

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Харківський національний університет радіоелектроніки
Факультет Комп'ютерних наук
Кафедра Програмної інженерії

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти

другий (магістерський)

«Дослідження методів машинного навчання для розпізнавання об'єктів та
прогнозування ймовірності та напрямку їх руху у режимі реального часу»
(тема)

Виконав:

студент 2 курсу, групи ІПЗм-21-2
Ванда М.А.
(прізвище, ініціали)

Спеціальність 121- Інженерія програмного
забезпечення
(код і повна назва спеціальності)

Тип програми Освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Керівник доц. Кравець Н.С.
(посада, прізвище)

Допускається до захисту
Зав. кафедри

(підпис)

З.В. Дудар
(прізвище, ініціали)

2023р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Програмної інженерії
(повна назва)

Рівень вищої освіти другий (магістерський)

Спеціальність 121- Інженерія програмного забезпечення
(код і повна назва спеціальності)

Тип програми Освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Інженерія програмного забезпечення
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« ____ » _____ 2023 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студента Ванди Максима Анатолійовича
(прізвище, ім'я, по батькові)

- Тема роботи Дослідження методів машинного навчання для розпізнавання об'єктів та прогнозування ймовірності та напрямку їх руху у режимі реального часу.
затверджена наказом університету від 83Стз № 03.04.2023
- Термін подання роботи до екзаменаційної комісії 18.05 2023р.
- Вихідні дані до роботи характеристики архітектур нейронних мереж, AutoML Vision, вибірка датасетів, Kaggle, мова Python, документація по машинному навчанню.
- Перелік питань, що потрібно опрацювати в роботі аналіз предметної області та постановка задачі, визначення переваг кожної з архітектур нейронних мереж, обрання датасету для навчання, створення та навчання моделей нейронних мереж, проведення експерименту.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної галузі	23.01.2023	виконано
2	Постановка задачі	28.01.2023	виконано
3	Аналіз методів машинного навчання	12.02.2023	виконано
4	Пошук датасетів для дослідження	01.03.2023	виконано
5	Створення та тренування моделей	08.03.2023	виконано
6	Дослідження	17.03.2023	виконано
7	Оформлення статті	28.03.2023	виконано
5	Написання пояснювальної записки	02.04.2023	виконано
12	Підготовка презентації та доповіді	10.05.2023	виконано
13	Нормоконтроль	13.05.2023	виконано
14	Рецензування	16.05.2023	виконано
15	Занесення диплома в електронний архів	18.05.2023	виконано
16	Захист кваліфікаційної роботи	20.05.2023	

Дата видачі завдання 23 січня 2023 р.

Студент _____
(підпис)

Ванда М.А.
(прізвище, ініціали)

Керівник роботи _____
(підпис)

доц. Кравець Н.С.
(посада, прізвище, ініціали)

РЕФЕРАТ / ABSTRACT

Пояснювальна записка до кваліфікаційної роботи містить 72 сторінки, 25 рисунків, 20 джерел.

AUTOML VISION, CNN, GOOGLE, NAS, RNN, TESLA, WEB, АВТОМОБІЛІ, АВТОПІЛОТ, ВИЗНАЧЕННЯ ОБ'ЄКТІВ, КВАЛІФІКАЦІЙНА РОБОТА, МЕТОДИ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ, ПОШУК ОБ'ЄКТІВ, РОЗПІЗНАВАННЯ ОБРАЗІВ.

Об'єктом дослідження є методи машинного навчання для розпізнавання об'єктів та прогнозування ймовірності та напрямку їх руху у режимі реального часу, а також виявлення їх кількості з подальшою перспективою застосування в роботизованому авто чи автопілоті автомобіля.

Метою роботи є дослідження методів машинного навчання для розпізнавання об'єктів на основі існуючих продуктів з перспективою застосування у роботизованих автомобілях або автопілоті.

При виконанні роботи були використані загально-логічні, теоретичні та емпіричні методи. Загально-логічні використовувались для проведення аналізу предметної області, теоретичні були використані під час аналізу нейронних мереж, а емпіричні було застосовано під час проведення експериментального дослідження.

У результаті виконання кваліфікаційної роботи було розроблено та досліджено методи машинного навчання для розпізнавання об'єктів та прогнозування ймовірності та напрямку їх руху у режимі реального часу. Проведено експеримент з їх порівняння та визначення найкращої моделі для заданих завдань.

AUTOML VISION, CNN, GOOGLE, NAS, RNN, TESLA, WEB, CARS, AUTOPILOT, OBJECT DETECTION, QUALIFICATION WORK, OBJECT RECOGNITION METHODS, OBJECT SEARCH, PICTURE RECOGNITION.

The object of research is machine learning methods for recognizing objects and predicting the probability and direction of their movement in real time, as well as detecting their number with the further prospect of application in a robotic car or car autopilot.

The purpose of the work is to research machine learning methods for object recognition based on existing products with the prospect of application in robotic cars or autopilot.

General logical, theoretical and empirical methods were used in the performance of the work. General-logical ones were used for the analysis of the subject area, theoretical ones were used during the analysis of neural networks, and empirical ones were applied during the experimental research.

As a result of the qualification work, machine learning methods were developed and researched to recognize objects and predict the probability and direction of their movement in real time. An experiment was conducted to compare them and determine the best model for the given tasks.

Умови публікації пояснювальної записки

Я,

Ванда Максим Анатолійович
(прізвище, ім'я, по батькові)

студент(ка) групи ІІЗм-21-2 здобувач вищої освіти на другому
(магістерському) рівні

кафедра _____ програмної інженерії _____,
(повна назва кафедри)

заявляю: моя кваліфікаційна робота на тему

Дослідження методів машинного навчання для розпізнавання об'єктів та
прогнозування ймовірності та напрямку їх руху у режимі реального часу,
(назва роботи)

що буде представлена до ЕК для публічного захисту, виконана самостійно, в ній немістяться елементи плагіату і вона може бути опублікована в електронному архіві відкритого доступу EIArKhNURE. Всі запозичення з друкованих та електронних джерел мають відповідні посилання.

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

ЗМІСТ

Вступ.....	8
1 Аналіз предметної галузі та актуальність дослідження.....	11
1.1 Аналіз предметної галузі.....	11
1.2 Аналіз аналогів та актуалізація рішень	15
1.3 Актуальність та мета дослідження.....	24
1.4 Постановка задачі.....	25
2 Аналіз методів для дослідження.....	27
2.1 Обґрунтування методів досліджень та етапи його проведення	27
2.2 Огляд та вибір датасету	28
2.3 Огляд обраного інструменту для дослідження.....	30
2.4 Аналіз обраних методів для дослідження.....	34
3 Проведення теоретичного та експериментального дослідження.....	45
3.1 Розробка та проведення дослідження.....	45
3.2 Результати дослідження та можливість їх використання у практичній діяльності.....	55
Висновки	57
Перелік джерел посилання	59
Перелік джерел посилання за науковими напрямками науковців кафедри програмної інженерії.....	61
Додаток А Результати перевірки на унікальність тексту в базі ХНУРЕ.....	62
Додаток Б Слайди презентації.....	63
Додаток В Апробація результатів роботи.....	70
Додаток Г Експертний висновок результатів перевірки кваліфікаційної роботи на відповідність оформлення вимогам ДСТУ 3008:2015.....	72

ВСТУП

У сучасному світі людство досить швидко прогресує, адаптується та розвивається. Це стосується абсолютно кожної сфери діяльності, автомобільна індустрія не залишилася осторонь. Якщо взяти до уваги статистичні дані по кількості автомобілів, то можна дізнатися, що кілька років тому на дорогах світу було набагато менше транспортних засобів. Що зумовлювало відсутність проблем як для водіїв, так і для пішоходів.

У наші дні, більшість великих міст світу стикаються з проблемою збільшення кількості автомобілів. У сучасну епоху кожен у спільноті хоче мати комфортну подорож. Завдячуючи цьому багато людей користуються автомобілями для поїздок на роботу, що призвело до того, що міста переобтяжені автомобілями. Проблем, пов'язаних із зростанням кількості автомобілів, незліченні, але основною безсумнівно є збільшення кількості ДТП (дорожньо-транспортні пригоди).

За кожним автомобілем сидить водії, який несе відповідальність за безпечне пересування свого транспортного засобу. Більшість систем класифікації причин ДТП були зосереджені на помилках і діях водіїв, які і призвели до цього. Дорожньо-транспортні пригоди ж спричиняють значну частку смертності або отримання інвалідності та призводять до втрачених років життя більше, ніж більшість хвороб які знає людство.

Є безліч факторів впливу на водіїв, що спричиняють збільшення кількості ДТП: зниження дієздатності на довгостроковій основі (недосвідченість, старіння, хвороба та інвалідність, алкоголізм, зловживання наркотиками), ті, що знижують здатність на короткостроковій основі (сонливість, втома, гостра алкогольна інтоксикація, короткочасна дія ліків, переїдання, гострий психологічний стрес, тимчасове відволікання), ті, що сприяють ризикованій поведінці з довгостроковими наслідками (переоцінка своїх здібностей, поширене перевищення максимальної дозволеної швидкості, ігнорування правил дорожнього руху, невикористання заходів безпеки) та ті, що сприяють ризикованій поведінці з короткочасним впливом (помірне споживання етанолу, психотропні препарати,

злочини, пов'язані з транспортними засобами, суїцидальна поведінка, компульсивні дії). Усі ці фактори притаманні людині, що у свою чергу можуть призвести до летальних випадків на дорозі. Усі ці фактори мінімізуються при використанні роботизованого водія – автопілота [1].

Транспортні засоби можуть не тільки їздити самостійно, але й взаємодіяти між іншими подібними собі. Вони зможуть взаємодіяти з навколишньою інфраструктурою, світлофорами, центрами керування дорожнім рухом, пішоходами та іншими учасниками дорожнього руху. Усе це зможе позитивно вплинути на динаміку збільшення аварійних ситуацій, так як роботизований транспорт виключить людські фактори у ДТП. Все це розроблено за допомогою різноманітних систем розпізнавання об'єктів, що робить тему дуже актуальною в сучасних умовах.

У сучасному світі машинне навчання є однією з найбільш обговорюваних тем. Воно використовується в багатьох галузях, включаючи медицину, бізнес, науку та технології. Однією з найважливіших областей застосування машинного навчання є розпізнавання об'єктів. Розпізнавання об'єктів є процесом визначення ідентичності або класифікації об'єктів на зображеннях або відео. Цей процес може бути використаний для автоматичної ідентифікації людей, визначення видів тварин, розпізнавання автомобілів на дорогах та багато іншого.

Дослідження методів машинного навчання для розпізнавання об'єктів та прогнозування їх руху у режимі реального часу у автопілоті є важливою темою для розвитку безпілотних автомобілів та інших транспортних засобів. Система матиме можливість розпізнавати дорожні знаки для їх класифікації та ідентифікувати появу пішоходів на дорожній зоні. Після навчання модель може визначати клас об'єкта на зображенні. Система використовуватиме класифікацію, щоб ідентифікувати тип автодорожнього знака.

Одним з основних методів машинного навчання для розпізнавання об'єктів є навчання з вчителем на основі нейронних мереж. На основі цього працює інструмент для машинного навчання від компанії Google, створений для автоматизації процесу навчання моделей комп'ютерного зору. У цьому підході

модель навчається на наборі даних, що містять зображення об'єктів та відповідні мітки класів.

Прогнозування руху об'єктів може бути реалізовано за допомогою методів навчання з підсиленням, які використовуються в безпілотних автомобілях. У цьому підході модель навчається взаємодії з навколишнім середовищем, отримуючи винагороду за правильні дії та покарання за неправильні. Також можуть використовуватися методи комп'ютерного зору для виявлення та відстеження об'єктів. Ці методи можуть включати в себе детектування об'єктів, витягування ознак та класифікацію за рухом.

Метою роботи є дослідження методів машинного навчання для розпізнавання об'єктів та прогнозування ймовірності та напрямку їх руху у режимі реального часу. У ході дослідження буде проаналізовано методи, які можуть бути використані у роботизованому автомобілі для розпізнавання об'єктів на дорожній зоні.

Об'єктом дослідження є дорожні знаки за допомогою яких автопілоти орієнтуються у дорожньому просторі.

Предметом дослідження є актуальні методи нейронних мереж для розпізнавання об'єктів, задля використання їх у розумних автомобілях.

Завданням дослідження є аналіз існуючих методів для обробки та виявлення об'єктів, які необхідні автомобілю для безпілотного керування. Результати проведеного дослідження будуть корисні для подальших розробок систем які будуть впроваджені у роботизованому авто для моніторингу об'єктів на дорозі.

При виконанні роботи були використані загально-логічні, теоретичні та емпіричні методи. Загально-логічні використовувались для проведення аналізу предметної області, теоретичні були використані під час аналізу нейронних мереж, а емпіричні було застосовано під час проведення експериментального дослідження.

Кваліфікаційна робота апробована на 27-му Міжнародному молодіжному форумі «Радіoeлектроніка і молодь у XXI ст.».

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ ТА АКТУАЛЬНІСТЬ ДОСЛІДЖЕННЯ

1.1 Аналіз предметної галузі

Автопілот – це система управління транспортним засобом, що може автоматично керувати його рухом. Однією з важливих складових автопілота є система розпізнавання об'єктів, яка забезпечує визначення різних об'єктів у навколишньому середовищі, що допомагає системі приймати рішення про керування транспортним засобом.

Методи машинного навчання використовуються для розпізнавання об'єктів у автопілоті, тому що вони можуть забезпечити високу точність і швидкість обробки даних. Зазвичай використовуються такі методи машинного навчання, як нейронні мережі, глибоке навчання та алгоритми класифікації.

Одним з прикладів використання методів машинного навчання для розпізнавання об'єктів у автопілоті є система визначення дорожньої ситуації. Наприклад, система може використовувати камери, щоб отримувати зображення навколишнього середовища та застосовувати алгоритми класифікації для визначення різних об'єктів, таких як автомобілі, пішоходи, знаки дорожнього руху тощо.

Іншим прикладом може бути система автоматичного паркування, яка використовує нейронні мережі для аналізу зображення та визначення розмірів парковки, орієнтації автомобіля та інших об'єктів, які можуть вплинути на процес паркування.

Загалом, використання методів машинного навчання для розпізнавання об'єктів у автопілоті допомагає покращити безпеку на дорозі та забезпечити більш ефективне управління транспортним засобом та підвищити рівень безпеки на дорогах.

Згідно з звітом, опублікованим дослідницьким центром Bernstein, до 2040 року загальна кількість автомобілів на дорогах майже подвоїться. За прогнозами, до 2040 року автомобілі досягнуть позначки два мільярди[2].

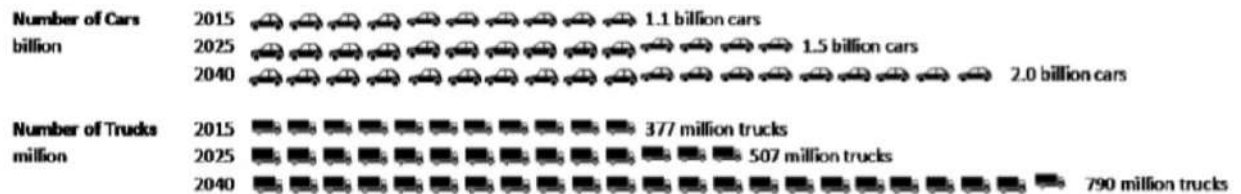


Рисунок 1.1 – Статистика зростання авто транспорту [2]

Людство пристосовується до всього та намагається зробити своє життя набагато комфортнішим та приємнішим. Людина витрачає багато часу на управління автомобілем і набагато ефективніше було б вирішити це. Це питання наразі і стоїть номером один у більшості компаній по конструюванні автомобілів: Google, Microsoft, Tesla, Mercedes, BMW, Bosch, Continental, Panasonic усі ці велетні ведуть титанічну роботу по розробці автопілотованого транспортного засобу. Про систему автономного управління автомобілем чути з усіх боків.

З моменту зародження автомобільної промисловості водії за законом повинні сидіти на передньому сидінні, поклавши руки на кермо. Наприклад, з 1968 року по всій Європі законодавство вимагає від автомобілістів тримати руки на кермі під час водіння, за винятком рухів, які потребують перемикання передач. Але незабаром це зміниться з появою безпілотних автомобілів. У найближчому майбутньому людина під час їзди на своєму авто може легально сісти на заднє сидіння, увімкнувши автопілот та снідати, доробляти щомісячний звіт, читати книгу, дивитися фільми або будь-яку іншу справу необхідну для неї, а водіння перейде на відповідальність автопілота – роботизованої системи управління транспортним засобом.

Але справжня перспектива автономних автомобілів полягає в потенціалі різкого зниження викидів CO₂. У недавньому дослідженні експерти ITDP визначили три тенденції, які, якщо їх застосувати одночасно, розкриють весь потенціал автономних автомобілів: автоматизація транспортних засобів, електрифікація транспортних засобів і спільне використання автомобілів (рис. 2). До 2050 року ці три революції можуть:

- зменшити затори (на 30% менше транспортних засобів на дорозі);

- скоротити транспортні витрати на 40% (щодо транспортних засобів, палива та інфраструктури);
- поліпшити прохідність і зручність життя;
- звільнити паркувальні місця для інших цілей (школи, парки, громадські центри);
- зменшити міські викиди CO₂ на 80% у всьому світі [3].

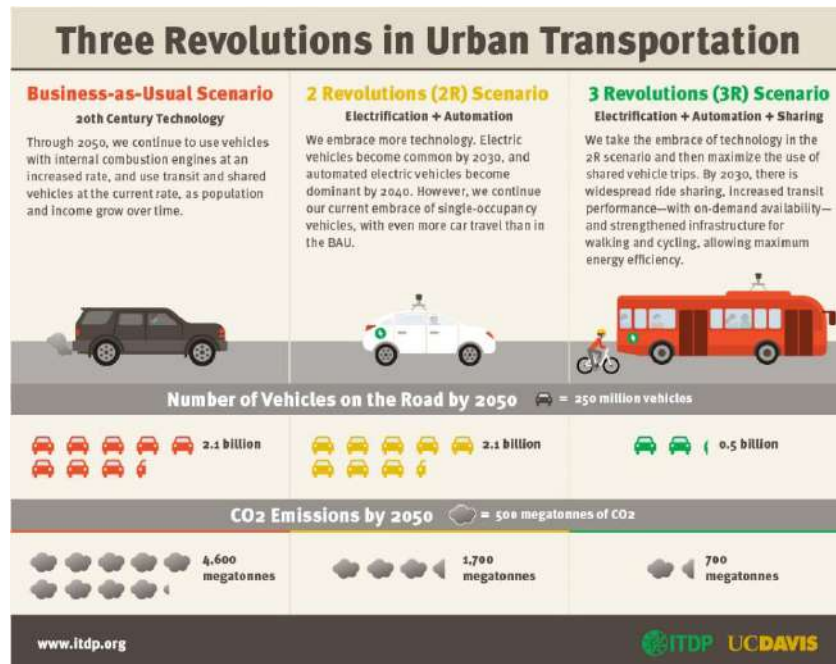


Рисунок 1.2 – Дослідження ITDP [3]

Робот може їздити набагато краще ніж людина. Він не має проблем зі здоров'ям, не вживає алкоголь, не має прогалин у навичках водіння, не відволікається на різноманітні фактори та багато іншого. Людський фактор має колосальний вплив при аварійних ситуаціях. Тому контроль за автомобілем у вигляді автопілота дуже допоможе людству у сучасному світі.

Автопілоти, розроблені на основі штучного інтелекту (ШІ), можуть бути програмовані для виявлення ризикових ситуацій на дорозі, таких як можливість зіткнення, перестроєння, порушення правил дорожнього руху тощо, і реагувати на них набагато швидше, ніж водій. Вони також можуть бути покращені безпечними засобами спілкування між автомобілями, такими як V2V (vehicle-to-vehicle) і V2X

(vehicle-to-everything) комунікації, що дозволяє обмінюватися даними про швидкість, напрямок руху та іншу інформацію, що допомагає уникнути аварій.

У четвертому кварталі 2018 року Tesla повідомила про одну аварію на кожні 2,91 мільйона миль, проїханих з увімкненим автопілотом. А при кермуванні людиною – одна аварія на кожні 1,76 мільйона миль у четвертому кварталі 2018 року та одна на кожні 1,58 мільйона миль у першому кварталі 2019 року. З цього квартального звіту компанії Tesla ми бачимо, що автопілот справляється набагато краще, ніж людина за кермом [4].

Автономні автомобілі покладаються на датчики, виконавчі механізми, складні алгоритми, системи машинного навчання та потужні процесори для виконання програмного забезпечення.

Вони оснащені різними сенсорами, камерами та датчиками. Наприклад, існують ультразвукові та інфрачервоні датчики. Деякі моделі навіть оснащені лазерними датчиками.



Рисунок 1.3 – Візуальна робота датчиків [Рисунок виконаний самостійно*]

За допомогою датчиків і камер машина сканує простір навколо себе кожні кілька мілісекунд. Системи сканування навколишнього середовища можуть

визначати навколишній ландшафт, дорожню розмітку, знаки та інші транспортні засоби у будь-якому місці.

Нові інновації також покращують безпеку автономних транспортних засобів. Наприклад, використання лідарних лазерів і допоміжних технологій водіння покращує можливості систем самостійного керування. Нові програми оцінювання таких агенцій, як Національне управління безпеки дорожнього руху, також стосуються безпеки автономних транспортних засобів, щоб зробити ці нові транспортні засоби безпечнішими на дорозі.

Автономні автомобілі створюють і підтримують карту свого оточення на основі різноманітних датчиків, розташованих у різних частинах автомобіля. Радіолокаційні датчики контролюють положення транспортних засобів, що знаходяться поблизу. Відеокамери виявляють світлофори, зчитують дорожні знаки, відстежують інші транспортні засоби та шукають пішоходів.

Датчики Lidar (виявлення світла та визначення дальності) відбивають імпульси світла від оточення автомобіля, щоб вимірювати відстані, виявляти краю доріг та ідентифікувати розмітку смуг. Ультразвукові датчики в колесах виявляють бордюри та інші транспортні засоби під час паркування.

1.2 Аналіз аналогів та актуалізація рішень

Автомобільна ніша є одна з найперспективніших та найпотрібніших у наш час, тому вона дає велику кількість проблем, які необхідно вирішувати.

Використання машинного навчання у автопілотах є дуже актуальним і постійно розвивається. Останні роки були свідками значного прогресу в цій галузі, зокрема завдяки зростанню доступності високопотужних обчислювальних ресурсів та розвитку нових алгоритмів машинного навчання.

Однією з найбільш важливих областей застосування машинного навчання у автопілоті є розпізнавання об'єктів, таких як інші автомобілі, пішоходи, дорожні знаки та інші перешкоди. Застосування глибоких нейронних мереж для розпізнавання об'єктів у реальному часі стає все більш поширеним. Іншою

важливою областю є розпізнавання голосу та розуміння мовлення, що дозволяє водієві взаємодіяти з автопілотом та керувати транспортним засобом за допомогою голосових команд.

Також у сучасних автопілотах все частіше використовуються методи машинного навчання для прогнозування поведінки інших водіїв та передбачення ризиків на дорозі. Це може допомогти попередити можливі аварійні ситуації та зменшити кількість дорожньо-транспортних пригод. Нарешті, існують також багато інших варіантів застосування машинного навчання в автопілоті, включаючи використання навчання з підкріпленням для оптимізації руху по маршруту, аналіз даних датчиків для виявлення невідповідностей в роботі системи та багато інших.

Однією з головних тенденцій використання машинного навчання у автопілоті є збільшення кількості датчиків та їх точності. Наприклад, розробники можуть використовувати лідари, радари та камери високої роздільної здатності для збору даних про навколишнє середовище та об'єкти на дорозі. Ці дані можуть бути використані для навчання моделей машинного навчання, які можуть розпізнавати об'єкти на дорозі, передбачати їх рух та робити рішення про керування транспортним засобом.

Провівши аналіз ніші, можна виокремити гіганта автомобільної індустрії, який працює над створенням автомобілів майбутнього. Компанія Tesla витрачає колосальні зусилля на створення повного автопілотованого автомобіля.

Tesla Autopilot — це набір передових функцій системи допомоги водієві (ADAS), запропонований Tesla, що відповідає стандарту SAE International Level 2 автоматизації транспортних засобів [5]. Його функції включають центрування смуги руху, зупинка на червоний, круїз-контроль з урахуванням дорожнього руху, автоматичну зміну смуги руху, напівавтономну навігацію на автострадах з обмеженим доступом, самостійне паркування та можливість викликати автомобіль із гаража чи паркувального місця. Кожен автомобіль як незалежна одиниця робить аналіз середовища навколо себе та аналізує об'єкти довкола.



Рисунок 1.4 – Розпізнавання об'єктів автопілотом [5]

Будь-яка з моделей Tesla на якій встановлено автопілот може триматися в межах своєї смуги, пересуваючись автономно по дорозі, уникати зіткнення за допомогою гальмування та руління, знайти місце для паркування та припаркуватися, регулювати свою швидкість залежно від інтенсивності руху, у тому числі до зупинки транспортного засобу і перебудовуватись на вимогу водія, достатньо включити поворотник. Унікальність Тесли полягає в тому, що вона вже може виконувати весь цей набір функцій, тоді як деякі автомобілі на ринку можуть виконувати лише окремі речі зі списку.

Ще одна по-справжньому розумна річ у системі автопілота є його функція навчання. Кожна модель, оснащена системою автопілотування, підтримує зворотний зв'язок у режимі реального часу з хмарою Тесла, завдяки цьому компанія постійно вдосконалює систему та покращує поведінку кожної машини. Інші автовиробники не мають цієї важливої ланки у своїх авто з напіваавтономними можливостями.

Контроль світлофора та стоп-сигналу виявляє світлофори або знаки, які сигналізують про зупинку та автоматично сповільнює автомобіль до зупинки під час наближення до них за допомогою автопілота під активним наглядом.

Коли ввімкнено керування світлофором і стоп-сигналом візуалізація водіння відображає майбутні світлофори, знаки зупинки або дорожню розмітку на перехрестях, де вашому автомобілю може знадобитися зупинитися. Коли ви наближаєтеся до перехрестя, навіть якщо світлофор горить зеленим світлом, на вашому автомобілі відобразатиметься червона лінія, яка вказуватиме, де транспортний засіб зупиниться, і транспортний засіб почне сповільнюватися.

Сповіщення на сенсорному екрані повідомлять про причину зупинки (знак «стоп» або світлофор), а також повідомлять приблизну відстань до моменту, коли ваш автомобіль зупиниться. Якщо сповіщення «Світлофор» не вказує світлий колір, автомобіль не підтвердив стан контролю дорожнього руху.

Автопілот дозволяє керувати, прискорюватися та гальмувати в межах смуги руху. Tesla має 12 ультразвукових датчиків далекого радіусу дії, розташованих навколо автомобіля для того, щоб «зондувати» те, що відбувається навколо, в діапазоні до 4,8 метрів у будь-якому напрямку та за будь-якої швидкості. Модель також має передній радар, фронтальну камеру, високоточний GPS, і систему, яка обробляє всі дані із зазначених пристроїв для формування віртуального образу дороги та обстановки попереду. Так як у автомобіля повністю цифрова комбінація приладів, остання здатна відображати те, що автомобіль "бачить" в режимі реального часу [6].



Рисунок 1.5 – Розміщення сенсорів Tesla [Рисунок виконаний самостійно*]

Завдяки функції самонавчання, кожен автомобіль Tesla з системою автопілота повідомляє автовиробнику в режимі реального часу про проблеми, це дає можливість виправляти всі помилки, що виникають в найкоротший термін і навіть превентивно видаляти потенційні помилки, перш, ніж інші автомобілі і користувачі відчують проблеми.

Нейронні мережі дивляться ці відео з усіх камер автомобіля, обробляють їх та роблять прогнози щодо того, що вони бачать. Нас цікавить дорожня розмітка, учасники руху, інші об'єкти та відстані до них, проїжджа частина, світлофори, дорожні знаки тощо.

Основне завдання, яке мережі вирішують в автомобілі, це розпізнавання образів. Для нас, людей, це дуже просте завдання. Ви дивитеся на зображення і бачите віолончель, човен, ігуану чи ножиці. Дуже легко та просто для вас, але не для комп'ютера. Причина в тому, що ці зображення для комп'ютера є масивом пікселів, де кожен піксель це значення яскравості в цій точці. Замість того, щоб просто бачити зображення, комп'ютер отримує мільйон чисел у масиві.

Нейронна мережа не розуміє, що зображено на фото або відео з камери. Для неї це лише набір відтінків від білого до чорного, які зберігаються у вигляді градієнтів кольорів. І щоб ідентифікувати об'єкт відео, необхідно обробити мільйони пікселів, розрізняючи їх за градієнтом кольору, і запам'ятати правильно пізнані поєднання. Нейронним мережам потрібна пам'ять, щоб зберігати вхідні дані, вагові параметри та функції активації. Тому потреби в обчислювальних ресурсах цих систем дуже великі.

Останні п'ять років, сучасні підходи до обробки зображень з використанням комп'ютерів також почали використовувати нейронні мережі, але в даному випадку - штучні нейронні мережі. Штучні нейронні мережі є грубе математичне наближення зорової кори. Тут також є нейрони, вони пов'язані один з одним. Типова нейронна мережа включає десятки чи сотні мільйонів нейронів, і кожен нейрон має тисячі зв'язків.

Взявши нейронну мережу і показавши їй зображення, наприклад, ігуану, мережа зробить прогноз про те, що вона бачить. На початку ініціалізації нейронних мереж всі зв'язки між нейронами та їх ваги встановлюються випадковим чином. Це призводить до того, що передбачення, зроблені мережею, також є випадковими. Може виявитися, що мережа думає, що це лев, але під час тренування відзначається, що на зображенні ігуана. Потім використовується математичний процес, який називається методом зворотного поширення помилки. Відновивши вагу кожного з цих з'єднань, як тільки оновлення буде завершено, ймовірність ігуани для цього зображення зростає, а ймовірність інших варіацій розпізнавання зменшиться.

Звичайно, робиться це не з єдиним зображенням. Необхідно мати великий набір розмічених даних. Зазвичай, це мільйони зображень, тисячі міток або близько того. Процес навчання повторюється знову і знову. Показуючи комп'ютеру зображення, він повідомляє свою думку, далі необхідно відкоригувати її, щоб мережа переналаштувалась.

Навчання зазвичай займає кілька годин чи кілька днів. Для успішного навчання потрібно дуже багато вхідної інформації. Для прикладу, взявши якусь

довільну собаку, людина дивлячись на цю картинку бачить тварину певної породи. Якщо переглянути ще кілька зображень інших собак, то можливо знайти серед них інших тварин такої ж породи. Для цього знадобився лише один приклад, але машини так не можуть. Для цього необхідна купа даних про собак та їх породи, тисячі прикладів, у різних позах, різних умовах освітлення, на різному фоні тощо. Без цих всіх даних комп'ютер не зможе вивчити потрібний шаблон та виявити необхідне.

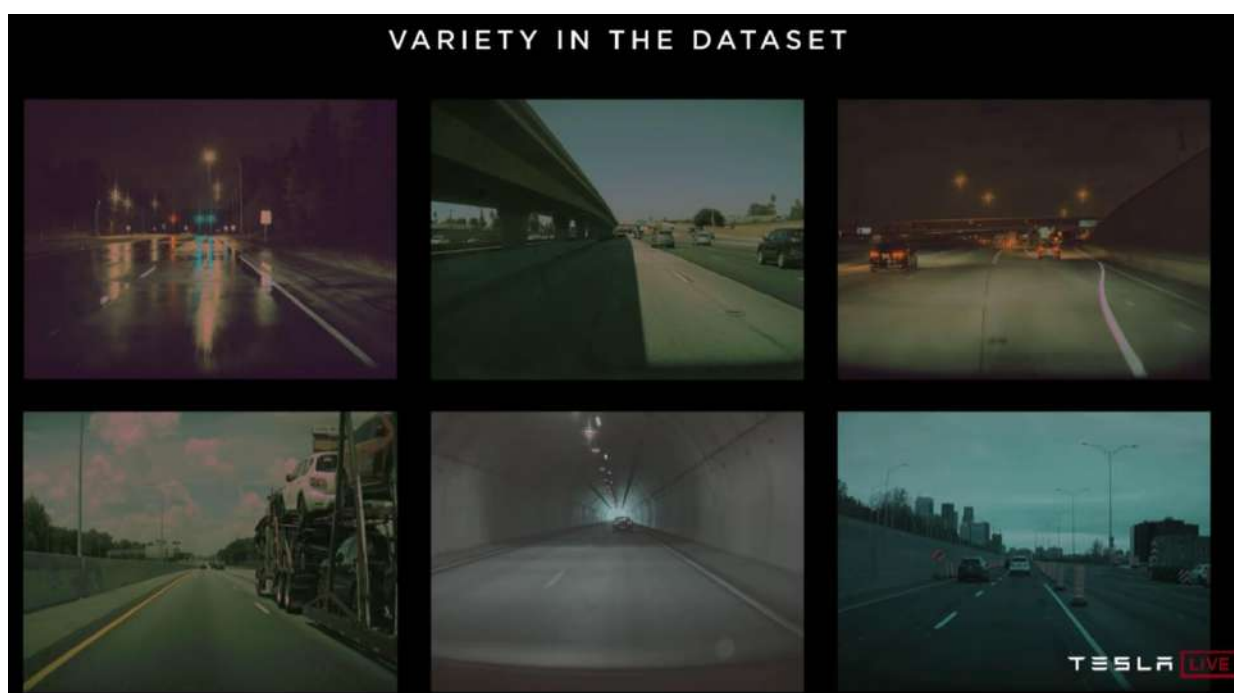


Рисунок 1.6 – Приклад набору даних для навчання автопілоту [5]

Ряд фахівців зазначають, що можливо б використовувати симуляцію, щоб отримати необхідні дані у потрібному масштабі. Вони використовувались для навчання досить успішно. Але в результаті, коли справа доходить до навчальних даних для нейронних мереж, ніщо не замінить реальних даних. Симуляції мають проблеми з моделюванням зовнішнього вигляду, фізики та поведінки учасників.

Реальний світ підкидає купу несподіваних ситуацій. Складні умови із снігом, деревами, вітром. Різні візуальні артефакти, які важко змоделювати. Зони ремонту дороги, кущі, поліетиленові пакети, що бовтаються по вітру. Там може бути багато людей, дорослі, діти та тварини упереміш. Моделювання поведінки та взаємодії

проблема не могла повторитися пізніше. Тепер, щоб виправити проблему, потрібно багато навчальних прикладів.

Необхідно відправити в систему більше зображень тунелів, правильно їх розмітити, додати в навчальну вибірку та перенавчити мережу, а потім завантажити в автомобілі. Цей цикл повторюється знову і знову. Цей ітеративний процес називається data engine. Включаючи мережу в тіншовому, виявляються неточності, додається більше даних та включаються в навчальну вибірку.

Планування траєкторії для автопілота робототехнічного пристрою схоже із завданням програмного помічника слабозорого: необхідно визначати місце розташування людини/пристрою та оцінювати навколишнє середовище, тобто вирішувати завдання одночасної навігації та побудови карти.

При цьому мається на увазі, що про середовище немає жодної валідної інформації. Можна використовувати лише бортові прилади, які фіксують зміни у навколишньому просторі. Для того щоб переміщатися з початкового положення в задане, необхідно створити і постійно підтримувати в актуальному стані карту місцевості та отримувати точну оцінку траєкторії свого руху.

Необхідно отримати дані з нейронної мережі і створити на їх основі семантичну мережу навколишнього простору, на якій кожен об'єкт отримує відповідні смислові мітки, переставши бути просто набором кольірних градієнтів. Наприклад, нейронна мережа автопілота Tesla вчиться розрізняти автомобіль, що йде попереду, з велосипедом, закріпленим на задньому бампері, приймаючи спочатку один об'єкт за два - власне автомобіль і велосипед.

Якщо дані організовані в семантичну мережу, значення навантаження редукується з двох об'єктів в один і, отже, в режимі експлуатації не потрібне додаткове впізнавання. Найважливіше тут саме редукування. Внаслідок такого редукування відпадає необхідність отримання та обробки детального зображення. Фіксуються лише основні ознаки.

1.3 Актуальність та мета дослідження

Актуальність дослідження методів машинного навчання для розпізнавання об'єктів полягає в широкому спектрі застосувань цих методів в різних галузях, таких як комп'ютерний зір, автономні автомобілі, відеоспостереження, розпізнавання образів, медицина, безпека, ігрова індустрія та багато інших.

Метою дослідження методів машинного навчання для розпізнавання об'єктів та прогнозування ймовірності та напрямку їх руху є створення ефективної та безпечної системи автопілоту для автомобілів, яка забезпечує автоматичне керування транспортним засобом на дорозі.

Застосування методів машинного навчання для розпізнавання об'єктів може вдосконалити процеси автоматизації, підвищити рівень безпеки, забезпечити ефективність виробничих процесів та покращити якість обслуговування.

У сучасних умовах компанії по впровадженню систем автопілотування у автомобілі зустрічаються з низкою проблем. Методи машинного навчання можуть бути не ідеальними, що може становити ризик для пішоходів та інших учасників дорожнього руху. Можуть бути проблеми з точністю, оскільки алгоритм може спиратися на попередній досвід. Це може бути проблемою в ситуаціях, коли автопілот повинен приймати швидкі рішення.

Дослідження машинного навчання для автономного водіння може бути дуже дорогим, оскільки необхідно розробляти, тестувати та налагоджувати різні моделі та алгоритми, щоб відповідати вимогам безпеки та надійності. У більшості випадків для навчання автопілота необхідно використовувати великі обсяги даних. Точність і надійність автономного керування може бути проблематичною, якщо вхідних даних мало.

Дослідження полягає у аналізі актуальних проблем, порівнянні методів нейронної мережі для розпізнавання об'єктів та дослідженні і порівнянні оптимальних варіантів для застосування у роботизованому автомобілі.

За допомогою методів машинного навчання можна створити моделі, які будуть розпізнавати об'єкти на дорозі та прогнозувати їх дії. Ці моделі можуть бути

навчені на великій кількості даних (зображень), що зібрані з датчиків, таких як лідари, радары та камери.

Це можна досягти завдяки використанню різноманітних методів машинного навчання. Використання цих методів дозволяє побудувати моделі, які можуть розпізнавати об'єкти на дорозі, передбачати їх рух та прогнозувати їх поведінку в майбутньому. Такі моделі можуть використовуватися для прийняття рішень про керування автомобілем, зокрема для уникнення аварій та забезпечення безпеки пасажирів. Ці методи можуть використовуватися для розпізнавання об'єктів, таких як інші транспортні засоби, пішоходи, велосипедисти, знаки дорожнього руху та інші елементи інфраструктури на дорозі.

Крім того, дослідження методів машинного навчання дозволяє покращити точність та швидкість роботи систем автопілоту, а також забезпечити їх адаптацію до різних умов дорожнього руху та змінюваних параметрів дороги. Отже, мета дослідження методів машинного навчання для розпізнавання об'єктів та прогнозування їх руху полягає в покращенні ефективності та безпеки систем автопілоту для автомобілів, що робить їх більш придатними для широкого використання на дорогах.

1.4 Постановка задачі

Мета дослідження – вивчення сучасних підходів машинного навчання для розпізнавання об'єктів та прогнозування їх руху. У наш час застосування нейронних мереж для цієї задачі є найактуальнішим та найперспективнішим серед усіх розробників новітніх авто.

Таким чином, мета дослідження полягає у отриманні знань та аналізу можливостей методів машинного навчання для розпізнавання об'єктів та прогнозування їх руху у режимі реального часу за допомогою різних типів нейронних мереж. Для роботи були обрані методи на основі CNN та NAS розглядаючи їх перспективність та актуалізацію у нинішній час.

Для проведення дослідження було використано набори даних основних дорожніх знаків з різних джерел, таких як зображення зняті з відеореєстраторів та довільні фото дорожніх знаків, які можуть попадати у поле зору системи розпізнавання об'єктів у автомобілі.

У результаті проведеного дослідження порівняно точності моделей машинного навчання для розпізнавання об'єктів та прогнозування ймовірності та напрямку їх руху у режимі реального часу використовуючи навчені нейронні мережі. Цей критерій є одним з найважливіших при виборі моделі для використання у автопілотованій системі.

2 АНАЛІЗ МЕТОДІВ ДЛЯ ДОСЛІДЖЕННЯ

2.1 Обґрунтування методів досліджень та етапи його проведення

Аналіз — це процес детального вивчення чогось, щоб зрозуміти його компоненти, зв'язки та функції. Це може включати дослідження, аналіз інформації, статистичний аналіз даних, опис процесів або явищ, а також визначення моделей і тенденцій.

Наукове дослідження — це систематичний процес збору та аналізу інформації для отримання нових знань, відкриття нових фактів або перевірки гіпотез. Наукове дослідження передбачає використання специфічних методів, які використовують емпіричні дані для організації інформації та перевірки гіпотез.

Наукові дослідження можна проводити в будь-якій галузі знань, від фізики та хімії до соціології та економіки. Вони є основним джерелом нових винаходів та інновацій, які допомагають вирішити проблеми, що стосуються суспільства та людства.

Наукові дослідження можуть бути теоретичними, спрямованими на розробку нових теорій і припущень, або прикладними, спрямованими на вирішення практичних завдань і поліпшення життя людей. Наукове дослідження, незалежно від його мети та масштабу, є важливим інструментом для розвитку знань і розвитку світу.

Наукові дослідження завжди пов'язані з вирішенням наукових проблем. Брак знань, фактів і суперечливі наукові думки є основою наукових досліджень. Добре, що хтось більш компетентний і вивчає питання, пов'язані з його практикою (спорт, освіта, організація, навчання чи виховання дітей). При цьому необхідно вивчати тестоздатність обраних суб'єктів.

Є достатня кількість досліджуваних для тестування експериментальних груп (експериментальна та контрольна групи), розробки приладів, правильне обладнання. Виховний процес експериментальної групи та ін. Для проведення наукового дослідження з обраної теми можна охарактеризувати декілька етапів, які можна узагальнити наступним чином:

Наукові дослідження за обраними темами можна розділити на кілька етапів, які можна підсумувати таким чином.

Постановка проблеми та формулювання питання дослідження. На цьому етапі аналізують розв'язувану проблему та визначають конкретні завдання дослідження.

Збір даних. У наукових дослідженнях важливо зібрати достатньо даних для створення та тестування моделей машинного навчання.

Попередній аналіз даних. На цьому кроці виконується попередній аналіз даних, наприклад візуалізація та статистичний аналіз.

Розробка моделей машинного навчання. Використовуючи зібрані дані, розробіть модель машинного навчання для вирішення проблем.

Навчання та перевірка моделі. Після того, як модель розроблена, її навчають на зібраних даних і перевіряють на тестових даних, щоб переконатися в її точності та продуктивності.

Оцінка та інтерпретація результатів. Після навчання та перевірки моделі оцініть результати та інтерпретуйте дані.

Вказує на результат. Останнім кроком є написання наукової статті та публікація результатів дослідження для подальшого вивчення та використання іншими науковцями.

Наукове дослідження — це самостійний процес, пов'язаний із пошуком нових знань і розв'язанням деяких проблем, які мають велике значення для людини. Ці дослідження включають збір і аналіз даних, огляд літератури та інформації з певної теми, перевірку гіпотез і проведення експериментів, а також проведення інтерв'ю з людьми.

2.2 Огляд та вибір датасету

Обрання правильного датасету є одним з найважливіших кроків у процесі навчання нейронної мережі. Вибір відповідного датасету може впливати на якість та точність навчаної моделі. Проте, на практиці можуть зустрічатися датасети з

неправильною або неповною інформацією, тому для підготовки до моделювання потрібно витратити багато часу на проведення операцій, таких як очищення, нормалізація та генерація змінних.

Набір даних має представляти завдання, які має виконувати нейронна мережу. Він повинен включати різні дані, які можуть зустрічатися в реальному світі, щоб модель могла вчитися на різних вхідних даних. Моделі можуть бути надмірно або недостатньо навчені через набір даних.

Розмір набору даних також є важливим фактором. Великі набори даних допоможуть моделям краще узагальнювати та знизити ризик перенавчання. У той же час надто маленький набір даних призведе до неефективного навчання моделі та низької точності.

Мітки в датасеті повинні бути якісними та достовірними. Якість мітки впливає на якість навчання моделі, тому важливі її точність і достовірність. Неправильний сигнал може призвести до неправильного навчання моделі та зниження її точності.

Якщо модель виконує завдання класифікації чи сегментації, важливо, щоб база даних мала баланс класів, тобто приблизно однакова кількість прикладів кожної категорії. Незбалансовані набори даних можуть призвести до упередженого навчання моделі та недостатньої точності у недостатньо представлених класах.

Датасет повинен містити різноманітність даних, таких як різні стилі, розміри, освітлення, пози, аугментації тощо. Це допоможе моделі навчитися адаптуватися до різних варіацій даних в реальному світі та забезпечити кращу навчаність моделі на нових, реальних даних.

Датасет повинен бути актуальним та відповідати сучасним реаліям задачі. Використання застарілих даних може призвести до погіршення точності моделі на нових даних, що можуть бути різні від даних в датасеті.

Важливо враховувати безпеку даних при виборі датасету. Датасет повинен бути зібраний відповідно до відповідних правил та законів, не містити конфіденційної, приватної або неприпустимої інформації. Дотримуючись рекомендацій при обранні датасету для навчання нейронної мережі може

допомогти покращити якість та точність моделі, що в свою чергу може мати великий вплив на результати та застосування моделі в реальному світі.

2.3 Огляд обраного інструменту для дослідження

AutoML Vision – це інструмент, розроблений компанією Google, який дозволяє автоматизувати процес створення моделей машинного навчання для задач комп'ютерного зору[7].

Завдяки AutoML Vision, можливо навчити систему розпізнавати об'єкти на зображеннях, класифікувати їх за категоріями та визначати різні параметри, такі як розмір та форма об'єктів. Крім того, інструмент може підібрати найкращу архітектуру моделі для конкретної задачі.

Система використовує алгоритм нейронно-архітектурного пошуку Neural Architecture Search, який дозволяє знаходити оптимальну нейронну мережу для конкретної задачі. AutoML Vision надає доступ до різноманітних інструментів, що дозволяють швидко та якісно навчати нейронні мережі.

Neural Architecture Search (NAS) - це процес автоматичного пошуку оптимальної архітектури нейронної мережі для конкретної задачі[8].

Зазвичай, для створення нейронних мереж необхідно вручну визначати їхню архітектуру, яка складається з кількох шарів та вузлів, що може бути важким та часовим затратним завданням. Застосування NAS дозволяє автоматизувати цей процес, визначаючи оптимальну архітектуру нейронної мережі за допомогою алгоритмів машинного навчання.

Алгоритми NAS базуються на принципі пошуку оптимальної архітектури шляхом знаходження набору параметрів та змінних, що відповідають заданій метриці якості. Ці параметри можуть включати кількість та розмір шарів, типи функцій активації, коефіцієнти регуляризації та інші.

Пошук нейронної архітектури — це завдання автоматичного пошуку однієї чи кількох архітектур для нейронної мережі, яка відносно швидко дасть моделі з хорошими результатами для даного набору даних. Пошук нейронної архітектури

зараз є новою областю. Проводиться багато досліджень, існує багато різних підходів до завдання, і немає єдиного найкращого методу в цілому — або навіть єдиного найкращого методу для спеціалізованих проблем, таких як ідентифікація об'єктів на зображеннях.

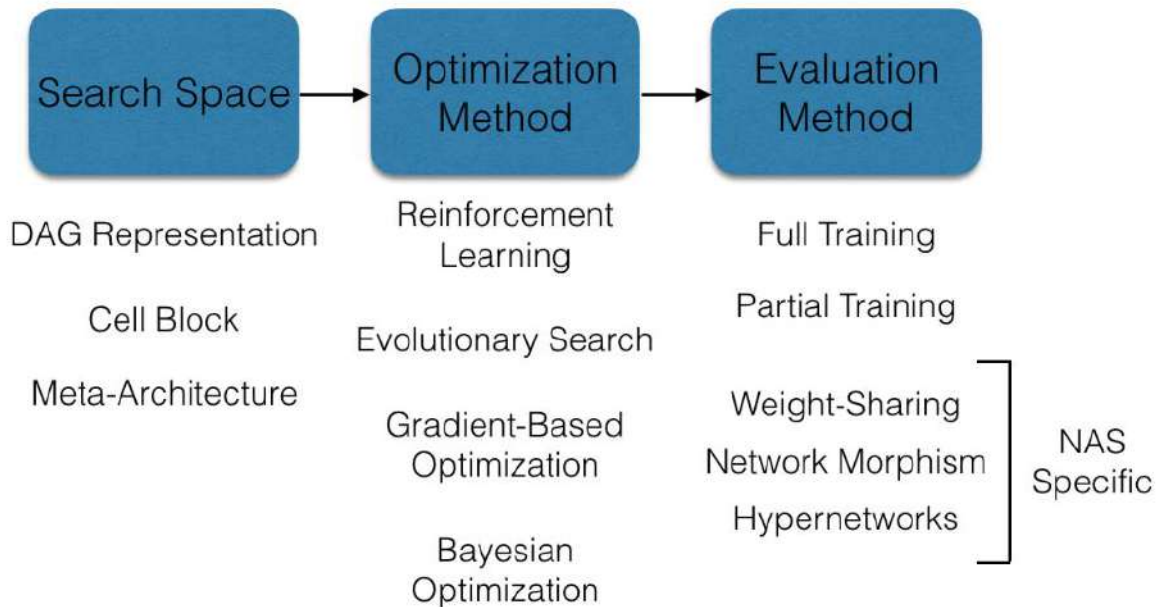


Рисунок 2.1 – Компоненти NAS [8]

Застосування NAS дозволяє зменшити затрати часу та зусиль на проектування нейронних мереж, підвищити їхню точність та зменшити ризик людської помилки при визначенні їх архітектури. NAS використовується в багатьох галузях, де використовуються нейронні мережі, таких як обробка зображень, обробка мовлення, розпізнавання мови та інші [9].

Найпростіший спосіб оцінки продуктивності нейронних мереж — це навчання та перевірка мереж на основі даних. На жаль, це може призвести до обчислювальних вимог порядку тисяч GPU-днів для пошуку нейронної архітектури. Способи скорочення обчислень включають нижчі оцінки точності (менше епох навчання, менше даних і зменшені моделі); екстраполяція кривої навчання (на основі лише кількох епох); тренування з розминки (ініціалізація ваг, копіювання їх із батьківської моделі); і одноразові моделі з розподілом ваги (підграфи використовують вагові коефіцієнти одновибіркової моделі). Усі ці

методи можуть скоротити час навчання до кількох GPU-днів, а не до кількох тисяч GPU-днів. Однак зміщення, внесені цими наближеннями, ще недостатньо зрозумілі.

Google AutoML Vision є гарною системою яка має цілий ряд переваг:

- автоматична оптимізація: AutoML Vision здатний автоматично оптимізувати модель машинного навчання, що дозволяє досягти кращих результатів в короткий проміжок часу;
- розширена функціональність: AutoML Vision має декілька інструментів, що дозволяють користувачам проводити аналіз, візуалізацію даних та тестування моделей машинного навчання;
- швидкість навчання: AutoML Vision може працювати з великими обсягами даних та швидко навчати моделі машинного навчання;
- висока точність: AutoML Vision використовує найновіші алгоритми машинного навчання та може досягати високої точності при класифікації.

Google AutoML Vision використовує глибоке навчання (deep learning [10]) для створення та навчання моделей машинного навчання. Глибоке навчання базується на нейронних мережах, які здатні розпізнавати складні шаблони у вхідних даних та автоматично навчатися відновлювати ці шаблони. Вона використовує алгоритми нейронної мережі, щоб автоматично вивчити приклади даних та знаходити закономірності, які можуть допомогти розв'язати задачі комп'ютерного зору.

Машинне навчання відбувається в два етапи: ми навчаємо модель, а потім просимо модель робити прогнози. Навчання — це процес, коли комп'ютер переглядає всі наші дані, щоб з'ясувати зв'язок між усіма наявними у нас значеннями, і у великих наборах даних це може зайняти багато часу — години, а потенційно — набагато довше. Прогнозування виконується на пристрої: ми передаємо йому навчену модель, і він використовуватиме попередні результати, щоб оцінити нові дані.

Навчання з вчителем (supervised learning) - це тип машинного навчання, в якому модель навчається на основі вхідних даних та відповідних їм вихідних значень. Наприклад, можна використовувати навчання з вчителем для

розпізнавання зображень, де вхідні дані - це зображення, а вихідні значення - це мітки, які вказують на те, що зображено на зображенні.

Навчання без вчителя (unsupervised learning) - це тип машинного навчання, в якому модель навчається на основі вхідних даних без відповідних їм вихідних значень. Наприклад, можна використовувати навчання без вчителя для кластеризації даних, де модель групує вхідні дані в різні кластери на основі спільності їхніх характеристик.

Підсилений (reinforcement learning) - це тип машинного навчання, в якому модель навчається на основі взаємодії з динамічним середовищем. Наприклад, можна використовувати підсилений навчання для навчання робота, де модель вчиться діяти в різних ситуаціях на основі відмінностей у нагородах та штрафів, які отримує робот за свої дії.

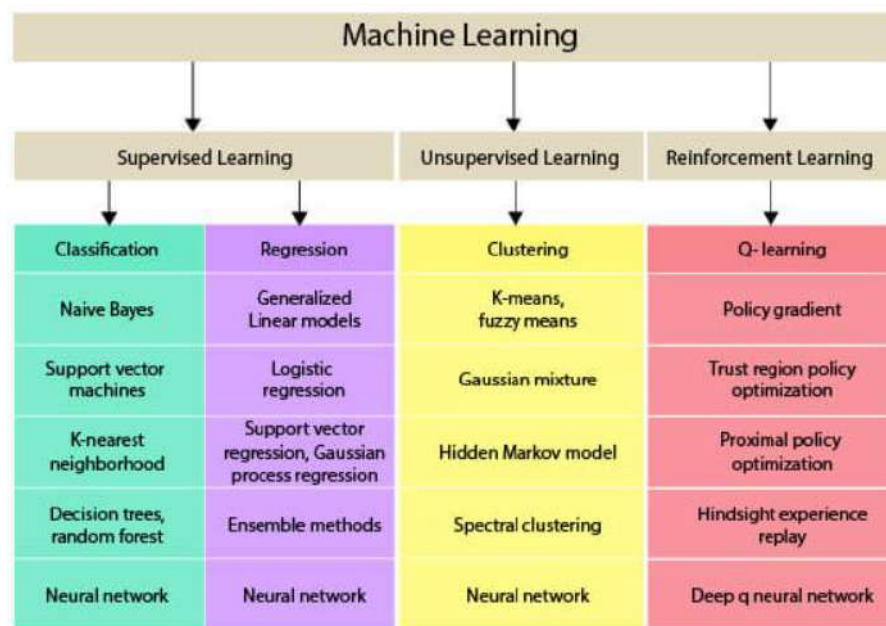


Рисунок 2.2 – Основні типи машинного навчання [Рисунок виконаний самостійно*]

Загалом, кожен тип машинного навчання має свої виклики та може використовуватися для різних задач. Для системи розпізнавання дорожніх знаків та ідентифікації пішоходів можна використовувати здебільшого навчання з

учителем (supervised learning), так як цей тип машинного навчання добре підходить для задач класифікації та регресії, що потрібні для автопілота.

Зазвичай, для навчання автопілота використовують нейронні мережі, зокрема глибокі нейронні мережі (deep neural networks). Вони можуть ефективно розпізнавати та аналізувати великі обсяги даних, що надходять від датчиків автомобіля та допомагають зробити рішення про поведінку автомобіля на дорозі.

2.4 Аналіз обраних методів для дослідження

Машинне навчання стає все більш популярним у сфері комп'ютерного зору та обробки зображень. Однією з найбільш захоплюючих областей є розпізнавання об'єктів [9]. Розпізнавання об'єктів є процесом визначення ідентичності або класифікації об'єктів на зображеннях або відео. Цей процес може бути використаний для автоматичної ідентифікації дорожніх знаків, визначення видів об'єктів, розпізнавання автомобілів на дорогах та багато іншого.

Це має вирішальне значення для багатьох застосувань, таких як автономні транспортні засоби, безпілотні літальні апарати та системи безпеки.

Системи автопілота в сучасному світі стають все популярнішими, а їх розвиток та вдосконалення полягає в основному в застосуванні методів машинного навчання. При виборі методу машинного навчання для систем автопілота слід керуватись вимогами конкретної системи.

Системи автопілота покладаються на алгоритми машинного навчання, щоб сприймати навколишнє середовище, приймати рішення та вживати відповідних дій. Для досягнення цих цілей використовуються різноманітні методи машинного навчання, такі як глибоке навчання та навчання з підкріпленням.

Штучні нейронні мережі — це мозок, який стоїть за деякими з найскладніших програм штучного інтелекту (ШІ). Але це не означає, що розуміння різних типів має бути складним. Що стосується штучних нейронних мереж обчислювальних систем, які імітують компоненти мозку, існують серйозні відмінності в типах. Розуміння цих відмінних форм, їхніх нюансів і різноманітних застосувань може

змінити успіх і провал у наступній ініціативі ШІ та машинного навчання. У машинному навчанні кожен тип штучної нейронної мережі пристосований для виконання певних наборів завдань.

Алгоритми глибокого навчання використовують штучні нейронні мережі для вивчення великих обсягів даних, що дозволяє системі розпізнавати шаблони та приймати рішення, тоді як навчання з підкріпленням використовує стратегію навчання, яка використовує позитивний і негативний зворотний зв'язок для посилення правильних дій. Глибоке навчання довело свою ефективність у розпізнаванні зображень і виявленні об'єктів, які є важливими для систем автопілота. Один з найбільш популярних методів машинного навчання, який використовується в системах автопілота, це глибоке навчання, зокрема з використанням згорткових нейронних мереж (CNN – Convolutional neural network).

CNN — це тип нейромережі, спеціально розроблений для обробки зображень та виявлення різноманітних візуальних ознак [11]. Він використовує оператори згортки та пулінгу для автоматичного виявлення різних рівнів абстракції в зображеннях. CNN може автоматично вивчати різноманітні характеристики, такі як краї, форми, текстури тощо з вхідних зображень. Він також може бути розширений до багатоканальних зображень, таких як RGB зображення, де можуть бути використані декілька шарів згортки та пулінгу для роботи з кожним кольоровим каналом окремо.

Цей тип глибокого навчання показав великі перспективи в системах автопілота та використовується у багатьох автомобільних нішах. CNN можуть розпізнавати особливості зображень і приймати рішення на основі цих характеристик. Їх використовували при розробці автономних транспортних засобів, дронів та інших безпілотних систем. Вони показали себе дуже ефективними для розпізнавання зображень та виявлення об'єктів, що робить їх добре придатними для використання в автономних транспортних засобах та інших подібних системах.

Архітектура аналогічна структурі підключення нейронів у людському мозку та була натхненна організацією зорової кори. Окремі нейрони реагують на стимули

лише в обмеженій області поля зору, відомої як рецептивне поле. Набір таких полів перекривається, щоб охопити всю візуальну область. CNN зазвичай має три рівні: згортковий рівень, рівень об'єднання та повністю зв'язаний рівень.

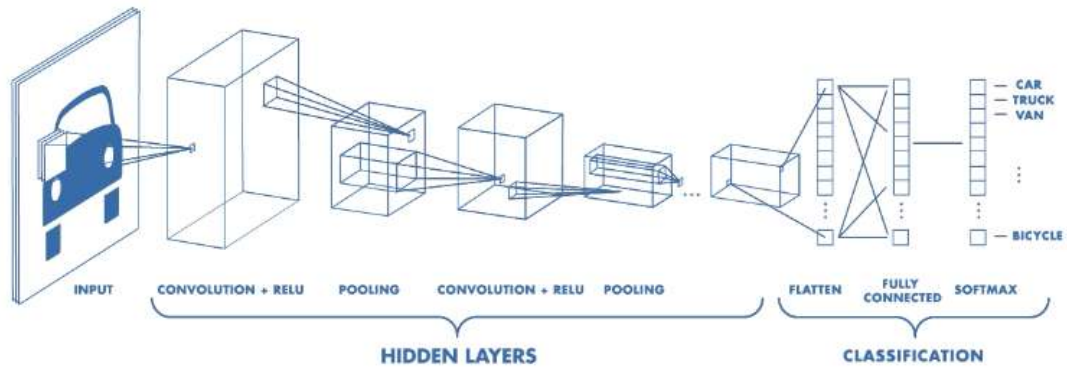


Рисунок 2.3 – Візуальне представлення архітектури Convolutional neural network [20]

Рівень згортки є основним будівельним блоком CNN. Він несе основну частину обчислювального навантаження мережі. Цей рівень виконує скалярний добуток між двома матрицями, де одна матриця є набором параметрів, які можна дізнатися, інакше відомих як ядро, а інша матриця є обмеженою частиною сприйнятливого поля. Ядро просторово менше, ніж зображення, але більш глибоке. Це означає, що якщо зображення складається з трьох (RGB) каналів, висота і ширина ядра будуть просторово малими, але глибина поширюється на всі три канали.

Під час прямого проходу ядро проходить по висоті та ширині зображення, створюючи представлення зображення цієї сприйнятливої області. Це створює двовимірне представлення зображення, відоме як карта активації, яка дає відповідь ядра на кожну просторову позицію зображення.

Тривіальні рівні нейронної мережі використовують множення матриці на матрицю параметрів, що описують взаємодію між блоком введення та виведення. Це означає, що кожен вихідний блок взаємодіє з кожним вхідним блоком. Однак

згорткові нейронні мережі мають розріджену взаємодію. Це досягається шляхом зменшення розміру ядра, наприклад, зображення може мати мільйони чи тисячі пікселів, але під час його обробки за допомогою ядра ми можемо виявити значущу інформацію, яка складається з десятків чи сотень пікселів. Це означає, що нам потрібно зберігати менше параметрів, що не тільки зменшує вимоги до пам'яті моделі, але й покращує статистичну ефективність моделі.

Рівень об'єднання замінює вихідні дані мережі в певних місцях шляхом отримання сумарної статистики найближчих виходів. Це допомагає зменшити просторовий розмір представлення, що зменшує необхідну кількість обчислень і ваги. Операція об'єднання обробляється для кожного фрагмента представлення окремо. Нейрони в цьому шарі мають повний зв'язок з усіма нейронами в попередньому та наступному шарі. Ось чому його можна обчислити як зазвичай шляхом множення матриці з наступним ефектом зміщення. Існує два типи об'єднання: максимальне об'єднання та середнє об'єднання. Max Pooling повертає максимальне значення з частини зображення, охопленої ядром. З іншого боку, Average Pooling повертає середнє значення всіх значень із частини зображення, охопленої ядром. Max Pooling також працює як шумозаглушник. Він повністю відкидає шумні активації, а також виконує усунення шумів разом із зменшенням розмірності. З іншого боку, Average Pooling просто виконує зменшення розмірності як механізм придушення шуму. Отже, ми можемо сказати, що Max Pooling працює набагато краще, ніж Average Pooling.

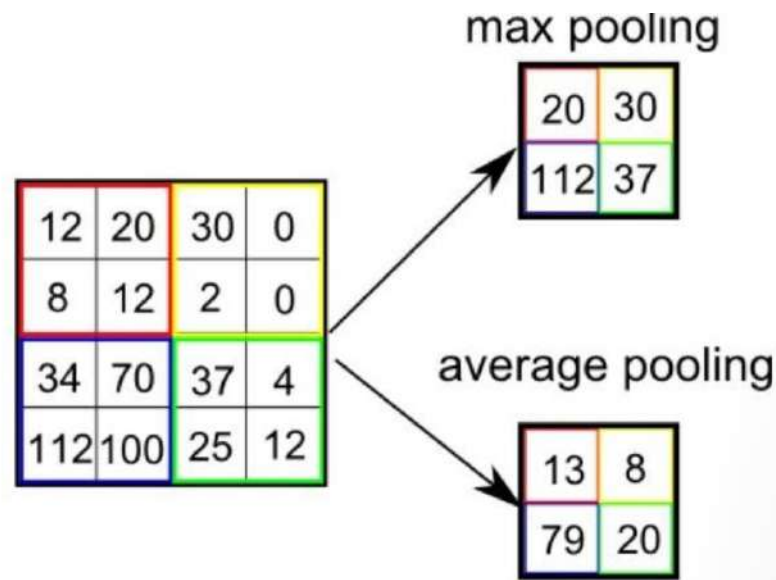


Рисунок 2.4 – Типи пулів [20]

Згортковий рівень і рівень об'єднання разом утворюють і-й рівень згорткової нейронної мережі. Залежно від складності зображень кількість таких шарів може бути збільшена для ще більшого захоплення деталей низького рівня, але ціною більшої обчислювальної потужності.

Розглянемо базову архітектуру CNN та створення її рівнів за допомогою мови програмування Python та використанням Python Keras:

```

from keras.models import Sequential
from keras.layers import Conv2D, MaxPooling2D, Flatten, Dense
model = Sequential()
model.add(Conv2D(32, (3, 3), activation='relu', input_shape=(64, 64, 3)))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dense(num_classes, activation='softmax'))
model.compile(loss='categorical_crossentropy', optimizer='adam',
metrics=['accuracy'])
model.summary()

```

Розглянемо приклад роботи моделі по розпізнаванню автомобіля на рисунку 2.5.

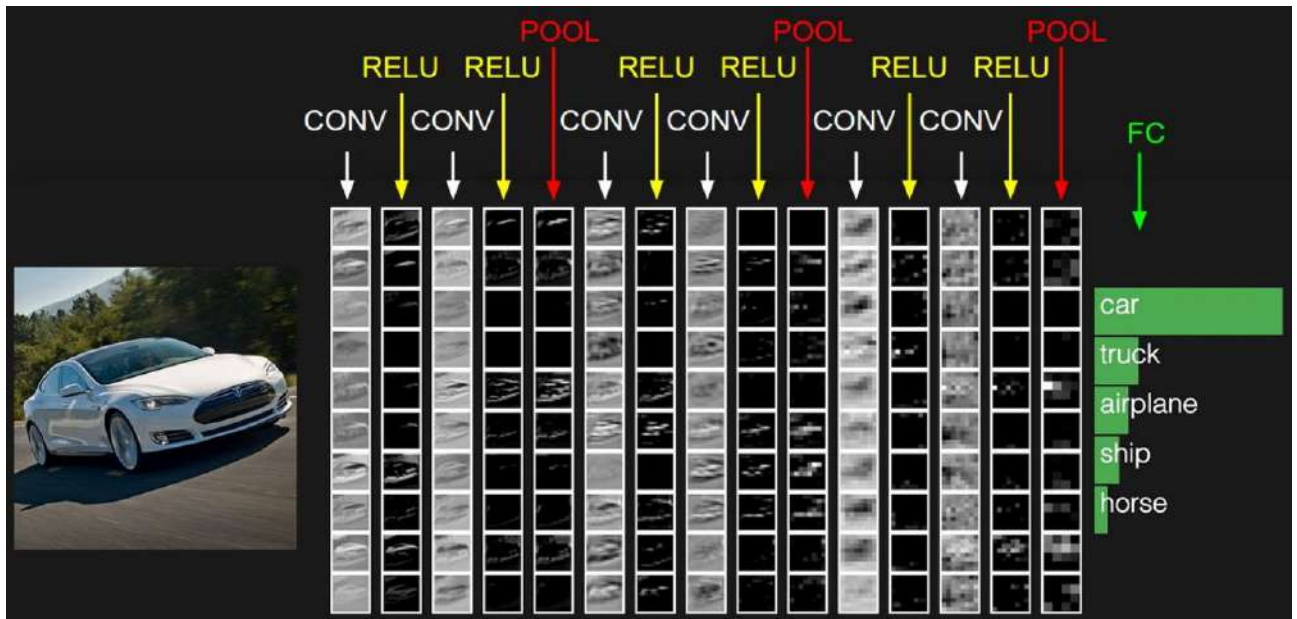


Рисунок 2.5 – Приклад роботи по розпізнаванню автомобіля [tesla.com]

Переваги CNN:

Автоматичне розпізнавання візуальних особливостей. CNN можуть самостійно вивчати різні візуальні особливості, такі як форми та текстури і не потрібно визначати ці функції.

Просторова варіація. CNN можуть розпізнавати об'єкти незалежно від їхнього положення на зображенні. Він добре працює для великих зображень, оскільки може розпізнавати об'єкти в різних частинах зображення.

Можливість повторного використання. CNN можна використовувати з попередньо навченими ваговими коефіцієнтами (наприклад, з використанням звичайних архітектур, як-от ResNet, VGG, Inception тощо), більш ефективними методами та кращими методами з меншою кількістю навчальних даних.

Ефективність в роботі з багатоканальними зображеннями. CNN може бути легко розширений до роботи з багатоканальними зображеннями, такими як RGB зображення, де можуть бути використані декілька шарів згортки та пулінгу для роботи з кожним кольоровим каналом окремо.

Недоліки CNN:

Багато параметрів. CNN може мати кілька параметрів. Зокрема, глибокі архітектури вимагають великих обсягів навчальних даних і обчислювальних ресурсів.

Обмежена робота з послідовними даними. Одна з основних переваг CNN - це також його недолік. Він не підходить для роботи з послідовними даними, такими як текст або часові ряди, де важлива залежність в часі.

CNN відмінно підходить для задач класифікації зображень, таких як визнання об'єктів, розпізнавання облич, класифікація зображень медичних досліджень та багато інших. Також може бути використаний для виділення ознак з зображень, які можуть бути використані в інших додатках, таких як виявлення об'єктів, відстеження об'єктів, рекомендації товарів на основі зображень та ін. CNN може бути використаний для розпізнавання об'єктів в реальному часі, наприклад у системах відеоспостереження, автономних автомобілях, системах допомоги водіям тощо.

Навчання з підкріпленням — ще один метод машинного навчання, який використовувався в системах автопілота. Алгоритми навчання з підкріпленням навчаються, отримуючи зворотний зв'язок на основі виконаних дій. У системах автопілота навчання з підкріпленням можна використовувати, щоб аналізувати дані отримані з навколишнього середовища.

RNN (Recurrent neural network) — це тип нейромережі, призначений для роботи з послідовними даними, такими як текст, звук або часові ряди. Він використовує рекурентні шари, що дозволяють зберігати попередній стан мережі та використовувати його для обробки нових вхідних даних. RNN може добре виконувати послідовні завдання, такі як розпізнавання мови, машинний переклад, аналіз настроїв, прогнозування часових рядів та інші.

RNN можуть бути використані для виявлення різних об'єктів на дорозі, таких як інші автомобілі, пішоходи, велосипедисти тощо. Вони можуть аналізувати відеопотік з камер автомобіля та використовувати рекурентні залежності для визначення рухових шаблонів об'єктів, їхньої швидкості, напрямку руху та інших характеристик. Також можуть бути використані для прогнозування дорожньої

обстановки, такої як визначення рівня трафіку, стану дороги, погодних умов тощо. Це може допомогти автопілоту приймати розумні рішення про швидкість руху, шлях, часи зупинок та інші аспекти керування автомобілем.

Однак, використання RNN в автопілоті також може мати свої виклики та недоліки. RNN вимагають великої кількості даних для ефективного тренування. Отримання великого обсягу даних для тренування може бути складним завданням, особливо в разі рідкісних подій на дорозі, таких як аварії або несподівані дії інших учасників руху. Використання RNN в автопілоті вимагає високої точності та надійності, оскільки невірні прогнози або помилки можуть мати серйозні наслідки на безпеку на дорозі. Робота зі збором, обробкою та аналізом великого обсягу даних може вплинути на точність та надійність RNN.

Автопілоти зазвичай використовують комплексні системи, що включають різні алгоритми та методи, такі як обробка даних зі сенсорів, комп'ютерного зору, планування маршруту та інші. Взаємодія RNN з іншими алгоритмами може бути складною та вимагати додаткових зусиль для вирішення супутніх проблем.

Остаточний успіх використання RNN в автопілоті залежить від багатьох факторів, таких як якість даних, тренування моделей, обчислювальні ресурси та взаємодія з іншими компонентами системи автопілоту. Додаткове дослідження, розробка та оптимізація моделей RNN можуть допомогти досягти високої точності, ефективності та надійності автопілотів на основі RNN.

RNN мають властивість "запам'ятовування", що дозволяє зберігати інформацію про попередні обчислення. Усі входи та приховані шари використовують однакові параметри, так як RNN виконує однакові завдання на всіх вхідних даних, створюючи вихід. Це дозволяє зменшити складність параметрів, відрізняючись від інших нейронних мереж.

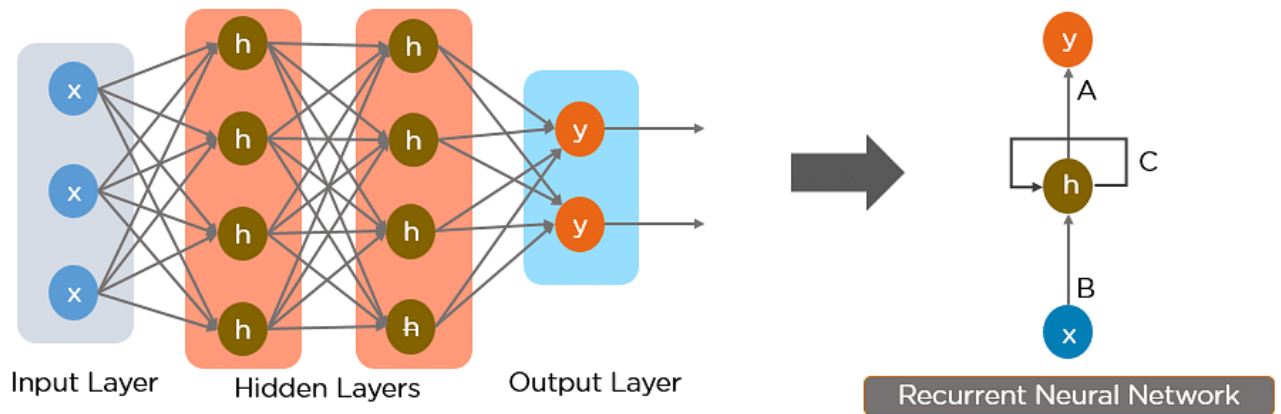


Рисунок 2.6 – Візуальне представлення рекурентної моделі [Рисунок виконаний самостійно*]

Вхідний рівень «x» приймає вхідні дані для нейронної мережі, обробляє їх і передає на середній рівень. Середній шар «h» може складатися з кількох прихованих шарів, кожен зі своїми функціями активації, вагами та зміщеннями. Рекурентна нейронна мережа стандартизує різні функції активації, ваги та зміщення, щоб кожен прихований шар мав однакові параметри. Потім, замість створення кількох прихованих шарів, він створить один і повторить його стільки разів, скільки потрібно.

Розглянемо базову архітектуру RNN та створення її рівнів за допомогою мови програмування Python та використанням Python Keras:

```
from keras.models import Sequential
from keras.layers import LSTM, Dense
model = Sequential()
model.add(LSTM(128, activation='tanh', input_shape=(timesteps,
input_dim)))
model.add(Dense(num_classes, activation='softmax'))
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=[
'accuracy'])
model.summary()
```

Переваги RNN:

RNN розроблені спеціально для роботи з послідовними даними, такими як текст, аудіо, відео, часові ряди тощо, де залежність в часі є важливою. RNN можуть моделювати довготривалі залежності між послідовними даними.

RNN можуть бути використані для різноманітних задач, таких як машинний переклад, генерація тексту, прогнозування часових рядів, аналіз емоцій у тексті та багато інших

RNN можуть бути розширені до глибоких архітектур, таких як LSTM (Long Short-Term Memory) та GRU (Gated Recurrent Unit), що дозволяє моделювати складніші залежності в послідовних даних.

Недоліки RNN:

RNN можуть стикатися з проблемою зникаючих градієнтів, коли навчання моделі на довгих послідовностях може бути вкрай складним. Це може призводити до проблеми з пам'яттю довготривалих залежностей, особливо в довгих послідовностях.

RNN можуть бути вимогливими до ресурсів відносно обчислювальної потужності, особливо при роботі з великими послідовностями даних, такими як відео або аудіо. Вони можуть бути менш ефективними в порівнянні з іншими методами машинного навчання, такими як CNN, на великих обсягах даних.

У RNN всі входи мають однакові ваги, що може бути недостатнім для деяких завдань, де різні входи можуть мати різну важливість або взаємодія між входами може бути складною.

Основна відмінність між CNN і RNN полягає в їх здатності обробляти часову інформацію, яка з'являється в таких послідовностях, як речення. Рекурентні нейронні мережі розроблені для цієї мети, але згорткові нейронні мережі не можуть адекватно пояснити часову інформацію. Таким чином, CNN і RNN служать абсолютно різним цілям, і структури нейронних мереж, що відповідають цим різним програмам, також різні.

CNN використовують фільтри всередині згорткових шарів для перетворення даних, тоді як RNN є прогнозними, повторно використовуючи функції активації з інших точок даних у послідовності для генерації наступного виходу в серії.

Результати цього дослідження дозволяють нам створювати системи, які здатні відстежувати об'єкти та передбачати їх рух у режимі реального часу. Наприклад, система, яка використовує глибоке навчання, може автоматично

ідентифікувати дорожні знаки, виявляти пішоходів і автомобілі та відстежувати їхнє пересування.

Ці системи мають незліченну кількість застосувань, особливо в областях, де безпека є ключовою проблемою. Наприклад, системи машинного навчання, які виявляють і прогнозують рух об'єктів у реальному часі, можна використовувати в автономних транспортних засобах, щоб уникнути аварій і забезпечити безпеку пасажирів і пішоходів. Їх також можна використовувати в системах безпеки для виявлення потенційних загроз і відповідного реагування.

Найкращий метод машинного навчання для систем автопілота залежить від конкретних вимог системи. Глибоке навчання, зокрема CNN, продемонстрували великий потенціал для розпізнавання зображень і виявлення об'єктів, що робить їх добре придатними для використання в автономних транспортних засобах та інших подібних системах. Навчання з підкріпленням також можна використовувати для таких же цілей у автомобільній індустрії, але воно вимагає значних обчислювальних ресурсів і більш широких даних для досягнення точних результатів.

Однак CNN зазвичай використовуються для обробки відео та зображень, тоді як RNN використовуються для обробки послідовностей даних, таких як текстові повідомлення та голосові команди. Оскільки основним завданням автопілота є виявлення об'єктів і збір інформації з датчиків, CNN можуть виконувати ці завдання більш ефективно. Вони можуть ефективно виконувати завдання виявлення об'єктів на зображенні, наприклад виявлення дорожньої розмітки, світлофорів, автомобілів та інших об'єктів на дорозі. Тому для дослідження та порівняння у роботі будуть використовуватись саме CNN.

Методи машинного навчання для розпізнавання об'єктів і прогнозування руху в режимі реального часу швидко розвиваються та мають широкий спектр застосувань у різних галузях промисловості. Оскільки аналізи, дослідження та розробки все ще тривають, можна очікувати появи ще більш складних систем, здатних робити точніші прогнози та забезпечувати ще більшу безпеку у нашому житті.

3 ПРОВЕДЕННЯ ТЕОРЕТИЧНОГО ТА ЕКСПЕРИМЕНТАЛЬНОГО ДОСЛІДЖЕННЯ

3.1 Розробка та проведення дослідження

Основна мета для використання NAS (Neural Architecture Search) полягає у знаходженні найбільш оптимальної архітектури для конкретних потреб, що використовуються в нейронних мережах. Для досягнення цієї мети використовується процес, який полягає у програмуванні, налаштуванні нейронної мережі та вивченні того, що працює добре, після чого це завдання автоматизується для виявлення більш складних архітектур. NAS — це підполе AutoML, яке інкапсулює всі процеси, які автоматизують проблеми машинного навчання, а отже і глибокого навчання. Велика кількість дослідницьких робіт стосується автоматизації пошуку архітектур нейронної мережі в різних галузях і різних проблемах. Вже сьогодні багато ручних архітектур були витіснені архітектурами, створеними на основі NAS: Object detection — Image Processing, Image classification — Image Processing та багато інших.

Методи NAS досліджують багато потенційних рішень із різною складністю і, отже, дуже дорогі з точки зору обчислень. Чим більший простір їх пошуку, тим більше є архітектур для тестування, навчання, оцінки. Ці методи потребують величезних ресурсів і днів, щоб знайти достатньо хорошу модель. Їхній метод навчання з підкріпленням Zoph et al. використовував 800 графічних процесорів NVIDIA K40 протягом 28 днів. Починаючи з перших методів, з'явилися нові моделі з набагато коротшим часом пошуку. Наприклад, Progressive Neural Architecture Search продемонстрував подібні найсучасніші результати з 5–8 разів швидшим часом пошуку. Щоб знайти цю архітектуру, потрібно приблизно 7 годин, що зменшує кількість GPU-годин більш ніж у 50 000 разів порівняно з NAS.

У нейронної мережі є три можливі класифікації (рис. 3.1). Перша класифікація передбачає навчання мережі за допомогою одного ярлика на вході, щоб забезпечити успішне навчання. Друга класифікація полягає в тому, що всі можливі точки на вхідному зображенні задаються, що дозволяє нейронній мережі

бути більш точною. Третя класифікація здійснює пошук об'єкта за допомогою автоматично заданих ярликів.

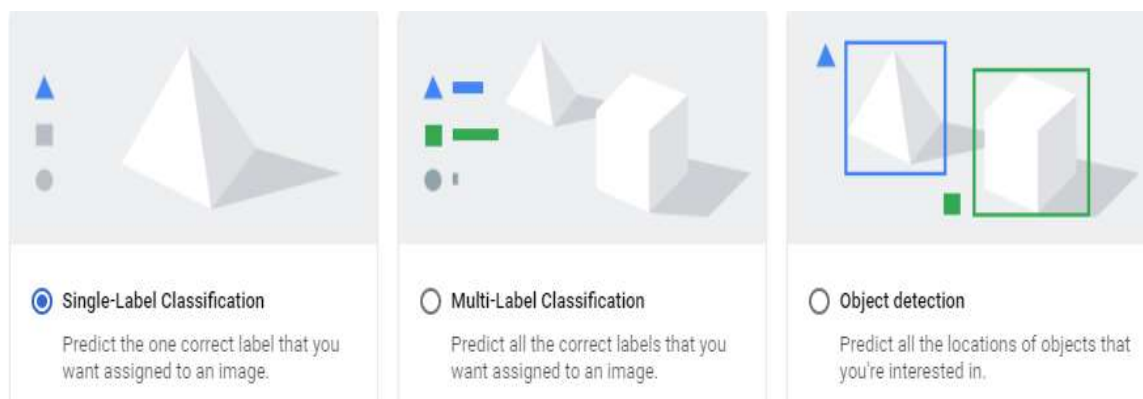


Рисунок 3.1 – Класифікація для навчання моделі [Рисунок виконаний самостійно*]

Для навчання власної нейронної мережі з розпізнавання різноманітних об'єктів необхідно створити Buckets, в які можна завантажити вхідні дані для навчання мережі, які будуть мати певну класифікацію (рис. 3.2).



Рисунок 3.2 – Класифікація вхідних даних [Рисунок виконаний самостійно*]

Потім скопіюєм зображення кожного типу та кожної класифікації і створимо «брудний» (dirty) набір даних, помістивши необхідні нам вхідні дані у необхідну папку для цієї класифікації (рис. 3.3).

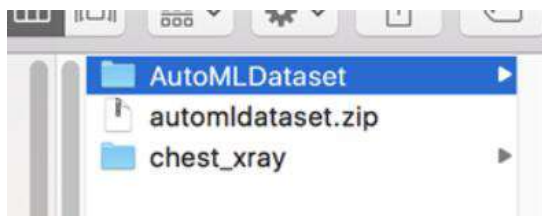


Рисунок 3.3 – Класифікація даних та створення automldataset.zip (Рисунок виконаний самостійно*)

Після класифікації наших вхідних даних в нових папках, заархівуємо всі дані у папку AutoMLDataset, щоб створити automldataset.zip для роботи з нашою моделлю. У цих віджетах можна зберігати зображення у звичайних форматах для зчитування (JPG, PNG, GIF, BMP, ICO, ZIP).

Для досягнення ідеального навчання мережі рекомендується використовувати приблизно 100 зображень на кожен тип, мінімальний же поріг навчання починається від 10 вхідних даних, але не більше 500 на кожен. Це дозволить автоматично визначати правильну позначку істинності для кожного зображення, дивлячись на ім'я папки, в якій воно міститься. У нашому дослідницькому наборі ми візьмемо меншу кількість вхідних даних для економії обчислювальних ресурсів та часу.

Низькі вхідні дані можуть призвести до низької точності та поганого охоплення дисперсії функцій, що може призвести до низької ефективності навчання мережі. Кінцевим результатом маркування має бути чітке призначення категорії для всіх завантажених зображень, які система використовує для навчання (рис. 3.4). Важливо використовувати зображення високої якості, щоб уникнути помилок і зайвого шуму під час навчання і подальшого розпізнавання.

Основна причина полягає в тому, що більш висока якість зображення дозволяє моделі машинного навчання легше розрізняти та визнавати різні об'єкти та деталі на зображенні. Якщо зображення містить багато шуму, розмивання або інших дефектів, то модель може помилково інтерпретувати ці дефекти як важливі ознаки та отримати неправильний результат.

Також, якість зображень може впливати на швидкість навчання моделі. Зображення низької якості можуть займати більше часу на обробку, що збільшує час, потрібний для тренування моделі під час виконання на нових зображеннях.


<input type="checkbox"/>	Thumbnail	Labels
<input type="checkbox"/>		main(1)
<input type="checkbox"/>		stop(1)
<input type="checkbox"/>		stop(1)
<input type="checkbox"/>		stop(1)
<input type="checkbox"/>		stop(1)
<input type="checkbox"/>		stop(1)
<input type="checkbox"/>		main(1)

Рисунок 3.4 – Список завантажених даних з отриманими мітками [Рисунок виконаний самостійно*]

Навчання даної розробленої моделі зайняло близько двох годин, включаючи навчання вузла, налаштування моделі та створення інфраструктури системи.

Після успішного завершення побудови та навчання моделі, ми отримали нейронну мережу зі здатністю до виявлення об'єктів на рівні точності 75% (рис. 3.5). Таку точність можна пояснити тим, що ми використовували мінімальну кількість вхідних даних для навчання, замість рекомендованих 100, що може вплинути на точність розпізнавання.

Total images	21
Test items	4
Precision 	75%
Recall 	75%

Use the slider to see which confidence threshold works best for your model on the precision-recall tradeoff curve.

[Learn more about these metrics and graphs.](#)

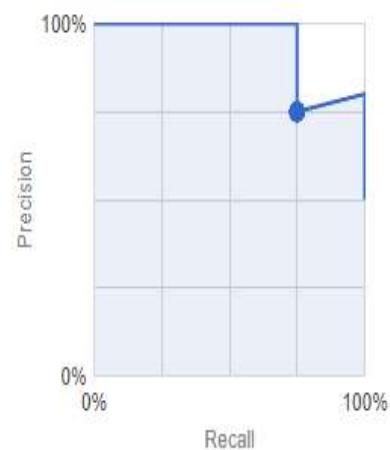


Рисунок 3.5 – Точність моделі [Рисунок виконаний самостійно*]

При завантаженні тестових даних ми отримали точність розпізнавання на рівні 0.97, 0.67 та 0.86. Варно зазначити, що рівень розпізнавання дуже залежить від якості зображення. На прикладі знака 2.2 проїзд без зупинки заборонено можемо побачити точність 0.61, що зумовлено поганою якістю зображення, нанесенням водермарки та паданням тіні на нього (рис. 3.6).



Рисунок 3.6 – Розпізнавання знака 2.2 на зображенні поганої якості [Рисунок виконаний самостійно*]

Але якщо обрати для розпізнавання зображення більш якіснішого рівня, то і вірогідність успіху у розпізнаванні буде значно вище (рис. 3.7).



Рисунок 3.7 – Розпізнавання знака 2.2 на зображенні нормальної якості
[Рисунок виконаний самостійно*]

Побудуємо матрицю розпізнавання після перевірки моделі на двох дорожніх знаках: 2.2 проїзд без зупинки заборонено та 2.3 головна дорога (рис. 3.8).

stop	67%	33%
main	-	100%

Рисунок 3.8 – Confusion matrix (матриця похибки) [Рисунок виконаний самостійно*]

Як можемо побачити з таблиці, знак 2.3 головна дорога розпізнавався моделлю зі 100% вірогідністю. А ось зі знаком 2.2 проїзд без зупинки заборонено виникли проблеми. Як згадувалось раніше, зумовлено це низькою якістю зображень знака 2.2 у тестовому наборі.

Для порівняння та проведення дослідження створимо аналогічну модель за допомогою CNN. Вона є гарним інструментом у глибокому та машинному навчанні

і застосовується у багатьох областях, таких як комп'ютерний зір, обробка зображень, розпізнавання образів, медична діагностика та багато іншого. Основна ідея CNN полягає у використанні сверток, що походять по вхідному зображенню, для виділення важливих ознак та структур. Далі ці ознаки оброблюються шарами підвищення абстракції, що дозволяє моделі виявляти більш складні зразки на зображенні. У цілому, CNN є важливою моделлю для роботи з великими наборами даних, зокрема зображеннями та відео.

Для навчання нашої моделі ініціалізуємо два списки `data` та `labels` (рис. 3.9).

```
data = []
labels = []
```

Рисунок 3.9 – Списки для збереження даних [Рисунок виконаний самостійно*]

Ці списки будуть зберігати вхідні дані та їх мітки для класифікації зображення. Далі ми перебираємо всі класи та додаємо відповідні зображення та мітки до списку `data` і `labels`, що показано на коді нижче:

```
for num in range(0, classes):
    path = os.path.join('train', str(num))
    imagePaths = os.listdir(path)
    for img in imagePaths:
        image = Image.open(path + '/' + img)
        image = image.resize((30,30))
        image = img_to_array(image)
        data.append(image)
        labels.append(num)
```

Як і у першому методі ми також використаємо 80% вхідних даних для зображення, а 20% для тестування моделі.

Далі побудуємо архітектуру для створення нашої CNN моделі. Модель складається з декількох шарів, загалом включаючи згорткові шари, шари

підвибірки, шари активації та пов'язані шари, що продемонстровано на фрагменті коду нижче.

```

model = Sequential()
    model.add(Conv2D(filters=32, kernel_size=(5,5), activation='relu',
input_shape=inputShape))
    model.add(Conv2D(filters=32, kernel_size=(5,5),
activation='relu'))
    model.add(MaxPool2D(pool_size=(2, 2)))
    model.add(Dropout(rate=0.25))
    model.add(Conv2D(filters=64, kernel_size=(3, 3),
activation='relu'))
    model.add(Conv2D(filters=64, kernel_size=(3, 3),
activation='relu'))
    model.add(MaxPool2D(pool_size=(2, 2)))
    model.add(Dropout(rate=0.25))
    model.add(Flatten())
    model.add(Dense(500, activation='relu'))
    model.add(Dropout(rate=0.5))
model.add(Dense(classes, activation='softmax'))

```

Після навчання нашої моделі за допомогою функції `model.fit()`, ми отримали модель яка на 25 епосі дала точність 92% (рис. 3.10).

```

loss: 0.2679 - accuracy: 0.9205 - val_loss: 0.0425 - val_accuracy: 0.9888

```

Рисунок 3.10 – Точність моделі [Рисунок виконаний самостійно*]

Протестуємо обрану модель на тестовому наборі зі знаком 5.19.1 пішохідний перехід (рис. 3.11). Під час етапу підготовки даних було використано зображення з відкритих джерел та ручне маркування кожного знаку, було проведено попередню обробку зображень та створено `Buckets` для навчання моделі.



Рисунок 3.11 – Перевірка моделі на тестовому наборі [Рисунок виконаний самостійно*]

На основі отриманих результатів можна зробити висновок, що CNN модель є ефективним та гарним інструментом для розпізнавання об'єктів на дорозі для автопілота. Проте, в процесі навчання необхідно враховувати відповідність кількості вхідних даних та їх якості, що може вплинути на результати точності моделі.

У майбутньому можна розглядати можливість покращення моделі шляхом використання додаткових методів та алгоритмів навчання. Варто спробувати додати більше шарів до моделі, щоб збільшити її якість та точність. Наприклад, додавання шару BatchNormalization або Dropout може допомогти уникнути перенавчання. Збільшення вхідних даних до тренувального набору за допомогою аугментації даних може допомогти покращити роботу моделі. Наприклад, можна використовувати збільшення, зменшення, повороти, зміну контрастності та яскравості тощо.

За допомогою інформаційно-вказівних знаків, які сигналізують про появу пішоходів (рис. 3.12) система може прогнозувати їх появу на дорожній зоні, що може значно знизити рівень небезпеки на дорозі.



Рисунок 3.12 – Знаки за допомогою яких можливо прогнозувати появу пішоходів [Рисунок виконаний самостійно*]

У більш складних системах можливе інформування водія у разі виявлення пішохода близько до дороги. Зображення можуть бути оброблені з використанням досліджених методів виявлення об'єктів, щоб визначити наявність пішоходів та їхню поведінку. Для прикладу, якщо людина пересувається паралельно дорозі, то і вірогідність його виходу на зону автомобілів низька, але якщо траєкторія руху є перпендикулярною до дороги, то є вірогідність виходу пішохода [18].

Пристрої сканування, такі як радар і лазерні датчики, можна використовувати для виявлення присутності людей у зоні та надсилання цієї інформації системам керування розумного автомобіля. Дані можна обробляти за допомогою методів машинного навчання, щоб передбачити рух і місцезнаходження пішоходів на основі інформації від датчиків.

У той же час, щоб успішно використовувати машинне навчання, потрібно мати достатній обсяг даних, які мають відношення до завдання прогнозування. Також важливо розробити ефективні та точні алгоритми та оновити їх на основі оновлень даних і змін середовища.

Також для прогнозування можливе використання GPS з мобільних додатків, які збирають дані про місцезнаходження користувачів. Після чого, дані можуть збиратися та передаватися на систему автомобіля та прогнозувати появу пішохода на дорожній зоні. Отримані дані у подальшому можуть бути оброблені за допомогою методів машинного навчання, які будуть прогнозувати потенційну появу пішоходів на певній ділянці дороги на основі різноманітних вхідних даних (погодні умови, період часу, розміщення офісних або державних закладів, тощо).

Застосування машинного навчання для прогнозування присутності пішоходів на дорозі потенційно може зменшити кількість дорожньо-транспортних пригод і підвищити безпеку дорожнього руху для всіх учасників дорожнього руху.

3.2 Результати дослідження та можливість їх використання у практичній діяльності

За результатами дослідження ми отримали точність моделі яка була побудована за допомогою NAS – 75%, а модель на основі CNN – 92%. Така різниця обумовлена багатьма факторами. CNN зазвичай складаються з послідовних стеків згорткових шарів, які дозволяють моделі автоматично ідентифікувати візуальні характеристики, важливі для розпізнавання зображення. Модель NAS, яка використовує складнішу архітектуру, може мати знижену ефективність і точність архітектури моделі. Такі параметри моделі, як кроки навчання та кількість епох, впливають на точність моделі. Якщо модель навчена надто мало або надто повільно, вона може не використовувати весь свій потенціал, що призведе до зниження точності.

Ієрархічна структура згорткових шарів дозволяє виявляти та аналізувати властивості зображення на різних рівнях деталей. У порівнянні з моделлю NAS, у якої генерація архітектури відбувається шляхом еволюції, CNN дозволяє розробити більш точну та ефективну архітектуру для розпізнавання об'єктів які будуть використовуватись у автопілоті.

Під час навчання моделі CNN застосовуються методи BatchNormalization та Dropout, що допомагають уникнути перенавчання та забезпечити більш точні результати.

CNN використовує згорткові шари та шари підвибірки, що дозволяє ефективно виявляти та виділяти ознаки на зображенні. У порівнянні з моделлю NAS, яка використовує більш складну архітектуру з великою кількістю шарів, CNN може працювати швидше та з більшою точністю.

Використання такої моделі для розпізнавання об'єктів має практичну цінність і може знайти своє застосування в різних сферах.

Одним з головних застосувань CNN для розпізнавання об'єктів на дорозі є підвищення безпеки на дорогах та запобігання аварій. Застосування автоматизованих систем розпізнавання дорожніх знаків дозволяє швидко та ефективно ідентифікувати дорожні знаки та інші об'єкти на дорозі, що може бути особливо корисним для автопілотованих автомобілів.

У сфері логістики, модель може бути використана для автоматичного розпізнавання номерних знаків на транспортних засобах та під час відстеження руху вантажівок, що допоможе забезпечити ефективність логістичних операцій.

У сфері містобудування, модель може бути використана для розпізнавання дорожньої інфраструктури, що дозволить покращити планування та управління міським транспортом та збільшити комфорт міського життя.

Дану модель можливо ефективно використовувати для прогнозування появи пішоходів на дорожній зоні по виявленню спеціальних дорожніх знаків, що регулюють місця для переходу. Також можливо налаштувати модель на розпізнавання спеціальних елементів, які ідентифікують місце для переходу для сповіщення водіїв [19].

Таким чином, використання моделі CNN для розпізнавання об'єктів на дорозі може мати значний практичний вплив у різних сферах та допомогти покращити ефективність та безпеку руху.

ВИСНОВОК

Майбутнє досить швидко стає сьогоднішнім, проблеми які ще рік тому нікого не хвилювали, зараз стають досить актуальними. Сьогодні ми все частіше чуємо про автономні автомобілі та інші робототехнічні пристрої, здатні самостійно пересуватися навіть дорогами загального користування.

Вони застосовують системи машинного зору з урахуванням нейронних мереж, однієї з найбільш просунутих напрямів штучного інтелекту. Ці системи здійснюють розпізнавання та ідентифікацію людей та предметів за фотографіями, зображеннями з відеокамер та лідарів. Розпізнавання об'єктів може досить ефективно допомогти людству зі створенням ідеального автопілотованого транспортного засобу.

Машинне навчання у сфері розпізнавання об'єктів є невід'ємною частиною систем автономного водіння в автомобілях. Щоб автопілот був корисним у реальному світі, система повинна мати можливість розпізнавати різні об'єкти на дорозі та в навколишньому середовищі, наприклад: пішоходи, дорожні знаки, будівлі та інші перешкоди.

Як правило, автопілотовані системи навчаються з використанням великих обсягів даних, таких як зображення з камер, дані з лідарів, радарів та інших датчиків. Після навчання система може розпізнавати об'єкти на дорозі та приймати рішення щодо керування транспортним засобом.

У даному дослідженні було розглянуто та проаналізовано методи для створення та навчання системи, яка зможе використовуватися у розумному авто. Головними факторами для такої системи належать якість, швидкість та масштабованість. Також важливим фактором для навчання є велика кількість вхідних даних.

Розпізнавання образів є досить складною та перспективною нішею. У сучасному світі сотні компаній моделюють та тестують системи задля досягнення високих показників у розпізнаванні з перспективою використання у різних напрямках.

У перспективі реалізована система може прогнозувати та сповіщати водіїв про неочікувані фактори. Для прикладу, це може бути поява пішоходів або іншого транспорту, який має перевагу у русі.

За результатами дослідження найкращі результати показала система, побудована на основі CNN. Але при цьому потрібно враховувати різноманітні обмеження, такі як відсутність достатньої кількості даних для навчання, складність дорожньої ситуації, погодні умови та інші. Однак є можливості для ще більшого покращення результатів та отримання кращої точності. Для прикладу це збільшення кількості вхідних даних, покращення якості даних для навчання, що може бути розглянуто у майбутніх дослідженнях.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Human factors contributing to the road traffic accident occurrence / sciencedirect – URL: <https://www.sciencedirect.com/science/article/pii/S2352146520302192> (дата звернення: 10.12.2022).
2. Increase in road traffic / Jack Bernstein – URL: <https://bernsteininjurylaw.com/blog> (дата звернення: 27.12.2022).
3. Three Revolutions in Urban Transportation / ITDP – URL: <https://www.itdp.org/2017/05/03/3rs-in-urban-transport/> (дата звернення: 07.01.2023).
4. Звіти Tesla / itc.ua – URL: <https://itc.ua/ua/tag/zviti-tesla/> (дата звернення: 27.01.2023)
5. Future of Driving / Tesla autopilot – URL: <https://www.tesla.com/autopilot> (дата звернення: 27.01.2023)
6. Autopilot and Full Self-Driving Capability / Tesla – URL: <https://www.tesla.com/support/autopilot> (дата звернення: 27.03.2023).
7. Google Cloud – URL: <https://cloud.google.com/vision/automl/docs> (дата звернення: 27.03.2023).
8. oreilly / What is neural architecture search? – URL: <https://www.oreilly.com/content/what-is-neural-architecture-search/> (дата звернення: 23.03.2023)
9. Wiki / Neural architecture search – URL: https://en.wikipedia.org/wiki/Neural_architecture_search (дата звернення: 17.02.2023).
10. Deep Learning (Adaptive Computation and Machine Learning series) / Ian Goodfellow, Yoshua Bengio, Aaron Courville – 2016 – 800 с.
11. Розпізнавання образів і машинне навчання / Бішоп К. М. – 2020 – 419 с.
12. Artificial Vision and Language Processing for Robotics / Alvaro Morena Alberola – 2019 – 356 с.

13. Machine Learning Models Efficiency Analysis for Image Classification Problem / Smelyakov, K – 2022.
14. Computer Vision: Algorithms and Applications / Richard Szeliski – 2022 – 725 с.
15. Neural Computing and Applications / John MacIntyre – 2021 – 425 с.
16. pyimagesearch / Real-time object detection with deep learning – URL: <https://pyimagesearch.com/2017/09/18/real-time-object-detection-with-deep-learning-and-opencv/> (дата звернення: 24.03.2023)
17. arxiv / A Survey of Deep Learning-based Object Detection – URL: <https://arxiv.org/abs/1907.09408> (дата звернення: 24.03.2023)
18. Towardsdatascience / People Tracking using Deep Learning – URL: <https://towardsdatascience.com/people-tracking-using-deep-learning-5c90d43774be> (дата звернення: 27.03.2023)
19. Towardsdatascience / Object Detection with 10 lines of code – URL: <https://towardsdatascience.com/object-detection-with-10-lines-of-code-d6cb4d86f606> (дата звернення 29.03.2023)
20. Towardsdatascience / A Comprehensive Guide to Convolutional Neural Networks – URL: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> (дата звернення: 29.03.2023)

**ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ ЗА НАУКОВИМИ НАПРЯМАМИ
КЕРІВНИКА ТА НАУКОВЦІВ КАФЕДРИ ПРОГРАМНОЇ ІНЖЕНЕРІЇ**

13. Machine Learning Models Efficiency Analysis for Image Classification Problem / Smelyakov, K – 2022.