.2.4. Бібліотеки WiringPi та softPwm….…………………...10  2.3. Опис нейронної мережі – алгоритму зворотного  розповсюдження помилки……………………………………..10  2.4. Бібліотека JSON………………………………………………...10  2.5. Висновки до розділу……………………………….…………...10  РОЗДІЛ 3 ПРОГРАМНА ТА АПАРАТНА РЕАЛІЗАЦІЯ МОДУЛЯ  РОЗПІЗНАВАНЯН ЖЕСТІВ РУК ДЛЯ КЕРУВАННЯ  КВАДРОКОПТЕРОМ………………………………………………...10  3.1. Реалізація концепції модуля розпізнавання жестів …..……...10  3.2. Програмна реалізація модуля……………………..…………...10  3.2.1. Клас AltAzimuth…………………………….…………...10  3.2.2. Клас HandTracking………………………….…………...10  3.2.3. Клас MediaPipeHandsDetector.……………..…………...10  3.2.4. Класи NeuralNetwork та Perceptrons……….…………...10  3.3. Інструкція по експлуатації……………………………………...10  3.4. Висновки до розділу……………………………………………...10  ВИСНОВКИ…………………………………………………………..…………...10  СПИСОК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ ВИКОРИСТАНИХ ДЖЕРЕЛ...10  ДОДАТОК А………………………………………………………….…………...10 | |
|  | *5* |
| ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ*,* СКОРОЧЕНЬ*,* ТЕРМІНІВ  ЕОМ – Електронна обчислювальна машина  ШІ – Штучний інтелект  ПЗ – Програмне забезпечення  *IoT* – Internet of Things, Інтернет речі  ОПЗ – Оперативний запам’ятовуючий пристрій  ПЗЗ – Пристрій із зарядним зв’язком  *CCD* – Charge-coupled device  КМОН – Комплементарна структура метал-оксид-напівпровідник  *CMOS* – Complementary-symmetry/metal-oxide semiconductor  Мп – Мегапіксель  *GIF* – Graphics Interchange Format  *BMP* – Bitmap Picture  *JPEG* – Joint Photographic Experts Group  *PNG* – Portable Network Graphics  *RGB* – Red, Green, Blue  ЗПП – Зворотнє поширення помилки  ЦАП – Цифро-аналоговий перетворювач  *LCD* - Liquid crystal display  *HSV* - Hue, Saturation, Value  *GPIO* - General-purpose input/output  *USB* - Universal Serial Bus  *HDML* - Handheld Device Markup Language  *DSI* - Display Serial Interface | |
|  | *6* |

|  |  |  |  |
| --- | --- | --- | --- |
| *CSI* - Camera Serial Interface  *PWM* – Pulse-width modulation  ШІМ – Широтно-імпульсна модуляція  *JSON* - JavaScript Object Notation | | | |
|  | | *5* | |
| **ВСТУП**  На сьогоднішній день, дедалі обширним та тісним стає вплив електронних систем на повсякденне життя людей. Ріст попиту людей на задоволення своїх потреб як у побуті, так у вирішенні низки проблем економічного та логістичного характеру безпрецедентно спонукає до розвитку науково-технічної галузі інтерактивних систем людино-комп’ютерної взаємодії. Виникає дедалі більший запит на системи, які забезпечують керування людиною різноманітних комплексів, і вимагають дедалі більш простого інтерфейсу для взаємодії з цими комплексами. Можна виділити кілька причин для спрощення систем інтерфейсу керування людино-машинних комплексів, а саме: легкість в освоєні персоналом простого та зручного інтерфейсу, зменшення витрат на навчання персоналу для керування комплексом, інтуїтивно зрозумілий інтерфейс зменшує ризик виникнення помилок під час експлуатації. Таким чином, простий та інтуїтивно зрозумілий інтерфейс надає значно кращі можливості для керування, що у свою чергу, має ліпший вплив на економічні витрати, зокрема під час освоєння та експлуатації подібних систем.  Однією з таких систем інтерфейсу керування є керування за допомогою жестикуляції рук людини, що у свою чергу складаються з інтерактивної системи розпізнавання жестів. Інтерфейси керування, які полягають у керуванні за допомогою жестикуляції рук починають лише впроваджуватися у життєві процеси людини, але дедалі більше досліджень здійснюють у цій області. Зокрема, як зазначалося раніше, дана система дозволяє значно спростити керування комплексом, звівши це до інтуїтивно зрозумілого рівня. Також дані системи надають можливість виключити фізичний контакт оператора з комплексом, що досить актуально, зважаючи на епідеміологічний стан у світі.  Керування жестикуляційним способом значно краще відображається на здатності людини виконувати кілька операцій одночасно – здатність людини здійснювати розподіл уваги і здійснювати кілька видів діяльності. | | | |
|  | | *7* | |

|  |  |
| --- | --- |
| Ключовим аспектом у виконанні одночасно кількох операцій є увага. Увага – це така форма психологічної діяльності людини, що виявляється у спроможності свідомості зосереджуватися на певних об’єктах, які мають те чи інше значення для людини у певний момент часу. Також, одна з умов виконання такої «багатозадачності», полягає у тому, що одна з таких операцій буде виражена у вигляді міміки кінцівок рук. Звісно, людина здатна виконувати різноманітні дії, які людина може виконувати одночасно, але найбільш успішно вона виконує одночасно дві дії, одна з яких – фізична робота, наприклад моторика рук.  Зокрема можна навести приклад подібних ситуацій, коли людина успішно та ефективно виконує кілька дій одночасно. Наприклад вона здатна в один і той же час мити посуд і вести розмову із співрозмовником на кухні, спілкуватися по телефону па в’язати світер, керувати машиною та спілкуватися з іншою людиною по голосовому зв’язку через гарнітуру. Людина не використовує великих розумових зусиль задля виконання привичних та інтуїтивних задач, наприклад, розглядаючи у даному випадку миття посуду. Вона без перешкод здатна вести діалог на різноманітні теми, що одночасно не несе ніякого негативного впливу на виконання другої задачі (миття посуду).  Окрім спілкування, людина здатна і на інші інтелектуальні задачі, одночасно з виконанням задач із залученням моторики рук. Завдяки довгому шляху еволюції людина набула навичок ефективно здійснювати розподіл уваги між виконанням розумової роботи та фізичної праці рук і переключення уваги з однієї діяльності на іншу стає непомітним для людини. Хоча варто зауважити, що на багатозадачність людини впливає безліч чинників, зокрема фізичний і психологічний стан людини і навколишнє середовище: індивідум не здатний якісно виконувати навіть одну задачу, у разі, якщо є певні чинники, котрі будуть негативно впливати на зосередженість (концентрацію) – інженер не здатний вести розрахунки для проєктування каркасу будівлі, якщо на вулиці будуть здійснюватися гучні будівельні роботи.  Одним із ключових аспектів формування стійкої багатозадачності людини є стійкість уваги, що можна розтлумачити як здатність протягом тривалого часу | |
|  | *8* |

|  |  |
| --- | --- |
| утримуватися на якомусь об’єкті, дії. Розглядаючи дане визначення у контексті багатозадачності, можна сказати, що це здатність людини приділяти достатню кількість уваги протягом певного періоду часу для кожної роботи, яка виконується у даний момент, за який людина встигає виконати певний мінімальний обсяг робіт, при якому кожна з виконуваних дій людиною буде здійснюватися без проблем чи затримок, і не перешкоджатиме виконанню іншої дії, яку здійснює людина у цей час.  Одночасно, така успішна діяльність людини у плані багатозадачності втрачає ефективність, якщо дії, які виконує людина, будуть носити виключно характер розумової праці. Значно важкіше, а більшості випадках – неможливо, здійснювати одночасно кілька інтелектуальних задач – складно проводити відразу кілька обчислень, проводити постановку різних задач кільком групам людей і т.п.. В такому випадку відразу проявляється зниження ефективності виконання кожної з робіт.  Одним з можливих рішень у застосуванні інтерактивних систем керування жестами, на думку автора, є керування літальними апаратами – мультикоптерами. Натхненний швидким та буремним розвитком коптеробудування та появою нових алгоритмів, методик та інструментів у галузях комп’ютерного зору та глибокого навчання (deep learning), автор розробив власний модуль розпізнавання жестів рук для керування квадрокоптером. Даний вид керування дозволить розширити можливості по управлінню літальними апаратами даного типу, а не лише застосовувати загальноприйняті пульти керування на радіоуправлінні.  Таким чином, інтерактивна система розпізнавання жестів – це комплекс, який розкриває нові можливості для керування різноманітними системами. Системи такого типу інтуїтивно зрозумілі, завдяки тому, що керування здійснюється жестикуляцією рук оператора. Сам процес розпізнавання жестів рук ділиться на кілька етапів, протягом яких система отримує інформацію, аналізує її, обробляє і на основі неї здійснює висновок щодо жестикуляцій. Інтерактивні системи розпізнавання жестів є складним, але у той же час, надзвичайно вагомим знаряддям людино-машинного інтерфейсу. | |
|  | *9* |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **РОЗДІЛ 1**  **АНАЛІЗ ЗАСОБІВ, МЕТОДІВ ТА АПАРАТНОГО ЗАБЕЗПЕЧЕННЯ, ЯКІ ЗАСТОСОВУЮТЬ ПРИ СТВОРЕННІ ІНТЕРАКТИВНИХ СИСТЕМ КЕРУВАННЯ ЖЕСТАМИ**  **1.1. Вступ до розділу**  Буремний розвиток технологічного прогресу людства дедалі більш сильніше розпочинає впливати на повсякдення людей. Зокрема, завдяки революції в методологіях розробки систем штучного інтелекту та розвитку архітектури обчислювальних машин, стало можливо проєктування, розробка та впровадження різноманітних систем, які можна інтегрувати у побут. Дані системи представляють собою складний комплекс поєднань найрізноманітніших технологій та процесів, тісно інтегрованих та пов’язаних між собою, які призначені вирішувати задачі, донині вирішення яких було неможливе, не застосувавши окремо виділених ресурсів, або потребувало зусиль, невиправдних як з економічної точки зору, так з точки зору затраченого на це часу. Однією з «гілок» у системах даного типу є інтерактивні системи керування жестами, які можуть бути представлені як окремими модулями, вбудовані в більш складні програмно-апаратні комплекси, так і самостійними системами, на які покладено завдання розпізнання певних об’єктів, наприклад жестів людських рук.  Надалі буде більш детально описано, що собою представляють дані системи, які зразки даних систем існують та як саме працюють модулі керування жестами у таких системах. Також буде оглянуто теоретичне підґрунтя інтерактивних системи розпізнавання жестів, загальна схема та принцип роботи подібних структру. | | | | | | | | | |
| Кафедра КСУ | | | | НАУ 21 02 22 000 ПЗ | | | | | |
| ***Виконав*** | *Левчук В.В.* |  |  | Аналіз засобів, методів та апаратного забезпечення, які застосовують при створенні інтерактивних систем керування жестами | ***Літера*** | | | ***Лист*** | ***Листів*** |
| ***Керівник*** | *Глазок О.М.* |  |  | *Д* |  |  | *10* | *59* |
| ***Консульт.*** |  |  |  | СП-436 123 | | | | |
| ***Н. контроль*** | *Тупота Є.В.* |  |  |
| ***Зав. Каф.*** | *Литвиненко О.Є.* |  |  |

|  |  |
| --- | --- |
| **1.2. Огляд загальної схеми роботи інтерактивних систем розпізнавання жестів, теорія і приклади**  «Що таке інтерактивна система керування жестами?» - задасть доволі слушне питання пересічна людина. Інша людина могла б припустити і дати відповідь на це питання наступним чином: «Це система, яка здатна розрізняти певні жести оператора, і на основі них робити ті чи інші дії». І насправді це буде цілком вичерпна відповідь на дану квестію. Якщо більш детально розглянути дане питання, то під визначенням «інтерактивна система керування жестами» розуміється система, яка здатна виявляти та розрізняти серед стохастичного набору рухів людини набір конкретних жестів, проводити аналіз даних жестів та інтерпретувати їх у вибірку команд або іншого вигляду інформацію, яка може використовуватися я к самою системою так і іншими системами чи комплексами.  Розглянемо загальну структуру таких систем і її основні частини (рис. 1.2.1.)    Рис. 1.2.1. - Загальна схема інтерактивної системи керування жестами  У даній системі наявні наступні основні вузли:   * Камера; * Буфер кадру; | |
|  | *11* |

|  |  |
| --- | --- |
| * Система виявлення жестів; * Система розпізнавання жестів; * Інтерфейс виведення інформації; * Відеоадаптер та Графічний дисплей; * Елемент управління камерою.   Також присутній необов'язковий модуль «Елемент управління камерою».  Розглянемо більш детально кожен із наведених пунктів, розберемо роль даних елементів у цілісній системі і взаємодію між даними блоками.  **Камера.** Камера є основним елементом, який забезпечує систему інформацією у вигляді кадрів відеопотоку або окремих зображень. Кадри зображення є основною інформацією, якою оперує система.  Необхідно описати основні характеристики, за якими обирають камери, для систем керування подібного типу. Серед основних характеристик камер, можна виділити: тип матриці, розмір матриці, фокусна відстань та кут огляду, роздільна здатність, додаткові можливості.  Більшість камер володіють одним із двох типів матриць: ПЗЗ-матрицю (англ. CCD) або КМОН-матрицю (англ. CMOS). Основна відмінність - спосіб зчитування інформації. У ПЗЗ-матриць зчитування відбувається послідовно, у КМОН-матриць зчитування відбувається у довільний спосіб. Переваги ПЗЗ полягають у більш якіснішому зображені, ніж у КМОН, проте у останніх формування кадру зображення відбувається значно швидше.  Фізичний розмір матриці визначає якість як самої матриці, так якість формування зображення. Проте, якість зображення залежить саме від розміру окремого елемента матриці – комірки. Збільшення розміру даної комірки призводить до покращення характеристик якості кадру відео чи зображення.  Фокусна відстань та кут огляду – два взаємопов’язані параметри, які визначають дальність та кут огляду камери. Зі зменшенням фокусної відстані зростає кут огляду, і зменшується кількість деталей на кадрі, які є можливість оглянути, і навпаки.  Параметр роздільної здатності. Роздільна здатність вимірюється у | |
|  | *12* |

|  |  |
| --- | --- |
| мегапікселях (Мп). Один мегапіксель уміщає у собі один мільйон пікселів. Кількість мегапікселів визначає розмір і якість зображення.  Камери можуть володіти додатковими можливостями, які збільшують спектр застосувань камери або змінюють параметри камери задля збільшення якості змімкі. Наприклад, можливість монтувати додаткові об’єктиви, містити інфрачервоний фільтр, що дозволить проводити зйомку у нічний час доби, мати додаткову підсвідку або спалах.  ***Буфер кадру***. Буфер кадру слугує місцем, куди поміщається зображення, отримане у ході зйомки. Зображення з камери може зберігатися у різноманітних форматах, які вирізняються якістю, у якій зберігають зображення.  Чим більш якісне зображення, тим системі керування жестами знадобиться здійснити більше роботи над зображенням, протягом всього робочого циклу: формування кадру, попередня обробка кадру, опрацювання кадру системою виявлення жестів. Це у свою чергу може сильно відобразитися на продуктивності окремих блоків, що впливатиме на всю систему в цілому. Таким чином, від формату зображення напряму може залежати робота всієї системи і необхідно підібрати формат, що забезпечить оптимальний рівень роботи системи.  Наразі, найбільш популярними форматами зображень є: GIF, BMP, JPEG, PNG.  Формат GIF. Вимовляється як «гіф». Має можливість зберігати дані з відсутністю втрати якості. Він не залежить від апаратних характеристик ЕОМ, що дозволяє використовувати його на різних платформах. Проте є лише підтримка 256 кольорів, але це забезпечує невелику вагу зображення. Також формат може зберігати кілька кадрів в одному файлі і час, протягом якого має бути відображений кожен кадр.  Формат BMP. Вимовляється як «бі-ем-пі». Розроблявся для операційних систем Windows та OS/2 і може зберігати лише один кадр, проте на кожен піксель в різних файлах може припадати різна кількість біт, що формує глубину кольору. Наразі доступні бітності 1, 2, 4, 8, 16, 32, 48, 64.  Формат JPEG. Вимовляється як «джейпег». Формат пропонує «золоту середину» між якістю та розміром картинки. Дозволяє проводити зжимання як з | |
|  | *13* |

|  |  |
| --- | --- |
| втратами так і без. У режимі стиснення є ймовірність появи різноманітних шумів що може зашкодити здійснювати аналіз зображень.  Формат PNG. Застосовують для зберігання зображень, у яких не має відбуватися втрат при зжимані. Підтримує різну глибину зображень, 16, 24 та 48 біт. Колір подається у форматі RGB, включаючи підтримку альфа-каналу.  Можна зробити висновок, що кожен із форматів надає користувачу свої переваги, але одночасно має і певні недоліки. Вибираючи формат, необхідно брати до уваги можливості інтерактивної системи керування жестами.  ***Система виявлення жестів.*** Отримавши інформацію з камери і провівши збереження кадру в одному із форматів, отримали зображення, яке слугуватиме одиницею інформації у інтерактивній системі розпізнавання жестів.  Дане зображення необхідно попередньо обробити спеціальними засобами.  Варто розглянути, що собою представляють жести. Жест – це така дія або рух людини, що забезпечує передачу інформації чи емоції без застосування мовлення. Жестом можна назвати будь-яку послідовність рухів людини, які можуть бути передати певну інформацію, або нести певний сенс. Людина використовує різноманітні жести, такі як: маніпулятивні, комунікаційні, емоційні, ілюстративні жести.  Тепер варто зрозуміти, чим являється жест, у системі розпізнавання і яких форм він набуває.  Жест є одиницею інформації. Як було сказано раніше, перед початком розпізнавання та класифікації жеста, необхідно певним чином обробити кадр зображення, щоб мати змогу виявити сам жест. Лише після того, як кадр був оброблений, можна отримати кадр із розпізнаним силуетом руки і розглядати такий кадр як одиницю інформації, яку надалі можна використовувати як для роботи розпізнавання і класифікації жестів, так і для безпосереднього розпізнавання логіки жеста. Засоби і методи, які використовують для виявлення жеста на кадрі зображення, буде детально висвітлено у пункті 1.3.  У більш ранніх технологіях по розпізнаванню і керуванню, використовувалися спеціальні маркери на кінчиках пальцях (рис. 1.2.2.) | |
|  | *14* |

|  |  |
| --- | --- |
| Рис. 1.2.2. – Рука із кольоровими маркерами  що дозволяло відслідковувати ключові точки на руці. Ці маркери представляли собою певні невеликі елементи визначеної кольорової гамми. Система аналізувала зображення на наявність цих «кольорових точок», і у разі, якщо колір даних точок входить у визначений діапазон кольорів, точка ідентифікувалася як ключовою. На кадрі зображення визначалася вибірка таких точок, які формували двомірну матрицю, готовою до обробки засобами розпізнання жестів. Система аналізувала матрицю і у разі успіху ідентифікувала ці точки як жест. Даний підхід ускладнював експлуатацію системи користувачем через ймовірність попадання шумів, які можуть ідентифікуватися як ключові точки пальців.  На сьогоднішній день, методологія розпізнавання жестів лежить на основі комп’ютерного зору. Модель може набувати вигляду однотонних силуетів. Проте, сам силует не використовується безпосередньо при розпізнаванні образа руки через складність і велику витрату ресурсів, при обробці такого кадру і ідентифікації на кадрі жесту руки. Зазвичай силует обробляють спецзасобами, які забезпечують отримання контурів руки.  Такий спосіб представлення інформації більш кращий, у порівнянні з усім силуетом руки, через те, що системі необхідно опрацьовувати значно меншу кількіть елементів (тобто - контури). Хоча, для повноцінної роботи системи розпізнавання, цього може бути недостатньо. | |
|  | *15* |

|  |  |
| --- | --- |
| Також наявний спосіб розпізнавання жестів на основі скелетної моделі кисті руки (рис. 1.2.3.).    Рис. 1.2.3. – Скелетна модель кисті руки  Він представляє симбіоз попередніх методик отримання жесту руки, з виділенням ключових точок (точок ламаної) у контурах та обчислення центру мас та побудови скелету по цим даним.  Після обробки та виявлення контурів жеста, зображення передається безпосередньо у систему виявлення жестів, яка може бути представленням наборів алгоритмів, зокрема і алгоритмів штучного інтелекту. На виході отримуємо результат у вигляді матриці *n*-мірності.  ***Система розпізнавання жестів.*** Розпізнавання – це процес, що здійснює групування та оброблення просторово-часової інформації, отриманої на попередніх етапах і здійснення остаточного висновку щодо чи є це жестом, і якщо так, то до якого набору жестів її віднести. Для цього використовуюють алгоритми штучного інтелекту. Перед початком роботи такої системи, необхідно провести її навчання. Під навчанням розуміють процес отримання такої системи, яка здатна реагувати на дані і приходити висновку. Для розпізнавання переважно застосовують алгоритм зворотнім поширенням помилки (ЗПП).  На кінець, нейромережа надає відповідь, що саме за жест було виявленно. Після цього необхідно провести класифікацію жеста, щоб встановити, яку саме | |
|  | *16* |

|  |  |
| --- | --- |
| інформацію необхідно передати на вихід через інтерфейс виведення інформації. Також, на основі отриманого результату, система може передавати певну графічну інформацію, яка може відображатися на графічному дисплеї, за необхідністюі.  ***Інтерфейс виведення інформації.*** Фінальним елементом системи по керування жестами є інтерфейс, через який система здійснює надсилання інформації отримувачу. Отримувачем може слугувати будь-яка система, який потребує даних від даної інтерактивної системи. На етапі розробки необхідно визначити, який тип даних має надавати мережа - чи це буде «класичний» тип даних, наприклад цілочисельне значення, певний символ або значення з плаваючою комою, або користувацький тип даних, такий як структура або клас. Не обов’язково система має повертати цифровий тип даних, це також може бути аналоговий сигнал, як наприклад радіочастота.  ***Відеоадаптер та Графічний дисплей.*** Даний елемент не є обов’язковим у системі. Зазвичай застосовується у разі відладки, або відображати зображення відеокамери, на яке накладено певні спецефекти.  Вивід графічної інформації розподіляється на два етапи – передача інформації від системи керування жестами у відеоадаптер і передача зображення від відеоадаптера на графічний дисплей. Під відеоадаптером розумієтьсяпрограмний блок, який забезпечує отримання інформації від системи розпізнавання жестів, її аналіз, обробка початкового кадру, отриманого від камери, і накладення на нього графічних елементів та передача зображення на вивід графічному дисплею.  ***Елемент управління камерою.*** Ще один додатковий модуль, який не є обов’язковим. Залежно від реалізації системи, камера може бути як статично закріпленою, так такою, яка має здатність обертатися. Якщо камера має бути рухомою, необхідно передбачити елемент, що надасть таку можливість. Для можна використати альт-азимутальну установка (рис. 1.2.4.). | |
|  | *17* |

|  |  |
| --- | --- |
| Рис. 1.2.4. – Альт-азимутальні установки із закріпленими камерами  Альт-азимутальна установка – це такий пристрій, що має вертикальну та горизонтальну осі обертання. Обертання може здійснюватися як вручну так і за допомогою спецзасобів, наприклад сервоприводів.  Розглянемо приклади використання інтерактивних систем розпізнавання жестів у різноманітних проєктах.  Наразі поширене явище створення пристроїв, відомих як інтернет речі (англ. Internet of Things, IoT). Це система, у якій різноманітні об’єкти здатні бути пов’язаними між собою, мають доступ до виходу у мережу Інтернет (або в локальну мережу), володіють різноманітними датчиками та сенсорами. Спільна робота допомагає у виконанні різноманітного кола задач, з якими людина стискається протягом всього циклу дня. Приклад IoT – система «розумний дім». Система «Розумний дім» - це система, яка складається з певного набору пристроїв, які розташовуються у будинку, які призначені виконувати набір дій з списку повсякденних завдань, без участі людини. «Розумний дім» може мати набір різноманітних датчиків: датчики освітлення, руху, вологості і т.п.. Одним із таких датчиків може бути інтерактивний модуль керування жестами. Система, буде захоплювати та відстежувати рухи рук господаря та розпізнавати набір жестів, що дозволить керувати технікою у домі, наприклад дял збільшення гучності на медіаплеєрі, керування освітленням або увімкненням побутових пристроїв лише одним рухом руки. | |
|  | *18* |

|  |  |
| --- | --- |
| Варто згадати системи, що впроваджені у мобільних телефонах та мобільних додатках. Зокрема, у них здійснюється відстеження рук, розпізнавання та класифікація жестів рук, що надає змогу: здійснення знімку екрана, початок та закінчення відеозапису екрана телефона, розблокування пристрою і т.п..  Також прикладом може слугувати система керування жестами у бортових комп’ютерах сучасних автомобілів. За функціоналом подібні до мобільних систем, проте у автомобілі система забезпечує керування мультимедією автомобіля. Прості жести, такі як "вказати" або "змахнути" надають можливість викликати певні функції - наприклад, прийняти або відхилити дзвінок.  **1.3. Теорія розпізнавання образів**  Виявлення жестів є окремим складним завданням, яке передбачає застосування цілого набору різноманітних операцій та алгоритмів.  Що собою представляє теорія розпізнавання образів? Це набір теоретичних основ та методик ідентифікації та класифікації предметів, процесів, явищ, за якими закріплена певна вибірка певних специфічних ознак, які допомагають точно ідентифікувати сутність із сукупності інших сутностей. Варто дати визначення, що саме розуміється під терміном «об’єкт». У теорії розпізнавання образів, об’єктом можна назвати виокремлену частину Всесвіту, що має певну форму та вид, за допомогою чого можна виділити набір ознак, притаманних саме цьому об’єкту. Відібравши вибірку ознак, необхідно вказати процедуру, яка описуватиме змінну кожної ознак. Наприклад, переважна більшість рослин мають зелене забарвлення і це є ключовою ознакою рослин. Також ознакою рослин є наявність листків. Проте, при настанні осені, колір листків змінюється від зеленого до помаранчевого, червоного, а через деякий час рослини взагалі позбуваються листків. Проте це не заважає нам ідентифікувати дані об’єкти як рослини. Отже колір та листки – це ознаки, а змінна кольору та наявність листків є варіацією цих ознак. Таким чином, кожний об’єкт – це виокремлена сутність серед інших сутностей із набором ознак, які однозначного ідентифікують об’єкт, а для кожної ознаки відоме правило, за яким ознака може варіюватися. | |
|  | *19* |

|  |  |
| --- | --- |
| На сьогоднішній день, задача розпізнавання є складним завданням, як у теоретичному розумінні, так і у практичному втілені. При формуванні теоретичної бази, розглядають поєднання спостережень можливостей живих істот та розвиток і дослідження математичної моделі, яка описує роботу розпізнавання сутностей. Наприклад, камера, яка слугує ключовим елементом у системах розпізнавання, є прототипом людського ока.  Описати систему виявлення образів (жестів) можна кількома пунктами. Перш за все, необхідно вказати, який саме об’єкт буде підпадати під розпізнавання, і виділити основні ознаки в цих об’єктів. Надалі будується математична модель, яка описуватиме той чи інший об’єкт. Вибирається необхідний алгоритм, або набір алгоритмів, які здійснюватимуть виявлення об’єктів. Здійснюється перетворення початкових даних для успішного опрацювання цих даних вибраним алгоритмом.  Розглядаючи розпізнавання об’єктів у контексті розпізнавання жестів руки, очевидно, що об’єктом розпізнавання є кисть руки. Ознаки руки – наявність пальців, спецефічна форма долоні (силует, що нагадує коло, овал) та заданий спектр кольорів людської шкіри. За даними ознаками, людина здатна без проблем ідентифікувати кисть руки серед набору інших об’єкті.  Тепер необхідно провести побудову математичної моделі кисті руки. Математична модель – це певні співвідношення, задані через математичні правила, які описують досліджуваний об’єкт. Математичну модель кисті руки можна подати у вигляді структури, яка складається з набору менших підструктур – точок, які описуються кількома параметрами, що описують просторове розміщення цих точок.  Наступним етапом в побудові системи виявлення жестів, є побудова математичного алгоритму, який виявлятиме жест. Для вирішення даної задачі наявний вельми великий набір математичних структур, які здатні вирішити дану проблему. Вони являють одні з найбільш ранніх варіантів алгоритмів, запропонованих вирішити задачу розпізнавання, проте зараз вони слугують основою для інших, більш продвинутих алгоритмів, або застосовуються разом з іншими алгоритмами, зокрема з моделями штучного інтелекту. | |
|  | *20* |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Виявлення об’єкта зазвичай є процесом «порівняння» по певному зразку із частинами зображення на якому необхідно знайти подібний об’єкт. Об’єкт має характеристики, які його однозначно ідентифікують: колір, форма, розташування у просторі, певні спеціальні характеристики (наприклад, пальці на долоні однозначно ідентифікують об’єкт як кисть руки) і т.п.. Об’єкти підпадають під класифікацію. Розглядаючи відеопотік, ставиться задача виділити об’єкти певного класу. Можна проводити класифікацію по формі. Класифікацію по формі можна розділити на два напрямки: по заданому шаблону та по аналітичній заданій формулі  Пошук по заданому шаблону. Шаблон має набувати форми, подібної до форми об’єкта, який має знаходитися на зображені. Такий шаблон, із необхідно заданою формою, переміщається по зображенню, одночасно проводиться розрахунок положення шаблону на кадрі по певній заданій формулі, і у разі, якщо значення результату є вищим якогось шуканий об’єкт  Пошук по аналітичній заданій формулі полягає у застосуванні певної математичної формули, яка вміщає у собі опрацювання певної кількості точок на зображенні. Після опрацювання даного набору точок, фігура, задана математичної формулою, просто здвигається по координатним осям і обрахунок проводиться заново. Під час опрацювань точок функції, відбувається розгляд характеристики кожної точки, наприклад переглядається інтенсивність яскравості точок.  Можна проводити класифікацію по просторовому розміщенню. При здійсненні пошуку об’єкта в області кадр, здійснюється припущення, що шукані області представляють собою довгасті, витягнуті у певному напрямку набір пікселів, які мають схожу яскравість. Наприклад, об’єкт може містити певні мітки, які вирізнятимуть об’єкт, або якісь інші деталі, виявивши ці деталі, можна легко здійснити припущення розрашуванню інших елементів, відносно даного елемента. Тобто, якщо існує метод, який забезпечує виділення в кадрі, наприклад, очей людини або носа, то, використовуючи ці деталі, можна зробити припущення, де знаходяться інші елементи об’єкта | | | | | | | | | | | |
|  | | | | | | | | | | | *20* |
| Класифікація по кольору є одним із основоположних методів по виявленню об’єктів. Зазвичай використовується разом з іншими методами для отримання максимального ефекту. Об’єкти підлягають ідентифікації через аналіз їх кольору – вони мають певну кольорову гаму. Завдяки різним способам представлення кольорів у зображені є можливість проводити глибокий і точний аналіз зображення. Для роботи подібної системи необхідно вибрати кольорову модель і задати необхідні діапазони для координат кольора, який шукається.  Вищеперелічені способи, використовуються як один із основних елементів при виявленні об’єктів. Зазвичай їх поєднують між собою і застосовують разом із іншими математичними функціями по обробці зображень,  Всі ці перетворення здійснюють лише частину роботи по виявленню об’єктів. Надалі, необхідно створити спеціальну математичну функцію, яка зважаючи на всі ці критерії (як наприклад, розмір, форма, колір, розміщення об’єкта), могла проводити автоматичне розпізнавання об’єкту. Для цього використовують алгоритми штучного інтелекту, зокрема нейронні мережі.  На завершення, виявлення об’єктів це складна процедура, яка відбувається в кілька етапів, а саме: вибору моделі об’єкта, який відстежується, побудова математичної моделі цього об’єкта, вибір алгоритмів по реалізації виявлення, здійснення підготовки вхідних даних.  **1.4 Концепція та теоретичне представлення нейронних мереж**  Як варто трактувати поняття «Штучний інтелект»? Даний термін є обширним і має кілька трактувань. Одне з них, полягає в тому, що штучним інтелектом можна назвати таку систему, яка має здатність виконувати творчі функції, які вважаються притаманні лише людині. Наприклад, розпізнавання зорових образів. Для людине не є вельми важко розпізнати той чи інший об’єкт, навіть за умови, що раніше людина не стискалася з ним. Кожен з нас, при виявленні незнайомої квітки, може точно сказати, що даний об’єкт – це рослина, і відноситься до квіток. Завдяки тому факту, що людина має можливість однозначно ідентифікувати ключові елементи рослини, такі як стебло, листки, | | | | | | | | | | | |
|  | | | | | | | | | | | *20* |
| бутон, вона може встановити, що даний об’єкт – квітка. На жаль, для обчислювальних систем це є важкою задачею і донині. Для вирішення задач подібного типу використовують нейронні мережі.  Штучні нейронні мережі – це такі системи, які є прототипами справжніх біологічних нейромереж. Батьком нейронних мереж по праву вважається Френк Розенблат, який описав у своїй книзі «Принципи нейродинаміки: перцептрони і теорія механізма мозку» загальний принцип роботи мозку людини, увів і описав прототип біологічного нейрону – перцептрон, і надав розгорнуте поняття роботи набору таких перцептронів – нейромережі. Його праця надовго заклала фундамент у принципах формування штучного інтелекту, як такого. Надалі буде оглянуто один із основоположних алгоритмів нейронних мереж – алгоритму зворотного поширення помилки (ЗПП). Детальний опис роботи цього алгоритму буде наведено у Розділі 2.  Алгоритм зворотнього поширення помилки – алгоритм реалізації штучної нейронної мережі на основі перцептронів за Розенбладом. Нейронна мережа за Розенблатом, складається із певної вибірки перцептронів, які організовані у структури – шари. нейромережа, розроблена за алгоритмом ЗПП, складається з кількох різних шарів персептронів, названих S, A, та R елементи (або вхідний, прихований та вихідний шари відповідно). Кількість прихованих шарів і перцептронів, у них, може бути довільною. Між кожним перцептроном у шарі є зв'язок, який сполучає його з перцептронами наступного шару. Через дані зв’язки відбувається обмін інформацією. Самі вузли мережі, персептрони, зберігають значення обчислень з попередніх шарів.  Варто розглягути, чим взагалі є перцептрон. Перцептрон має кілька вхідних зв’язків та один вихідний, хоча кількість зв’язків може бути різною. Перцептрон є прототипом реального біологічного нейрона, який працює завдяки збудженню зв’язків – дендритів, і нейрон, в залежності, наскільки активні дендрити, збуджує або не збуджує вихідний відросток – аксон. Робота перцептрона є аналогічною. Перцептрон збирає поточну інформацію із вхідних зв’язків, пропускає її через певну порогову функцію і передає оброблену інформацію на вихідний зв’язок (аксон). Таким чином, інформація проходить по всій нейромережі – від вхідного шару, проходить через приховані шари і оброблена | | | | | | | | | | | |
|  | | | | | | | | | | | *20* |
| інформація виводиться на вихідний шар. Після цього, отриманий результат аналізується, і у разі помилки – значення ваг персептронів корегується. Проводячи велику кількість ітерацій цих дій, нейромережа проходить етап навчання, що надалі дозволяє їй виконувати завдання із великою точністю.  Підбиваючи підсумок, нейронні мережі є прототипами справжніх біологічних нервових систем. Основуючись на перцептронах – прототипах нейронів, дані мережі здатні проводити вирішення складних задач, зокрема, розпізнавання образів та їх класифікація.  **1.5. Застосування одноплатних комп’ютерів у модулях керування**  **жестами.**  При створенні інтерактивних систем керування жестами, необхідно вирішити питання апаратної частини, тобто на якій електронно-обчислювальній машині буде виконуватися програмний код.  Найбільшого поширення набуло використання одноплатних ЕОМ через задовільні характеристики та компактність. Одноплатний ЕОМ – це комплекс технічних засобів, елементи, які призначенні для опрацювання інформації в процесі вирішення обчислювальних задач, розміщенні на одній платі, що дозволяють реалізувати різноманітні компактні та автономні системи керування жестами. У таблиці 1.5.1 наведена характеристика деяких ЕОМ.  Таблиця 1.5.1  Порівняльна таблиця одноплатних ЕОМ   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | ЕОМ | Ціна | Процесор | Ядра | ОЗП | Пам’ять | | Raspberry Pi 3B | 35 $ | Broadcom BCM2837 | 4x A53, 1.2 Ghz | 1 Gb | MicroSD | | Banana Pi M3 | 74 $ | Allwinner A83T | 8x A7, 2GHz | 2 Gb | 8GB eMMC, MicroSD | | Orange Pi Prime | 30 $ | Allwinner H5 | 4x A53 | 2 Gb | MicroSD | | | | | | | | | | | | |
|  | | | | | | | | | | | *20* |
| Дані комп’ютери підтримують операційні UNIX-системи, а наявність різноманітних портів, надасть можливість підключати різноманітну переферію і модулі, що забезпечує певну «маневреність» при реалізації власних проєктів.  **1.6. Висновки до розділу**  Під час написання першого розділу, було наведено пояснення, що саме розуміється під терміном «інтерактивна система керування жестами». Описано загальну структуру такої системи, наведені основні і додаткові функціональні блоки. Кожен блок був детально описаний, наведено апаратні засоби, надані алгоритми роботи та методики, що використовуються при проєктуванні кожного з блоків. Також наведені відомості з теорії розпізнавання образів, наведені відомості про загальний принцип виявлення об’єктів та поясненно, якими інструментами та методами досягається результат такої роботи. Викладені загальні положення про нейронні мережі, наведені поширені алгоритми, які використовуються у системах керування жестами, які також застосовуються під час виявлення об’єктів. На кінець був проведений огляд апаратури, на якій, зазвичай, працюють системи розпізнавання жестів, а саме одноплатні комп’ютери. | | | | | | | | | | | |
|  | | | | | | | | | | | *20* |
| **РОЗДІЛ 2**  **ПРОЕКТУВАННЯ МОДУЛЯ РОЗПІЗНАВАННЯ ЖЕСТІВ РУК ДЛЯ КЕРУВАННЯ КВАДРОКОПТЕРОМ**  **2.1. Опис апаратних елементів, які застосовувалися під час проектування   модуля**  **2.1.1. Одноплатний комп’ютер.**  Основою для виконання програмного коду та керування елементами модуля, був обраний одноплатна ЕОМ Raspberry Pi 3B+ (рис. 2.1.1.1.).    Рис. 2.1.1.1. – Raspberry Pi 3B+  Даний одноплатний комп’ютер відповідає усім необхідним вимогам, які потрібні для успішного проектування та створення автономного модуля, а саме: Компактність, задовільні апаратні можливості, наявність інтерфейсу вводу-виводу загального призначення (GPIO), мале споживання електроенергії, архітектура комп’ютера підтримує роботу UNIX-подібних операційних систем. | | | | | | | | | | |
| Кафедра КСУ | | | | НАУ 21 02 22 000 ПЗ | | | | | | |
| ***Виконав*** | *Левчук В.В.* |  |  | Проектування модуля розпізнавання жестів рук для керування квадрокоптером | ***Літера*** | | | ***Лист*** | ***Листів*** | |
| ***Керівник*** | *Глазок О.М.* |  |  | *Д* |  |  | *21* | *59* | |
| ***Консульт.*** |  |  |  | СП-436 123 | | | | | |
| ***Н. контроль*** | *Тупота Є.В.* |  |  |
| ***Зав. Каф.*** | *Литвиненко О.Є.* |  |  |

|  |  |
| --- | --- |
| Компактність. Розміри даного одноплатного комп’ютера наступні: довжина 85 мм, ширина 56 мм, висота 17 мм.  Задовільні апаратні можливості. Апаратні характеристики даного пристрою задовольняють потреби, які виникають при роботі програмного забезпечення.  Характеристики Raspberry Pi 3B+:   * Чіп: Broadcom BCM2837B0; * Процесор: 64-бітний 4-ядерний ARMv8 Cortex-A53 з тактовою частотою 1,4 ГГц; * Відеопроцесор: VideoCore IV 3D; * ОПЗ: 1 Гб LPDDR2 (900 МГц); * Роз’єми: HDMI, 3,5 мм jack, USB 2.0 4 шт., роз’єм для microSD, 10/100 Ethernet, порт для дисплея DSI, потр для камери CSI-2, GPIO 40 pins; * Безпроводні інтерфейси: WiFi 2,4 ГГц і 5 ГГц IEEE 802.11.b/g/n/ac 2,4 ГГц, Bluetooth: 4.2 Classic.   Вважливим моментом є наявність інтерфейс вводу-вивода загального призначення (GPIO). Даний інтерфейс забезпечую можливість підключення до ЕОМ різноманітних модулей і їх керування, також надає можливість через піни здійснювати надсилання сигналів.  Даний комп’ютер здатний працювати при малому споживанні електроенергії – достатньо напруги в 5V та сили струму в діапазоні 1.5-2 A.  Операційна система, вибрана для роботи, є Raspbian OS – операційна система, заснована на основі Debian.  **2.1.2. Сервоприводи.**  Сервопривід – це пристрій, що за рахунок енергії допоміжного джерела здійснює механічне переміщення регулюючого органу відповідно до отримуваних від системи керування сигналів. Для забезпечення обертання рухомих частин модуля (альт-азимутальна установка), було вибрано сервоприводи Tower Pro MG90S (рис. 2.1.2.1.). | |
|  | *22* |

|  |  |
| --- | --- |
| Рис. 2.1.2.1. – Сервопривід  Характеристики:   * Робоча напруга: 4-6 V; * Пусковий момент: 1.8 кг\*см; * Кут повороту: 180˚; * Час повороту на 60: (при напрузі 4.8V без навантажень): 0.1 с, (при напрузі 6V без навантажень): 0.08 с; * Розміри: довжина: 32,3 мм , ширина: 33 мм, висота: 12,4 мм.     **2.1.3. Камера та альт-азимутальна установка.**  Одним із основних елементів модуля є камера, яка здійснює відеозахоплення. Вибір впав на камеру для Raspberry Pi, версії 1.3 (рис. 2.1.3.1.).    Рис. 2.1.3.1. – Модуль камери для Raspberry Pi, версія 1.3 | |
|  | *23* |

|  |  |
| --- | --- |
| Модуль камери має роздільну здатність матриці в 5 Мп, розмір матриці сягає 1/4" або 3,76 × 2,74 мм, максимальний розмір зображення, який може бути створений - 2592 × 1944 пікселів. Розміри модуля: довжина: 25 мм , ширина: 24 мм, висота: 9 мм.  Задля забезпечення фіксації і обертання камери, було використано альт-азимутальну установку (рис. 2.1.3.2.).    Рис. 2.1.3.2. – Альт-азимутальна установка  **2.1.4. Живлення.**  Живлення модуля відбувається за допомогою двох джерел. Для забезпечення роботи одноплатного комп’ютера *Raspberry Pi* та сервоприводів, використовується два зовнішніх акумулятора (повербанк) *Power Bank 3 10000 mAh* (рис. 2.1.4.1.).    Рис. 2.1.4.1. – Зовнішній акумулятор  Вихідна напруга: 5V, сила струму: 1.5 A, ємність батареї: 10000 mAh. | |
|  | *24* |

|  |  |
| --- | --- |
| **2.2. Технології, застосовані для обробки зображень, реалізації процеса виявлення жестів та керування альт-азимутальною установкою.**  **2.2.1. Бібліотека Mediapipe.**  Mediapipe – це кросплатформенна бібліотека рішень задач комп’ютерного зору, на основі використання бібліотеки OpenCV та алгоритмів ШІ.  OpenCV – бібліотека алгоритмів комп’ютерного зору та алгоритмів для опрацювання зображень та відеопотоків, яка написана на мові C/C++. Алгоритми штучного інтелекту, які використовуються у MediaPipe, забезпечуються завдяки бібліотеці для машиного навчання TensorFlow. Алгоритми TensorFlow містять інструменти для створення моделей нейроних мереж і їх навчання.  Робота даної бібліотеки організована у вигляді «конвеєра» або «трубопровода» (pipeline) – у систему подаються певні вхідні дані, які проходять через кілька вузлів обробки і користувач отримує на виході результат, залежно від заданих конфігурацій. Таким чином, робочою системою бібліотеки є пайплайн («трубопровід»), який зображується у вигляді графа. Сам граф складається вершин графа, їх також називають «калькуляторами», ребер графа. Ребрами графа є потоки, які передають інформацію, пакетів. Пакет – це одиниця даних, яка курсує від одного калькулятора до іншого, завдяки потокам, які пов’язують калькулятори.  Також, в роботі графа застосована синхронізація між пакетами, задля забезпечення коректної і стабільної роботи системи.  Розберемо детальніше концепцію даної бібліотеки. Бібліотека є набором графів, які складені з калькуляторів, тому, для успішного використання бібліотеки, необхідно розуміти принцип роботи вершин графа, а саме, як все влаштовано «під капотом».  Кожен з калькуляторів є вузлом графа. Для того, щоб використати калькулятор у графі, необхідно провести певну послідовність дій. Спочатку | |
|  | *25* |

|  |  |
| --- | --- |
| визначити новий підклас – калькулятор, який бажаємо створити. Надалі реалізувати обов’язкові методи класа. Визначити вхідні та вихідні потоки. На кінець, провести реєстацію нового калькулятора.  Розглянемо більш детальніше кожен із кроків. Для того, щоб створити калькулятор, необхідно створити клас, дочірній класу CalculatorBase. Даний клас є батьківським для усіх класів, що реалізовують роботу калькуляторів.  Надалі необхідно провести перевизначення,, як мінімум, чотирьох методів: GetContract(), Open(), Process(), Close(). Робота даної бібліотеки організована у вигляді «конвеєра» (pipeline) – у систему подаються певні вхідні дані, які проходять через кілька вузлів обробки і користувач отримує на виході результат,  Пакет – це структура, яка може містити довільну кількість різноманітних даних. Тип цих даних може бути як користувацьким типом даних (структура, клас), так і одним із типів даних, передбачених у С/C++ (bool, int, char, double і т.п.). Також, в роботі графа застосована синхронізація між пакетами, задля забезпечення коректної і стабільної роботи системи.  Синхронізація у MediaPipe досягається шляхом використання у пакетах часових міток і застосуванні механізму обробки цих міток. Виконання графа MediaPipe децентралізовано. Це забезпечує більш високу пропускну здатність за рахунок конвеєрної обробки.  Калькулятор буде здійснювати обробку найновіших пакетів, у яких буде співпадати часова мітка із тією міткою, яка приходить з вхідного потоку.  Граф насправді є поєднанням комбінації менших графів, які з’єднані потоками. Дані графи називають підграфами. Це можна уявити конструктор, який будується із безліч невеличких блоків – даних підграфів. В них здійснюється налаштування вхідних потоків, вихідних потоків та вхідних статичних пакетів. Після цього, всі дані вносяться у один конфігураційний файл з розширенням .pbtxt, який зберігає дані про граф. Таким чином, відбувається створення графа і побудова плану, по якому відбувається виконання методів класів (калькуляторів).  Під час проектування модуля керування жестами, був застосований граф MediaPipe Hands. Із структурою підграфу, який відповідає за виявлення кисті руки, можна ознайомитися на рисунку 2.2.1.1.. | |
|  | *26* |

|  |  |
| --- | --- |
| Рис.2.2.1.3. – Підграф виявлення кисті руки | |
|  | *27* |

|  |  |
| --- | --- |
| MediaPipe Hands забезпечує рішення точного відстеження рук і пальців. Він використовує засоби OpenCV та алгоритми машинного навчання, щоб надати 21 трьохмірний орієнтир руки з усього одного кадру і позначає область, у якій виявлена долонь. Даний метод вирішує проблему виявлення кисті руки з використанням різних стратегій. Детектор кисті руки – це нейромережа, котра пройшла навчання на величезній кількості даних, а самі дані підготовленні засобами OpenCV. Таким чином, забезпечується отримання 21 ключової точки на кисті руки (рис. 2.2.1.4.).    Рис. 2.2.1.4. – Ключові точки долоні, які виявляються MediaPipe Hands  Виявлення відбувається навіть частково видимих рук або у випадку, якщо одна рука перекриває іншу – завдяки регресії, тобто прямого передбачення координат.  Отже, алгоритм отримує у якості вхідних даних кадр відеопотоку, де ймовірно зображена кисть руки, і у разі, якщо кисть присутня – виявляє її і надає масив з 21 точки, кожна точка якої містить три координати, які описують положення точки у трьохмірному просторі. | |
|  | *28* |

|  |  |
| --- | --- |
| **2.2.2. Бібліотека OpenCV.**  **2.2.2.1. Загальні відомомсті.**  OpenCV – це бібліотека, яка забезпечує можливість використання алгоритмів комп’ютерного зору і деяких інших алгоритмів по обробці зображень. Дана бібліотека написана на мові C/C++ що надає чудову швидкодійність програмного коду. Вона надає можливість здійснювати інтерпретації зображень, калібровки камери по еталону, ліквідацію шумів на зображені, визначення форми різноманітних об’єктів, здійснювати розділення за кольорами і т.п..  При проектуванні об’єкта дипломної роботи, дана бібліотека використовувалася для вирішення низки задач, також вона застосовується у роботі раніше згаданої бібліотеки MediaPipe. Надалі буде описано вибірку алгоритмів, які були застосовані у проектуванні модуля розпізнавання жестів.  **2.2.2.2. Розмивання Гауса.**  Розмивання Гауса – це алгоритм, який здійснює розмиття зображення. Сам алгоритм був названий у честь німецького математика Карла Фрідріха Гауса.  Розмиття є інструментом обробки зображення. Основною метою застосування є зменшення деталізації зображення, що використовується для багатьох алгоритмів обробки зображень, завдяки зменшенню кількості шумів на зображені. Розмиття здійснюється завдяки застосування функції Гауса:  (2.2.2.2.1)  де *a*,*b*,*c –* довільні дійсні числа. | |
|  | *29* |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| На рисунку 2.2.2.2.1. можна побачити один із варіантів роботи даної функції:    Рис. 2.2.2.2.1. – Розмиття за Гаусом  **2.2.2.3. Кольорова модель HSV.**  Кольорова модель HSV (вимовляється як «ейчесві») – математична модель представлення кольорів зображення, у якій застосовується певна кількість кортежів (упорядкований набір фіксованої довжини) чисел. Модель представляється через три значення: Hue (тон), Saturation (насиченість), Value (яскравість). На рисунку 2.2.2.3.1. демонструється просторове розміщення даних координат.    Рис. 2.2.2.3.1. – Просторове розміщення Hue, Saturation, Value. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Зазвичай приймають значення у діапазоні від 0 до 100 (окрім тому, стандартний діапазон якого від 0 до 360), проте, за необхідності, діапазон може бути масштабований до необхідних значень.  Дана кольорова модель використовується для попередньої обробки зображень, зокрема для алгоритмів порогових операцій (алгоритми, які дозволяють виділити один певний кольоровий спектр із зображення). На рисунку 2.2.2.2.3.2. зображенно порівняння зображень, які використовують дві кольорові моделі – RGB та HSV (зліва – RGB, справа - HSV)    Рис. 2.2.2.3.2. – Кольорові моделі – RGB та HSV  **2.2.2.4. Порогові операції.**  Порогові операції – це методи сегментації заданих, через пороги, кольорів, тобто – виділення певного кольору на зображенні. Якщо розглянути рисунок просторового розташування HSV у попередньому розділі (рис. 2.2.2.3.1.), то можна побачити, як даний «циліндр» розділений на умовні сегменти. Даний сегмент описується парою координат Hue, Saturation, Value («мінімальний» та «максимальний» пороги), які і описують ці сегменти. Сегмент буде містити колір, який необхідно виділити. Після того, як сегмент задано, заданий колір на рисунку набуває білого забарвлення, а все інше – чорного. Зображення стає бінарним (зображення, яке складається лише з двох кольорів – білого і чорного). На рисунку 2.2.2.4.1. продемонстровано приклад виділення жовтого кольору. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Рис. 2.2.2.4.1. – Виділення жовтого кольору  Таким чином, за допомогою заданих порогів, виділяється сегмент, що забезпечує виокремлення потрібного кольору (кольорового спектру).  **2.2.2.5. Алгоритм виявлення контурів.**  Алгоритм виявлення контурів, що використовується в OpenCV, є алгоритмом японського математика Сатосі Сузукі. Алгоритм відомий як Suzuki85 (англ. Сузукі85). Даний алгоритм забезпечує можливість ієрархічних відносин між границями контурів та здатний вказувати зовнішні та внутрішні границі. Загалом, алгоритм працює на основі аналізу пікселів і сусідніх пікселей, до пікселя який наразі розглядається. Обробка пікселей відбувається зверху-вниз і зліва на право. На бінарному зображенні всі чорні пікселі позначаються умовним нулем, білі – одиницею. У ході опрацювання пікселей, можливі два варіанта, які можуть підсказати, яка саме границя наразі розглядається – зовнішня чи внутрішня (рис. 2.2.2.5.2.).    Рис. 2.2.2.5.2. – Визначення границі | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Якщо перший піксель – нульовий, а наспупний є відмінним від нуля, то це повідомляє про виявлення контуру. Якщо поточний піксель позначений як ненульовий а наступний є нульовим, то розглядаємо це як закінчення контуру. Після того, як була досягнута дана ситуація, іде сканування сусідніх пікселів, спочатку за годинниковою стрілкою, потім – проти годинникової. Перевіривши ще низку умов, встановлюємо номери для кожного з контурів. Також необхідно мати змінну-лічильник, яка буде відслідковувати поточну кількість контурів. Таким чином, можна позначити пікселі, які відповідають одним і тим самим умовам, однаковими значеннями. Пікселі, що мають однакові значення, формують окремий контур.  **2.2.2.6. Алгоритм Рамера-Дугласа-Пекера.**  При роботі алгоритму по виявленню контурів (наприклад Сузукі85) може бути ситуація, коли алгоритм виявить величезну кількість контурів, кожен з контурів буде доволі коротким, а самі контури утворять ламану, яка містить безліч точок (рис. 2.2.2.6.1.)    Рис. 2.2.2.6.1 – Контур фігури (фіолетовим)  Наявність величезної кількість точок може нашкодити швидкодійності програми, внаслідок збільшення часу опрацювання усіх точок. Для зменшення точок, застосовують алгоритм Рамера-Дугласа-Пекера. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Даний алгоритм здійснює апроксимацію ламаної до ламаної з меншою кількістю точок. Наприклад, якщо застосувати даний алгоритм на контурі з рисунку 2.2.2.6.1, то отримаємо більш гладкішу ламану, зображену на рисунку 2.2.2.6.2.    Рис. 2.2.2.6.2. – Апроксимований контур  Рамера-Дугласа-Пекера видно, що даний контур є ламаною з меншою кількістю точок переломів, що спрощує подальші оперції над контуром. Наша ламана, яку потрібно апроксимувати, розглядається як упорядкована вибірка з точок переломів. Також, задається довільне дійсне число ε > 0. Алгоритм рекурсивно проводить поділ ліній. Чим більше епсілон, тим більш апроксимованим буде контур. Вибирається дві точки (зазвичай початкова і кінцева точки ламаної) і між ними проводиться лінія. Відбувається пошук точки ламаної, яка є найбільш віддалена від даної лінії. Надалі поточні точки сполучаються із цією точкою і дана точка позначається як така, яка буде збережена . Проводиться перевірка відстані між іншими точками, які ще не були задіяними, і проведеними прямими. Якщо відстань між точною і прямою  менше за ε, то така точка відкидається. Надалі проводимо нову ітерацію до тих пір, поки будуть оброблені всі точки. Таким чином, отримуємо апроксимовану ламану. На рисунку 2.2.2.6.3. показано приклад апроксимації ламаної. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Рис. 2.2.2.6.3. – Приклад апроксимація ламаної  **2.2.2.7. Поєднання роботи алгоритмів.**  Тепер, оглянувши алгоритми, можна навести приклад їхнього поєднання задля досягнення певних результатів. Наприклад – виявити об’єкт із заданим кольором. На рисунку 2.2.2.7.1. проілюстровано застосування алгоритмів і їх результат. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Рис. 2.2.2.7.1. – Процес застосування алгоритмів  Маємо наступну послідовність дій:   1. Отримуємо зображення з камери у кольоровій моделі RGB (у OpenCV використовується BGR) 2. Застосовуємо Гаусівське розмиття, задля зменшення кількості шумів 3. Конвертуємо RGB в HSV-модель 4. З HSV-моделі отримуємо необхідний колір, застосувавши порогові операції | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| 1. З бінарного зображення отримуємо контури за допомогою аглоритма виявлення контурів (Сузукі 85) 2. Застосувавши алгоритм Рамера-Дугласа-Пекера, апроксимуємо контур   Отже, застосувавши алгоритми, отримали замкнутий колір, який позначає на зображенню місцерозташування шуканого об’єкта.  **2.2.3. Система збірки Bazel.**  Система збірки Bazel – це програмне забезпечення з відкритим кодом, що забезпечує можливість збірки та тестування програм. Має можливість створювати ПЗ для різних платформ. Підтримує наступні мови програмування: C, C++, Java , Go, Python , Objective-C и Bourne shell-скрипти. Під час розроблення Bazel, увага була сфокусована на досягнення прискорення збірки, зокрема застосовуючи розпаралелювання. Також є можливість через даний інструмент підключати код, бібліотеки і інші файли, які знаходяться віддалено, потрібні для компіляції. Основою Bazel є робоча область (WORKSPACE), конфігураційні файли збірки (BUILD) та цілі (targets).  WORKSPACE – файл, який має бути в кореневій папці проекту. В даному файлі можна задати глобальні налаштування.  BUILD – кожна окрема папка повинна мати свій власний BUILD-файл. Тут прописуються всі налаштування конфігурацій, підключення файлів з програмним кодом, бібліотек і т.п. і вказуються інші конфігураційні файли BUILD.  Target – цілі, або правила, які підлягають до виконання. Через них вказуються дії, які необхідно виконати у BUILD-файлах.  **2.2.4. Бібліотеки WiringPi та softPwm.**  Дані бібліотеки застосовуються для забезпечення зручної роботи з GPIO на Raspberry Pi. GPIO – це універсальний інтерфейс вводу-виводу, що надає | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| можливість обмінюватися інформацією із різноманітною апаратурою, завдяки змінні напрузі на пінах. (рис. 2.2.4.1).  GPIO pins Рис. 2.2.4.1 – GPIO Raspberry Pi  Через те, що безпосередня робота з пінами Raspberry Pi є доволі важкою, через необхідність вдаватися до низькорівневого програмування, функції з бібліотек WiringPi та softPwm надають легкий спосіб використання цих пінів.  WiringPi - це бібліотека керування за допомогою GPIO, написана на мові C для пристроїв Raspberry Pi. Дана бібліотека розроблялася виключно для керування загальним інтерфейсом вводу-виводу на Raspberry Pi. Функції у бібліотеці наближені до функцій, які використовуються для написання програмного коду для апаратно-програмного пристрою Arduino на мові С/С++.  softPwm – бібліотека, яка використовує бібліотеку WiringPi для додавання певних можливостей, зокрема для реалізації цифрової PWM-модуляції (ШІМ- модуляція). Широтно-імпульсна модуляція - процес керування шириною (тривалістю) високочастотних імпульсів за законом, який задає низькочастотний сигнал. Використовується для керування сервоприводів.  Принцип роботи сервопривода на основі ШІМ – подання імпульсів певною тривалістю. В залежності від тривалості імпульсу, буде задаватися поворот сервопривода. На рисунку 2.2.4.2 наглядно показано, як від тривалості залежить кут повороту сервопривода. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Рис. 2.2.4.2 – ШІМ сервоприводів  Як бачимо, залежно від тривалості сигналу, поворот буде здійснено уліво (варіант а) – тривалість менше тривалості «нейтральної позиції»), управо (варіант в) – тривалість більше тривалості «нейтральної позиції»), або прийняти положення без поворотів (варіант б) – тривалість рівна тривалості «нейтральної позиції»).  **2.3. Опис нейронної мережі – алгоритму зворотного розповсюдження  помилки.**  Метод зворотного поширення помилки – це ітеративний градієнтний алгоритм, який забезпечує мінімізацію помилки багатошарової нейронної мережі. Як було сказано раніше, алгоритм використовує градієнтний спуск, що надає можливість знаходження локального мінімуму функції. Сама нейромережа – це функція багатьох змінних, а локальний мінімум – є наближенням до правильного значення функції при тих чи інших вхідних даних.  В чому полягає ідея методу? Під час роботи нейромережі, сигнали (дані) від вхідного шару нейромережі поширюються по прихованим шарам до вихідного шару, де формується відповідь нейромережі.Відповідь порівнюється із «правильною» відповіддю розраховується помилка нейромережі. Після цього отримана помилка розпочинає поширюватися у зворотному напрямку – від вихідного шару по прихованим до вхідного шару. У процесі поширення, коригуються значення ваг та персептронів.  Розглянемо загальну структуру нейромережі (рис. 2.3.1). | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Рис. 2.3.1 – Загальна структура нейромережі  Нейромережа може мати довільну кількість нейронів і прихованих шарів і завжди по одному вхідному і вихідному шарах.  Отже, введемо позначення:  X = (x1,…,xk) - вхідні вектори, Xi є ℝn (2.3.1)  B = (b1,…bk) - здвиги (2.3.2)  A = (a1,…,ak) - правильні вхідні вектора, Ai є ℝm (2.3.3)  (X, A) - навчальна вибірка (2.3.4)  W - вектор ваг нейронної мережі (2.3.5)  N(W, X) - функція, яка описує нейронну мережу (2.3.6)  Y = N(W, X) - відповідь нейронної мережі, Y є ℝm (2.3.7) | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| D(Y, A) = 2  - функція помилки (2.3.8)  Di(Y) = D(Y, Ai) - функція помилки на і-му прикладі (2.3.9)  Ei(W) = Di(N(W, Xi)) - помилка мережі на і-му прикладі (2.3.10)  E(W) = - помилка мережі на всій навчальній вибірці (2.3.11)  Завдання алгоритму можна звести до наступнийх цілей, а саме знайти:  Вектор W такий, що E(W) → min(навчання на всій вибірці)  Вектор W такий, що Ei(W) → min(навчання на всій вибірці)  Для початку, необхідно зрозуміти, як саме працює персептрон – робоча одиниця в нейронній мережі. На рисунку 2.3.2 представлена загальна схема персептрона.    Рис. 2.3.2 – Схема перцептрону | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Перцептрон – це реалізація певної математичної функції (2.3.13):  (2.3.13)  де *f* – функція активація;  *wij* – ваги;  *xi* – вхідні значення;  *b* – здвиг;  *y* – результат обчислень пецсептрону.  До перцептрону підключені ваги *wij* (у даному випадку розглядається один перцептрон, і тому кількість ваг – *w11…wk1*, де перша цифра позначає номер вхідних даних (*xi*, *i є [1…k]*) а друга цифра позначає номер перцептрону, а через те що перцептрон один, то *j = 1*), які представляють собою дійсне число у межах (0…1). *xi… xk* – набір вхідних значень, які передається перцептрону. Це може бути як звичайні дійсні числа, так і результат роботи інших перцептронів.  *f* – функція активації (або порогова). Дана функція надає результат у діапазоні [0…1] або [-1…1]. Вона симулює роботу біологічного нейрону – нейрон можу бути незбудженим (значення перцептрону близьке до 0 або -1), або бути збудженим (значення перцептрону близьке до 1). Розенблат використовував у своїй реалізації персептрону функцію сегмоїди (2.3.14):  (2.3.14)  Графік функції сигмоїди зображений на рисунку 2.3.3.    Рис. 2.3.3 – Графік функції сигмоїди | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Функція сигмоїди надає результат у діапазоні [1…0].  Також, для реалізації роботи персептрону, можна використовувати інші функції активації, зокрема деякі з них наведен на рисунку 2.3.4.    Рис. 2.3.4 – Функції активації  Здвиг *b* – це спеціальна вага, значення якої множиться на 1 та додається до суми добутків значень вхідних перцептронів на відповідні ваги. Вона грає роль певного «стимулятора», щоб у разі обчислень перцептрону не отримувати значення, рівне 0. Тепер ознайомимося з процесом навчання.  Як було раніше зазначено, робота нейромережі базується на надані їй вхідних даних, які проходять через приховані шари нейромережі і прямують до вихідного шару, де сформовується остаточна відповідь. Дана відповідь залежить від значень ваг перцептронів. Якщо нейромережа надає невірний результат, необхідно відкорегувати ці значення. Вручну видозмінити їх майже неможливо – по-перше, невідомо, як саме потрібно змінити ту чи іншу вагу, а по-друге, кількість ваг може бути тисячі. На допомогу приходить алгоритм зворотного поширення помилки. Алгоритм надає можливість встановити числову різницю між відповіддю, що надала нейромережа і еталонної відповідді. Після цього, дану помилку нейромережі алгоритм «протягує» через усю мережу – від вихідного шару, через приховані шари і до вхідного шару, змінюючи ваги і значення персептронів. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Розглянемо рисунок 2.3.5.    Рис. 2.3.5. – Приклад нейромережі  Дана нейромережа представлена двома перцептронами n1 та n2, двома входами x1, x2, вагами w11, w12, w21, w22 та здвигами b1, b2  Проведемо умовні обчислення:  Значення першого перцептрону: *y1 = sigm(w11x1 + w21x2 + b1)* (2.3.15)  Значення другого перцептрону: *y2 = sigm(w12x1 + w22x2 + b2)* (2.3.16)  Також, нехай *S1 = w11x1 + w21x2 + b1,* (2.3.17)  *S2 = w12x1 + w22x2 + b2* (2.3.18)  Ми отримали відповідь нейромережі і тепер необхідно звіритися з еталонною відповіддю з навчальної вибірки (2.3.4). Для цього обчислимо функцію помилки двох змінних (2.3.19):  *Dk(y1,y2) = (y1 – a1)2 + (y2 – a2)2* (2.3.19)  Таким чинном, помилка нейромережі на певній навчальній вибірці (2.3.10) буде виглядати наступним чином:  *Ek(W) = Dk(y1(w11, w21, b1), y2(w12, w22, b2))* (2.3.20) | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Алгоритм ЗПП надає можливість корегувати ваги перцептронів через похідку *Ek* від необхідної ваги перцептрона:  (2.3.21)  Для цього необідно знайти часткові похідні. Розглянемо цей процес на прикладі нейромережі з попереднього рисунку 2.3.  Розглянемо знаходження похідної Ek від w21. Ми маємо помилку нейромережі (2.3.20), яка основується на функції помилки двох змінних (2.3.19). Для того, щоб знайти похідну Ek від w21, необхідно знайти часткові похідні Ek від y1 та *E*k від y2, а у свою чергу, треба буде знайти часткові похідні y1 від w21 та w21 та y1 від w21. Таким чином, розпишемо формулу (2.3.19) для цього окремого випадку:  (2.3.22)  Розпишемо часткові похідні:  (2.3.23)  (2.3.24)  (2.3.25)  (2.3.26)  Таким чином, сформуємо остаточний варіант:  (2.3.27)  Отже, формула (2.3.24) надає помилку вказаної ваги. Тепер необхідно модифікувати вагу: | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| (2.3.28)  Сформуємо загальні формули для знаходження похідної *Ek від* вказаної ваги *w:*  *Dk(y1,…yn) = (yi – ai)2 + … + (y2 – an)2* (2.3.29)  (2.3.30)  (2.3.31)  (2.3.32)  (2.3.33)  (2.3.34)  (2.3.35)  Таким чином, формули (2.3.29 - 2.3.35) задають правила для видозміннення ваг персептронів. Змінна здвигів відбувається за аналогічною послідовністю.  Проте, нейромережа може складатися з багатьох прихованих шарів. Необхідно надати можливість передавати помилку в глибину нейромережі. При фомуванні попередніх правил корегування ваг, ми брали значення помилки, сформоване в результаті роботи перцептрона y1, y2.  Проте, якщо розглядати вхідні значення (2.3.1.) то можна побачити, що у свою чергу це також є результатом роботи перцептрона і до цих перцептронів можна застосувати ті самі правила по корегуванню ваг (рис. 2.3.6.). | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| Рис. 2.3.6 – Нейромережа з прихованим шаром  Тепер, щоб відкорегувати ваги *v*, необхідно здійснити корегування значень перцептронів прихованого шару (*n11, n12*). Здійснити це можна, використавши формули (2.3.29 - 2.3.35), проте тепер необхідно знайти похідну *Ek* відвказаного значенян перцептрона *x*. Отримуємо формули для знаходження помилки значень перцептронів (2.3.36 - 2.3.42):  *Dk(y1,…yn) = (yi – ai)2 + … + (y2 – an)2* (2.3.36)  (2.3.37)  (2.3.38)  (2.3.39)  (2.3.40)  (2.3.41) | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| (2.3.42)  Таким чином, ми передаємо помилку на значення перцептронів, що знаходяться у прихованих шарах, і після цього ми мажмо змогу застосувати формули (2.3.26 - 2.3.32), які забезпечать корегування ваг перцептронів далі у нейромережі.  **2.4. Бібліотека JSON.**  Бібліотека JSON (вимовляється як «джейсон») – це бібліотека, яка надає текстовий формат обміну даними між різними ЕОМ. Окрім обміну, бібліотека надає можливості формувати файли для зберігання упорядкованої інформації і проводити читання таких файлів. Таким чином, її використання вимагається саме при записі та читанню файлів із структоризованим вмістом.  Під час реалізації предмету дипломного дослідження, передбачається використання бібліотеки для збереження даних, а саме: налаштувань нейромережі і навчальних даних.  **2.4. Висновки до розділу**  У другому розділі висвітлений перелік використаних технологій, які застосовувалися під час реалізації предмета дослідження дипломної роботи, наданий короткий опис, для кожної технології наведене коротке пояснення принципу її роботи. Окрім застосованих технологій і алгоритмів, розділ містить дані про обрані апаратні елементи, які будуть використані на етапі розробки модуля по керуванню жестами рук.  Був наданий перелік апаратних елементів, які будуть використовуватися у процесі реалізації модуля по керуванню жнстами. Одноплатний комп’ютер Raspberry Pi, сервоприводи, модуль камери, альт-азимутальна установка та | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| живлення. Висвітлено процес обробки зображень, що в майбутньому використовується для виявлення об’єктік (жестів). Наданий детальний опис теоретичних відомостей по бібліотекам MediaPipe та OpenCV. Надано перелік алгоритмів і методів по обробці зображень, а саме: розмивання Гауса, кольорова модель HSV, поргові операції, алгоритм виявлення контурів, алгоритм Рамера-Дугласа-Пекера. Для кожного з алгоритмів наведений їх короткий опис і укінці наданий приклад їх застосування для виявлення об’єктів. Наведений опис системи збірки Bazel, яка використовуватиметься під час компіляції і збірки програмного забезпечення. Наданий опис бібліотек WiringPi та softPwm, які забезпечують можливість простої взаємодії із загальним інтерфейсом вводу/виводу GPIO. Описано призначення бібліотеки JSON.  В останньому підрозділі висвітлені відомості про алгоритм штучного інтелекту, який забезпечить розпізнавання логіки виявлених жестів, а саме алгоритм зворотного поширення помилки (ЗПП). Коротко описана суть методу, і наведене детальне пояснення роботи даного алгоритму, з відображенням усіх формул та обчислень, разом з ілюстративним супроводом. Також, розглянуто невеликі приклади у застосуванні даних формул. | | | | | | | | | | | |
|  | | | | | | | | | | | *28* |
| **РОЗДІЛ 3**  **ПРОГРАМНА ТА АПАРАТНА РЕАЛІЗАЦІЯ МОДУЛЯ РОЗПІЗНАВАНЯН ЖЕСТІВ РУК ДЛЯ КЕРУВАННЯ КВАДРОКОПТЕРОМ**  **3.1. Реалізація концепції модуля розпізнавання жестів**  На початку робіт, то фізичному втіленню предмета дипломної роботи, а саме модуля розпізнавання жестів рук для керування квадрокоптером, необхідно було продумати загальну концепцію всієї системи. Модуль має відповідати наступним критеріям:   * Компактний розмір; * Автономність роботи; * Рухомий елемент, який забезпечуватиме достатній огляд для можливості розпізнання жесту; * Можливість кріплення модуля на коптер; * Інтерфейс, який забезпечить передачу даних від модуля до коптера.   Враховуючи дані вимоги, і загальну схему систем розпізнавання жестів (Розділ 1, рис. 1.2.1.) було обрано елементи для побудови (Розділі 2) і пропрацьована схема пристрою (рис. 3.1.1.). | | | | | | | | | | |
| Кафедра КСУ | | | | НАУ 20 13 20 000 ПЗ | | | | | | |
| ***Виконав*** | *Левчук В.В.* |  |  | Програмна та апаратна реалізація модуля розпізнавання жестів рук для керування квадрокоптером | ***Літера*** | | | ***Лист*** | ***Листів*** | |
| ***Керівник*** | *Глазок О.М.* |  |  | *Д* |  |  | *30* | *59* | |
| ***Консульт.*** |  |  |  | СП-436 123 | | | | | |
| ***Н. контроль*** | *Тупота Є.В.* |  |  |
| ***Зав. Каф.*** | *Литвиненко О.Є.* |  |  |

|  |  |
| --- | --- |
| Рис. 3.1.1. – Схема підключення компонентів модуля розпізнавання жестів  Отримавщи усі необхідні компоненти, здійснено фізичну збірку модуля (рис. 3.1.2)    Рис. 3.1.2. – Фізична реалізація модуля | |
|  | *31* |

|  |  |
| --- | --- |
| Після цього, було продумано інтегрування модуля у квадрокоптер і створення ескізу (рис. 3.1.3.).    Рис. 3.1.3 – Ескіз поєднання модуля і квадрокоптера  Тепер, маючи загальне представлення про систему вцілому, за допомогою програмних засобів у AutoCAD, була створена 3D модель квадрокоптера у поєднані з модулем розпізнавання жестів (рис. 3.1.4.). | |
|  | *32* |

|  |  |
| --- | --- |
| Рис. 3.1.4. – Модель у AutoCAD (Концептуальний)  Наступним етапом є створення програмного забезпечення, яке здійснюватиме функції по розпізнаванню жестів і надання необхідних команд квадрокоптеру.  Дане ПЗ має виконувати перелік вимог: здійснювати керування сервоприводами для керування альт-азимутальної установкою, реалізовувати процес отримання кадру з відеопотоку і здійснбвати попередню обробку, реалізовувати процес виявлення жестів, реалізовувати процес розпізнавання жестів, надавати інформацію на бортовий комп’ютер квадрокоптера.  В якості дисципліни для вивчення в програмному модулі: "Програмування мовою *С*++". Таким чином, для подальшої реалізації програмного коду, була створена діаграма прецедентів (відображена у обов’язковому графічному матеріалі.). На основі діаграми прецедентів, був створений програмний код (опис реалізації окремих класів у підрозділі 3.2.). Враховуючи написаний програмний код, була реалізована діаграма послідовностей (відображена у обов’язковому графічному матеріалі.). На основі програмного коду, для легшого розуміння структури програми, була реалізована діаграми класів (відображена у обов’язковому графічному матеріалі.). | |
|  | *33* |

|  |  |
| --- | --- |
| Загальний план роботи програми виглядає наступним чином. Після запуску програми через термінал з одним із аргументів: "default" (або не вказуючи ніяких параметрів), "learning", "make\_data", виконується один із сценаріїв програми.  Сценарій при запуску з аргументом "default" (або без аргументів). Програма запускається у звичайному режимі. Створюється об’єкти класів HandTracking та AltAzimuth і запускається основний метод processTracking(), у якому відбувається основна логіка програми. Паралельно, в окремому процесі відбувається виконання методу mainLoop(), об’єкта класа AltAzimuth, який забезпечує обертання альт-азимутальної установки. Поки не буде виявлена кисть, альт-азимутальна установка буде проводити сканування, у разі виявлення кисті – буде фокусувати камеру на кисті руки. Під час виконання методу processTracking(), здійснюються наступні дії:   * Об’єкт класу NeuralNetwork (нейромережа) здійснює завантаження конфігураційних даних; * Створюється об’єкт класу MediaPipeHandsDetector, що забезпечуватиме роботу по виявленню жестів; * Розпочинається робота циклу, який працюватиме до моменту, поки працює програма (або до аварійного завершення); * Отримуємо зображення з камери; * Передаємо зображення об’єкту класу MediaPipeHandsDetector для аналізу на наявність кисті руки. Якщо кисть була виявленна, вектор ключових точок кисті руки передається на обробку у нейромережу, яка аналізує дані і виявляє логіку жеста. У разі успіху, передає сигнал на пін; * Якщо жест був виявлений, здійснюється обробка кадру для подальшого знаходження центра мас кисті руки; * Відбувається обчислення центра мас кисті руки і відносно положенню центра мас на зображенні, здійснюється корегування положення камери; * Здійснюється відображення обробленого кадру зображення (якщо підключений дисплей); * Відбувається наступна ітерація циклу. | |
|  | *34* |

|  |  |
| --- | --- |
| Сценарій при запуску з аргументом "learning". Створюється об’єкт класу *HandTracking* і запускається метод *processLearning*. Під час роботи методу *processLearning*() здійснюється навчання нейромережі.  Сценарій при запуску з аргументом "make\_data" (або без аргументів). Створюється об’єкти класів *HandTracking* та *AltAzimuth* і запускається метод, який забезпечує створення навчальних даних. Паралельно, в окремому процесі відбувається виконання методу *mainLoop*(), об’єкта класа *AltAzimuth*, який забезпечує обертання альт-азимутальної установки. Поки не буде виявлена кисть, альт-азимутальна установка буде проводити сканування, у разі виявлення кисті – буде фокусувати камеру на кисті руки. У процесі здійснюється захоплення кисті руки, і, натискаючи клавіші 0-8, здійснюється формування навчальних даних, одного із 9 жестів.  Приклад роботи програми зображено на рисунку 3.1.6..    Рис. 3.1.6. – Результат роботи програми | |
|  | *35* |

|  |  |
| --- | --- |
| **3.2. Програмна реалізація модуля.**  **3.2.1. Клас *AltAzimuth***  Загальний опис класу *AltAzimuth.*  Клас AltAzimuth реалізовує логіку обертання альт-азимутальної установки, на яку вмонтована камера, яка забезпечує надання системі зображень. В загальному, в класі реалізовано два основних метода, які спрацьовують залежно від обставин. На початку роботи модуля, спрацьовує метод пошуку кисті руки (void searching()), який здійснює сканування області. Сканування області відбувається шляхом обертання камери горизонтально у межах 0˚ - 180˚. Вертикально, наведення альт-азимутальної установки здійснюється у межах 0˚ - 165˚. Якщо кисть була виявлена, розпочнеться виконання методу відстеження (void tracking()). Даний метод працює на основі місцерозташування центру мас точок кисті руки, відносно центру зображення. Надаючи загальний опис даному методу, обертання альт-азимутальної установки здійснюється наступним чином: якщо центр мас знаходиться правіше від центру зображення, камеру необхідно повернути ліворуч, і навпаки, якщо центр мас знаходиться лівіше від центру зображення – необхідно здійснити обертання праворуч. Подібним чином, логіка задана і для вертикального обертання – якщо центр мас вище – опускаємо камеру, якщо нижче – камера піднімається. Якщо центр мас знаходиться доволі близько від центру зображення (не далі відстані, що відповідає 15% половини висоти або ширини зображення.  Поля класу *AltAzimuth*.  *bool isWork –* змінна, яка відстежує роботи системи. Якщо значення змінної істино, то робота альт-азимутальної установки продовжується, інакше – припинити роботу.  *cv::Point point* – змінна, яка зберігає координати центра мас кисті руки. Це клас бібліотеки OpenCV, містить дві змінні *int x* та *int y.* Якщо дані змінні рівні -1, отже кисть не виявленна.  *cv::Point centre* – координати центра зображення. | |
|  | *37* |

|  |  |
| --- | --- |
| *int width, height* – змінні зберігають ширину і висоту зображення відповідно.  *int calibrationW, calibrationH* - відстанні, які будуть викоритовуватися при заданні зони у якій, при знаходжені point, сервоприводи не будуть рухатися (15% від половини ширини і довжини зображення).  *int pin0, pin1* – змінні зберігають номера пінів на GPIO Raspberry Pi, через які відбувається модуляція ШІМ (у роботі використані піни 24 та 23 відповідно).  *int range* – задає діапазон для ШІМ (у роботі використано значення 100).  int value – задає частоту для ШІМ (у роботі використано значення 50).  Методи класу *AltAzimuth*.  *void tracking(int delayQuantity = 50)* – даний метод реалізовує логіку обертання камери, для здійснення пошуку кисті руки. Здійснюється др того моменту, поки *isWork* є істиною і значення змінних *x* та *y* змінної *point* рівні -1. Аргумент *delayQuantity* задає швидкість обертання камери.  *void searching(int delayQuantity = 150)* – метод, що реалізовує обертання альт-азимутальної установки, що імітує процес відстеження кисті руки. Як було раніше сказано, логіка методу побудована на основі обчислення положення цетру мас кисті руки відносно центра кадра зображення. Аргумент *delayQuantity* задає швидкість обертання камери.  *void setPositionX(int delayQuantity)* – здійснює обертання альт-азимутальної установки у вертикальній площині. Аргумент *delayQuantity* задає швидкість обертання камери.  *void setPositionY(int delayQuantity)* – здійснює обертання альт-азимутальної установки у горизонтальній площині. Аргумент *delayQuantity* задає швидкість обертання камери.  *AltAzimuth(bool &isWork, cv::Point &point, int width = 640, int height = 480, pin0 = 23, pin1 = 24, range = 100, value = 50)* – конструктор класу. Також проводить налаштування GPIO за допомогою бібліотек wiringPi та softPwm.  *void mainLoop()* – головний метод класу. Залежно від умов, активує метод пошуку кисті руки, або метод відстеження.  *virtual ~AltAzimuth()* – деструктор класу. | |
|  | *38* |

|  |  |
| --- | --- |
| **3.2.2. Клас *HandTracking.***  Загальний опис класу *HandTracking.*  Клас *HandTracking* є головним класом у даному програмному забезпеченні. Він об’єднує у собі логіку всіх інших класів (окрім класу *AltAzimuth*, з яким зв’язаний опосередковано, надаючи дані про наявність кисті руки і центр мас). Даний клас може здійснювати три сценарії, залежно від аргументу, який задається у терміналі при запуску програми.  Якщо не було задано ніякого аргументу або задано аргумент «default», то розпочинається звичайне виконання роботи програми. Розпочинає виконання метод *processTracking()*. Спочатку здійснюється завантаження конфігураційних даних нейромережі і відбувається її ініціалізація. Створюється об’єкт класу *MediaPipeHandsDetector*, який відповідає за процес виявлення кисті руки і передачі даних у нейромережу на аналіз. Надалі відбувається входження програми у цикл, який завершується при умові, якщо змінна *isWork* набуде хибного значення. Ітерація циклу виглядає наступним чином:   * Отримуємо зображення із камери; * Дане зображення перевертається по діагоналі (через те, що розташування альт-азимутальної установки передбачається перевернутим). * У разі, якщо ніяке зображення не було отримано, це сигналізує закінчення роботи. Змінні *isWork* присвоюється значення *false* і ітерація циклу преривається оператором *break*. * Здійснюємо виклик методу *RunMPPGraph(cv::Mat &camera\_frame\_raw)* об’єкту класа *MediaPipeHandsDetector*. Даний метод здійснює обробку зображення і аналіз його на наявність кисті руки. У разі виявлення кисті руки, дані передаються у нейромережу на розпізнання жесту. * Якщо жест виявлено, здійснюємо обробку кадра зображення, для отримання центру мас кисті руки. * Проводимо обчислення центру мас кисті руки і передачу координат у об’єкт класа *AltAzimuth.* | |
|  | *39* |

|  |  |
| --- | --- |
| * Якщо жест не виявлено, у об’єкт класа *AltAzimuth* передаємо значенян координат центру мас рівні –1 (що означає, що кисть не виявлена). * На кінець ітерації, відображаємо оброблений кадр зображення.   Якщо був переданий аргумент «learning», то розпочинається процес навчання нейромережі. Розпочинає виконання метод *processLearning()*. Злдійснюєтьсявиклик функції *study(std::string type)*, який розпочинає процес навчання нейромережі. Після закінчення навчання, нейромережа зберігає дані.  Якщо був переданий аргумент «make\_data», то розпочинається процес створення навчальних даних для майбутнього навчання нейромережі. Робота схожа на роботу програми при запуску з аргументом «default», проте нейромережа неактивна і при натиснені клавіш цифр 0-8, здійснюється зберігання даних кисті руки з міткою відповідного жеста.  Поля класу *HandTracking*.  *bool isWork –* змінна, яка відстежує роботи системи. Якщо значення змінної істино, то робота класу продовжується, інакше – припинити роботу.  *bool isDetected* – змінна, яка приймає значення істини, у разі, якщо кисть руки була виявлена, інакше змінна приймає значення хибності.  *cv::VideoCapture capture* – змінна типу даних з бібліотеки OpenCV для захоплення відеопотоку.  *bool grab\_frames* – змінна позначає стан відеопотоку (*capture*). Якщо істина – процес захоплення відеопотоку працює, інакше – завершити процес захоплення відеопотоку.  *cv::Mat camera\_frame\_raw* – змінна зберігає кадр відеопотоку. *cv::Mat* – тип даних бібліотеки OpenCV і собою представляє матрицю, комірки якого описують зображення у певній кольоровій моделі (RGB, HSV).  *cv::Scalar lower{0,250,250}* та *cv::Scalar upper{5,255,255}* – змінні задають кольоровий спектр для порогових операцій.  *cv::Mat frameBlur, cv::Mat frameHSV, cv::Mat frameMask* – буферні змінні для обробки зображень.  *std::vector <std::vector<cv::Point>> contours* - змінна зберігає контури, які отримано після застосування алгоритма Сузукі85. | |
|  | *40* |

|  |  |
| --- | --- |
| *std::vector <std::vector<cv::Point>> contoursPoly* – змінна зберігає апроксимовані контури.  *cv::Point point* – змінна зберігає координати центра мас кисті руки.  *std::vector<int> hiddenLayers* – вектор, що містить кількість персоптронів у прихованих шарах. Кількість елементів *hiddenLayers* позначає кількість прихованих шарів.  *int inputPerceptrons* – зберігає кількість перцептронів у вхідному шарі.  *int outputPerceptrons* – зберігає кількість перцептронів у вихідному шарі.  *NeuralNetwork \*neuralNetwork* – вказівник на клас NeuralNetwork. На даний вказівник здійснюється динамічне виділення пам’яті, що забезпечує створення нейромережі.  Методи класу *HandTracking*.  *void calculateCenterOfMass()* – даний метод здйснює обчислення центру мас кисті руки, у разі її виявлення. Обчислення центру мас – це обчислення середнього арифметичного для координат *x* та *y* усіх ключових точок кисті руки.  Після підрахунку координат центру мас, по координатам рисується на зображені невелике рожеве коло, що візуально позначає центр мас на зображенні.Формат JPEG.  *void frameProcessing()* – метод здійснює обробку зображення для виявлення центру мас кисті руки. У процесі роботи метода застосовуються такі алгоритми, як: розмиття Гауса, конвертація у іншу кольорову модель (HSV), порогові операції, алгоритм знаходження контурів (Сузукі85), апроксимація контурів за алгоритмом Рамера-Дугласа-Пекера.  *HandTracking(int inputPerceptrons, int outputPerceptrons, std::vector<int> hiddenLayers)* – конструктор класу *HandTracking.* У даному конструкторі відбувається створення об’єкту класа NeuralNetwork і його ініціалізація. Зокрема вхідні аргумени задаються кількість перцептронів у вхідному та вихідному шару нейромережі, також задається кількість прихованих шарів і їхня кількість перцептронів.  *~HandTracking()* – деструктор класа *HandTracking.*  *void processTracking()* – основний метод класа. Містить у собі виклики | |
|  | *41* |

|  |  |
| --- | --- |
| ключових методів об’єктів класів і задає логіку роботи даного програмного забезпечення.  *void processLearning()* – метод, що реалізує процес навчання нейромережі і зберігання результату навчання.  *bool &getIsWork()* – метод повертає посилання на змінну siWork.  *cv::Point &getPoint()* – повертає посилання на змінну *point.*  **3.2.3. Клас *MediaPipeHandsDetector.***  Загальний опис класу *MediaPipeHandsDetector*.  Клас MediaPipeHandsDetector реалізовує логіку виявлення жестів за допомогою засобів бібліотеки MediaPipe. Принцип роботи бібліотеки, зокрема логіка виявлення кисті руки, висвітлено у Розділі 2. Загалом, принцип роботи даного класу полягає в наступному:   * Ініціалізація графу; * Створення захоплювачів, через які буде здійснено отримання Hand Landmarks, кількість виявлених долонь (у програмному коді бібліотеки був змінений параметр, який відповідав за кількість одночасно виявлених кисть рук на зображені до 1), обробленого зображення із графічними мітками ключових точок кисті руки; * Передача кадру зображення у граф і здійснення аналізу даного зображення графом на наявність кисті руки; * Отримання за допомогою захоплювачів необхідних даних; * Передача вектора з ключовими точками у нейромережу.   Поля класу *MediaPipeHandsDetector*.  *bool isDetected* – змінна, яка приймає значення істини, у разі, якщо кисть руки була виявлена, інакше змінна приймає значення хибності.  *mediapipe::CalculatorGraph graph* – граф MediaPipe.  *std::unique\_ptr<mediapipe::OutputStreamPoller> outputVideoPoller* –захоплювач, через який здійснюється отримання обробленого зображення із графічними мітками ключових точок кисті руки на ньому. | |
|  | *42* |

|  |  |
| --- | --- |
| *std::unique\_ptr<mediapipe::OutputStreamPoller> outputLandmarksPoller* – захоплювач, через який здійснюється отримання ключових точок кисті руки.  *std::unique\_ptr<mediapipe::OutputStreamPoller> outputHandCountPoller* – захоплювач, через який здійснюється отримання кількості виявлених кисть рук.  *std::string calculator\_graph\_config\_contents* - Змінна зберігає шлях до конфігураційного файлу (*\*.pbtxt*).  *mediapipe::CalculatorGraphConfig config* - Змінна зберігає конфігураційні дані.  *sizeframe\_timestamp\_us\_t* – змінна, що зберігає поточну часову мітку.  Методи класу *MediaPipeHandsDetector*.  *absl::Status Setup(std::string calculator\_graph\_config\_file)* – метод здійснює встановленя деяких конфігурацій для роботи графа.  *absl::Status Shutdown()* – метод для коректного завершення роботи об’єкта класу *MediaPipeHandsDetector.*  *MediaPipeHandsDetector(bool\_isDetected,\_std::string\_calculator\_graph\_config\_file = "mediapipe/graphs/hand\_tracking/custom\_hand\_tracking\_desktop\_live.pbtxt")* - конструктор класа MediaPipeHandsDetector. Здійснює встановлення змінної isDetected і виклик метода Setup().  *absl::Status RunMPPGraph(cv::Mat &camera\_frame\_raw)* – основний метод, що здійснює опрацювання кадру відеопотоку і надає можливіть отримати необхідну інформацію у вигляді вектора ключових точок кисті руки. Повертає *absl::StatusOk*, що є свідченням успішного виконання метода.  **3.2.4. Класи *NeuralNetwork* та *Perceptrons*.**  Загальний опис класу *Perceptrons.*  Даний клас є допоміжним для *NeuralNetwork* і реалізовує шар перцептронів. У тілі класу надано визначення структури *perceptron*, яка реалізовує окремий перцептрон. Клас *Perceptrons* містить у собі вектор, тип даних якого є *perceptron*. Даний вектор є втіленням шару перцептронів. | |
|  | *43* |

|  |  |
| --- | --- |
| Поля класу *Perceptrons*.  *struct perceptron* – структура, що реалізовує окремий перцептрон. Містить наступні змінні: *double value* – значення перцептрону, *double d\_value* – значення помилки перцептрону, *std::vector<double> weights* – вектор ваг, які ідуть від перцептрона, *std::vector<double> d\_weights* – вектор помилок ваг, які ідуть від перцептрона, *double derivative* – буферна змінна, яка викорисовується для обчислень.  *std::vector<perceptron> perceptronsLayer* – вектор, який реалізовує шар перцептронів.  *perceptron bias* – здвиг.  Методи класу *Perceptrons*.  *Perceptrons(int numbOfPerceptrons, int numbOfWeights)* – конструктор даного класу. Здійснює ініціалізацію шару пецептронів, в залежності вхідних аргументів.  *~Perceptrons()* – деструктор класу *Perceptrons.*  *int getQuantityOfPerceptrons()* – метод повертає кількість перцептронів у шару.  *int getQuantityOfWeights()* - метод повертає кількість ваг, що ідуть від будь-якого перцептрону у даному шару.  Загальний опис класу *NeuralNetwork.*  Клас *NeuralNetwork* здійснює реалізацію нейронної мережі на основі алгоритму зворотного поширення помилки. Детальний опис алгоритму наведено у другому розділі. Метод *void study(std::string type)* здійснює процес навчання.  Поля класу *NeuralNetwork*.  *int countLayers* - зберігає кількість слоїв у нейромережі.  *std::vector<Perceptrons\*> layers* - масив вказівників на слої нейромережі.  *std::string typeOfData* – задає тип жесту.  *int numbOfMotion* – задає номер жесту.  *std::vector<double> data* - вектор, що містить дані для вхідного шару нейромережі.  *std::vector<double> arr* - Буферна змінна, для зберігання значень вихідного шару. | |
|  | *45* |

|  |  |
| --- | --- |
| Методи класу *NeuralNetwork*.  *NeuralNetwork(int inputPerceptrons, int outputPerceptrons, std::vector<int> hiddenLayers)* – конструктор класу *NeuralNetwork.* За допомогою вхідних аргументів, задається кількість перцептронів на вхідному та вихідному шарах, кількість прихованих перцептронів і кількість перцептронів у кожному із прихованих шарів.  *~NeuralNetwork()* – деструктор класу *NeuralNetworІk.*  *void setInputPerceptrons(std::string type, std::vector<double> &data)* - функція встановлює значення на вхідному шарові персептрона. *vector<double> data* - вектор, що містить значення вхідного слою персептронів.  *void setPrimaryPerseptrons()* - метод виставляє початкові значення для значень персептронів.  *void setPrimaryWeights()* - метод виставляє початкові значення для ваг персептронів.  *double sigm(double x)* – метод, що реалізовує сигмоїду.  *double d\_sigm(double x)* - метод, що реалізовує похідну сигмоїди.  *void calculatePerceptrons(int begin, int end, int layer)* - метод обраховує персептрони наступного слою.  *void calculateNeuralNetwork()* - метод обраховує значення нейромережі.  *void setTrueAnswer(int numbPerceptron)* - метод виставляє в останьому слої нейромережі в змінних *d\_x* "правильні значення" персептронів. Потім відбувається обрахування *d\_x = x - d\_x*.  *void backPropagationBias(int layer)* - обчислення помилки для здвигу за методом ЗПП.  void backPropagationWeights(int layer) - обчислення помилки для ваг за методом ЗПП.  *void weightsDifference(double coefficient = 0.02)* – додає до ваг знайдені «посилки» (відбувається процес коренування ваг). *double coefficient* – задає швидкість навчання нейромережі.  *void backPropagationPerceptrons(int layer)* – метод, що реалізовує розповсбдження помилки на перцептрони. | |
|  | *46* |

|  |  |
| --- | --- |
| *void backPropagation()* - реалізація методу ЗПП.  *double levelOfEducation()* - функція визначає рівень навчиності нейромережі. Чим значення *double*, що повертає функція, ближче до нуля, тим більше навчина нейромережа.  *bool saveWeights(std::string path)* – збереження даних нейромережі, а саме ваг.  *bool loadWeights(std::string path, int numbOfLayer)* - завантаження даних нейромережі, а саме ваг.  *void loadData(std::string type)* – метод реалізує завантаження параметрів нейромережі, а саме кількість шарів нейромережі і кількість перцептронів у цих шарах.  *void study(std::string type)* – один із основних методів нейромережі. Здійснює виклик вибірки методів, що забезпечує повний цикл навчання за алгоритмом ЗПП.  *int getQuantityOfItemInFolder(std::string pathToFolder)* - технічний методі, для підрахування кількості навчальних даних у папці.  *void saveTrainData(std::string type , int numbOfFolder /\*Номер жеста\*/, const std::vector<double> &data)* – метод реалізовує логіку збереження навчальних даних по конкретному жесту.  *void loadTrainData(std::string type, int numbOfFolder, int numbOfFile, std::string &type\_out, int &numb\_out, std::vector<double> &data\_out)* - метод реалізовує логіку завантаження навчальних даних по конкретному жесту.  **3.3. Інструкція по експлуатації.**    Початок експлуатації модуля не є важким. Перш за все, необхідно переконатися у тому, що Raspberry Pi увімкнута. Після цього, необхідно відкрити термінал і перейти у теку, в якій знаходиться папка з проектом і зайти у дану папку. | |
|  | *47* |

|  |  |
| --- | --- |
| Після цього запустити команду:  *bazel-bin/mediapipe/examples/desktop/controle\_module/controle\_module <аргумент>*  де замість *<аргумент>* необхідно підставити один із наступних аргументів командного рядка:   * *default* – запускає модуль у звичайному режимі роботи; * learning – запускає процес навчання нейромережі; * *make\_data* – запускає процес формування навчальних даних. Для того, щоб створити навчальний жест, необхідно надати кисті руки необхідного вигляд необхідного жесту, і натиснути клавішк 0-8 на клавіатури, що позначає необхідний жест.   У разі, якщо не було надано ніяких аргументів, або аргумент був прописаний невірно, програма запускається у режимі *default.*  **3.4. Висновки до розділу.**  У даному розділу було надано детальний опис процесу реалізації предмету дипломного дослідження, а саме створення модуля розпізнавання жестів рук для керування квадрокоптером.  Було проведено роботу по висвітленню всього процесу побудови модуля. Детально наведена інформація по реалізації концепції модуля розпізнавання жестів, а саме наведенені критерії, за якими визначалося подальша розробка. Зокрема, модуль мав відповідати наступним вимогам:   * Компактність; * Автономність; * Рухома камера; * Можливість розмістити модуль на квадрокоптері.   Враховуючи перелічені вимоги, була вибрана апаратура, яка необхідна для успішної реалізації модуля. Таким чином, основою, на якій базуватиметься робота усього пристрою, став одноплатний комп’ютер Raspberry Pi 3B+, який має | |
|  | *48* |

|  |  |
| --- | --- |
| **ВИСНОВКИ**  Під час написання даної дипломної роботи, був пророблений комплекс дій, які забезпечили повне виконання предмету дипломної роботи, а саме створення модуля розпізнавання жестів рук для керування квадрокоптером. Модуль відповідає загальній структурі інтерактивним системам розпізнавання жестів. Модуль виконує необхідні задачі, а саме:  - здійснює захоплення кадрів відеопотоку та забезпечує обертання відеокамери на основі руху альт-азимутальної установки;  - здійснює виявлення кисті руки;  - здійснює процес розпізнавання жестів  - має можливість виводити оброблені кадри відеопотоку на графічний дисплей.  Вся робота по реалізації модуля розпізнавання жестів проводилася у кілька етапів, які забезпечили послідовне та структуризоване виконання поставленої задачі:  1. Було наведено пояснення, що саме розуміється під терміном «інтерактивна система керування жестами».  2. Здійснено вичерпне опрацювання теоретичного матеріалу по темі дипломної роботи. Була розглянута низка технічної літератури по комп’ютерному зору та штучному інтелекту, а саме опрацьовано і вивчено матеріал, який стосується нейронних мереж. Крім технічної літератури, проведене ознайомлення із науковими статтями, які стосуються теми дипломної роботи.  3. Надано опис алгоритмівта методик, які використовуються при проєктуванні модуля розпізнавання жестів. Висвітлено відомості з теорії розпізнавання образів, відомості про загальний принцип виявлення об’єктів та поясненно, якими інструментами та методами досягається результат такої роботи. Викладені загальні положення про нейронні мережі, більш детально розглянуто алгоритм зворотного поширення помилки (ЗПП), який використовувався при реалізації нейронної мережі. | |
|  | *49* |

|  |  |
| --- | --- |
| Наведені поширені алгоритми, які використовуються у системах керування жестами, які також застосовуються під час виявлення об’єктів.  4. Проведена робота по ознайомленню з апаратними засобами і елементами, які використані у процесі реалізації модуля розпізнавання жестів або могли б бути використані. Оглянуту технічні характеристики даних апаратних засобів.  5. Була проведена робота по збірці апаратної частини модуля розпізнавання жестів. Сам модуль збирався з апаратних елементів, таких як: одноплатний ЕОМ Raspberry Pi 3B+, сервоприводів, альт-азимутальної установки, камери та блоку живлення. Проведені роботи по перевірці працездатності даної системи.  6. Було здійснено написання програмного коду, що надавав можливість здійснювати необхідну логіку модуля, яка відповідає меті даної дипломної роботи. Для забезпечення якісної роботи програмного забезпечення, були обдумані функції, які має виконувати дане ПЗ. На основі даних функцій, була створена діаграма прецеднентів, опираючись на яку, розпочалася робота над програмним кодом. Була проведена робота по забезпеченню кращого розуміння коду, тому були створенні діаграми послідовності та класів. У ході написання програмного коду, були виконані роботи по перевірці даного програмного забезпечення.  7. На кінець, були проведені роботи по перевірці роботи програмного коду у зібраній апаратній системі.  Таким чином, виконавши попередньо описані дії, був створений модуль розпізнавання жестів для керування квадрокоптером. Також, передбачається подальша можливість здійснювати модернізацію розробленої системи, заради розширення функціоналу і можливостей системи. | |
|  | *50* |

|  |  |
| --- | --- |
| **СПИСОК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ ВИКОРИСТАНИХ ДЖЕРЕЛ**   1. Вапник В. Н. Теория распознавание образов. Статические проблемы обучения // «Наука», Москва, 1974. 2. Васильев В.И. Распознавающие системи. Справочник // «Наукова думка», Киев, 1983. 3. Фрэнк Розенблатт Принципы Нейродинамики: перцептроны и теория механизмов мозга, «Мир», Москва, 1965. 4. Фарсайт Д.А., Понс Ж. Компьютерное зрение: современный подход // Вильямс, 2004. 5. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей// Москва, 2018 г. 6. Кононюк А.Е. Общая теория распознавание Книга 1 Начала // «Освіта України», Киев, 2012. 7. Распознавание жестов на основе нейросетевого подхода [Електронний ресурс] – Режим доступу до ресурсу: <https://cyberleninka.ru/article/n/raspoznavanie-zhestov-na-osnove-neyrosetevogo-podhoda/viewer> 8. Реализация интерактивных систем распознавание жестов [Електронний ресурс] – Режим доступу до ресурсу: <https://cyberleninka.ru/article/n/raspoznavanie-zhestov-na-osnove-neyrosetevogo-podhoda/viewer> 9. Система распознавание жестов на основе нейросетевых технологий [Електронний ресурс] – Режим доступу до ресурсу: https://cyberleninka.ru/article/n/sistema-raspoznavaniya-zhestov-na-osnove-neyrosetevyh-tehnologiy/viewer 10. Проблемы распознавание языка жестов и методы их решения [Електронний ресурс] – Режим доступу до ресурсу: https://cyberleninka.ru/article/n/problemy-raspoznavaniya-yazyka-zhestov-i-metody-ih-resheniya/viewer | |
|  | *58* |

|  |  |
| --- | --- |
| 1. Особливості реалізації технології обробки даних для розпізнання жестів [Електронний ресурс] – Режим доступу до ресурсу: <https://cyberleninka.ru/article/n/osobennosti-realizatsii-tehnologii-obrabotki-dannyh-dlya-raspoznavaniya-zhestov/viewer> 2. MediaPipe [Електронний ресурс] – Режим доступу до ресурсу: https://mediapipe.dev/ 3. OpenCV [Електронний ресурс] – Режим доступу до ресурсу: https://docs.opencv.org/master/index.html 4. Ravesli [Електронний ресурс] – Режим доступу до ресурсу: https://ravesli.com/uroki-cpp/ 5. Бойченко С. В., Іванченко О. В. Положення про дипломні роботи (проекти) випускників Національного авіаційного університету. Київ : НАУ, 2017. 63 с. 6. ДСТУ 3008-95. Документація. Звіти у сфері науки і техніки. Структура і правила оформлення : видання офіційне, Київ : Держстандарт України, 1995. 38 с. | |
|  | *58* |

|  |  |
| --- | --- |
| **Додаток А**  **Лістинг коду реалізації головного вікна програмного модуля** | |
|  | *59* |