

**ПОЛТАВСЬКИЙ ДЕРЖАВНИЙ АГРАРНИЙ УНІВЕРСИТЕТ  
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ЕКОНОМІКИ, УПРАВЛІННЯ,  
ПРАВА ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ  
КАФЕДРА ІНФОРМАЦІЙНИХ СИСТЕМ ТА ТЕХНОЛОГІЙ**

**Пояснювальна записка**

до кваліфікаційної роботи на здобуття ступеня вищої освіти Бакалавр

на тему: «Розпізнавання типу дорожнього покриття»

Виконав: здобувач вищої освіти  
за освітньо-професійною програмою  
Інформаційні управляючі системи  
спеціальності 126 Інформаційні  
системи та технології  
ступеня вищої освіти Бакалавр  
групи 126ІСТбд41  
Кабак Д.І.  
Керівник: Слюсар В.І.  
Рецензент: Брикун О.М.

**Полтава – 2024 року**

## ВСТУП

*Актуальність* теми кваліфікаційної роботи підтверджується необхідністю розпізнавання типу дорожнього покриття в інтересах точного визначення типу покриття для безпеки, ефективності руху, оптимізації витрат енергії та прогнозування часу подолання маршрутів, що є критичним для розвитку робототехніки та автономного транспорту. Основні типи дорожнього покриття включають асфальт, бетон, бруківку та ґрунтові дороги. Для розпізнавання типу дорожнього покриття існує кілька технологій. До них відноситься візуальний аналіз, використання різних сенсорів, машинне навчання. Однак питання розпізнавання типу дорожнього покриття на основі нейронної мережі потребує додаткових досліджень. Все це свідчить про актуальність теми роботи.

*Метою* кваліфікаційної роботи є підвищення ефективності систем автопілотування безпілотних транспортних засобів за рахунок розпізнавання типу дорожнього покриття.

*Завданнями* кваліфікаційної роботи є:

- розробка моделі розпізнавання типу дорожнього покриття на основі нейронної мережі;
- порівняльна оцінка точності нейронних мереж розпізнавання типу дорожнього покриття;
- формування рекомендацій щодо реалізації розпізнавання типу дорожнього покриття за допомогою нейронної мережі;
- економічне обґрунтування прийнятих рішень.

*Об'єктом* дослідження є процес розпізнавання типу дорожнього покриття за допомогою нейронних мереж.

*Предметом* дослідження є точність розпізнавання нейронними мережами різної архітектури.

*Методами* дослідження є в рамках визначення інструментарію для розпізнавання типу дорожнього покриття і економічного обґрунтування прийнятих рішень використовувався аналітичний метод досліджень, а для

розробки архітектури нейронних мережі і формування датасету – моделювання.

*Інформаційна база* кваліфікаційної роботи сформована з Інтернет-ресурси, що містять інформацію про згорткові нейронні мережі, великі мовні моделі, датасети зображень дорожнього покриття.

*Практична значущість* роботи полягає у розробці моделі розпізнавання типу дорожнього покриття на основі нейронної мережі; порівняльній оцінці точності нейронних мереж розпізнавання типу дорожнього покриття; формуванні рекомендацій щодо реалізації розпізнавання типу дорожнього покриття за допомогою нейронної мережі – можуть бути використані для подальших досліджень за даною тематикою та при проектуванні реалізації систем автопілотування безпілотних транспортних засобів.

*Апробація результатів* відбувалася в рамках ХІХ щорічної студентської наукової конференції «Сучасні інформаційні технології та інноваційні методики в економіці, менеджменті та бізнесі» Полтавського державного аграрного університету (14 травня 2024 р., м. Полтава).

За результатами досліджень здійснено публікації тез доповідей.

*Структура кваліфікаційної роботи* логічно пов'язана з завданнями досліджень і містить вступ, три розділи основної частини, висновки, список використаних джерел, додатки. Загальний обсяг пояснювальної записки кваліфікаційної роботи складає 54 сторінки формату А4. Вона містить 27 рисунків.

# РОЗДІЛ 1

## АНАЛІЗ ОСОБЛИВОСТЕЙ РОЗПІЗНАВАННЯ ТИПУ ДОРОЖНЬОГО ПОКРИТТЯ

### 1.1 Аналіз варіантів застосування систем розпізнавання типу дорожнього покриття

Сучасний розвиток технологій, зокрема у сфері робототехніки [1] та штучного інтелекту [2], відкриває нові можливості для автоматизації та підвищення ефективності різних процесів. Одним з перспективних напрямків є розробка наземних дронів [3], які здатні автоматично орієнтуватися та пересуватися по різних типах дорожнього покриття.

Наземні дрони з функцією розпізнавання дорожнього покриття можуть бути використані в різних галузях, таких як логістика, комунальні служби, сільське господарство, рятувальні операції та багато інших. Наземні дрони, які здатні розпізнавати різні типи покриття, можуть автоматично адаптувати свою швидкість та траєкторію руху, що знижує ризик аварійних ситуацій та підвищує безпеку як для самого дрона, так і для оточуючих.

Знання про тип дорожнього покриття дозволяє дрону оптимізувати свої алгоритми пересування, що підвищує загальну ефективність його роботи.

По-перше, це дозволяє автомобілю або іншому типу UGV (Unmanned Ground Vehicle) [4] адаптувати свої системи керування для забезпечення максимальної безпеки та комфорту. Різні поверхні, такі як асфальт або бруківка, мають різні характеристики зчеплення, що впливає на гальмівний шлях та стійкість автомобіля. По-друге, розпізнавання типу поверхні допомагає оптимізувати витрати палива чи електроенергії, оскільки деякі дорожні покриття вимагають більше енергії для пересування. Третя причина полягає в покращенні навігації та планування маршруту, оскільки інформація про тип покриття може допомогти визначити найбільш підходящі шляхи для руху, уникаючи, наприклад, ділянок з поганим дорожнім покриттям, що

знижує ризик аварій та пошкодження транспортного засобу. Нарешті, це важливо для загального сприйняття довкілля автомобілем, що необхідно для забезпечення гладкої та безпечної інтеграції у дорожній рух, підвищення довіри користувачів та дотримання нормативних вимог.

Розпізнавання типу дорожнього покриття також відіграє ключову роль у прогнозуванні швидкості руху та часу подолання різних ділянок маршруту. Знання типу покриття дозволяє безпілотним транспортним засобам адаптувати свою швидкість відповідно до дорожніх умов, що не тільки сприяє безпеці, а й підвищує ефективність поїздки. Наприклад, на слизькому або нерівному покритті, такому як бруківка, транспортний засіб може заздалегідь знизити швидкість для запобігання втраті керування, в той час як на хорошому асфальті може рухатися швидше.

Точне розуміння часу подолання різних ділянок дороги є важливим для розрахунку оптимального маршруту і часу прибуття, що має велике значення як для особистого використання, так і для комерційних перевезень. Це допомагає у плануванні та гарантує більш точне дотримання розкладу, що особливо важливо для служб доставки та пасажирських перевезень.

Крім того, розуміння дорожніх умов допомагає у прийнятті рішень щодо обходу певних ділянок дороги, які можуть бути тимчасово або постійно непридатні для руху через погане покриття або надмірно зношувати шасі та трансмісію. Це покращує загальну безпеку та надійність подорожі, зменшуючи ймовірність несподіваних зупинок чи аварій, дозволяє продовжити технічний ресурс платформи двигуна.

Це особливо важливо для виконання завдань у міських умовах, де різноманітність дорожніх покриттів є значною. Правильне розпізнавання типу покриття допомагає уникнути зайвого зносу обладнання та зменшити витрати на технічне обслуговування дронів. Розробка ефективних алгоритмів розпізнавання дорожнього покриття сприяє загальному розвитку робототехніки та штучного інтелекту, стимулюючи інновації в цих сферах. Умови дорожнього покриття можуть змінюватися залежно від погодних умов,

сезону та інших факторів. Дрони, що здатні адаптуватися до цих змін, забезпечують більш надійну та стабільну роботу.

Крім наземних дронів варто також звернути увагу на реалізацію автопілотів. Вібрації є одним із найбільш критичних негативних явищ, що генеруються системами та засобами транспорту. Вони небажані, оскільки значно впливають на рівень безпеки і деградацію технічного середовища. Оцінка рівня вібрацій, які впливають на людське тіло, одна із найважливіших показників, підтверджують комфорт поїздки. Недотримання критеріїв впливу вібрацій на людське тіло, викликаного рухом транспортного засобу, може призвести до серйозних порушень сприйняття, що ускладнює виконання екіпажем оперативних дій. Умови оптимізації комфорту аналізу елементів крісел і несприятливих впливів на людське тіло важливі, але також слід пам'ятати про індекс оперативної готовності, оскільки втрата готовності, викликана занадто високими вібраціями, може пошкодити один з основних елементів, відповідальних за мобільні якості та ін.

## **1.2 Інструментарій для розпізнавання типу дорожнього покриття**

Розпізнавання типу дорожнього покриття – це процес ідентифікації матеріалу, з якого зроблено дорогу, і може включати різні методи та технології. Це може бути корисним для автомобільних навігаційних систем, планування інфраструктури міста, а також для покращення безпеки на дорогах. Основні типи дорожнього покриття включають асфальт, бетон, бруківку та ґрунтові дороги. Для розпізнавання типу дорожнього покриття існує кілька технологій.

Однією з них є візуальний аналіз, який базується на використанні камер для збору зображень дорожнього покриття. Після цього зібрані дані аналізуються за допомогою алгоритмів комп'ютерного зору [5]. Перший етап у процесі розпізнавання типу дорожнього покриття – це збір даних. На

транспортні засоби встановлюються камери, які знімають дорогу під час руху. Ці камери можуть бути як традиційними (RGB), так і спеціалізованими, такими як інфрачервоні або стереокамери. Вони дозволяють отримати зображення дороги в різних умовах освітлення та погодних умовах. Попередня обробка зображень. Зібрані зображення потребують попередньої обробки, щоб виділити корисну інформацію і зменшити шум. Цей етап може включати корекцію освітлення, фільтрацію шуму, вирівнювання контрасту та інші методи обробки зображень. Метою є отримання максимально чіткого та інформативного зображення для подальшого аналізу. Використання алгоритмів комп'ютерного зору. Після попередньої обробки зображень застосовуються алгоритми комп'ютерного зору. Це можуть бути як традиційні методи обробки зображень, такі як детекція країв, сегментація зображень та аналіз текстур, так і сучасні підходи на основі машинного навчання та глибоких нейронних мереж [6]. Традиційні методи, такі як детекція країв, дозволяють визначити межі між різними типами покриття, наприклад, між асфальтом і бетонною поверхнею. Аналіз текстур допомагає ідентифікувати особливості поверхні, такі як наявність тріщин, ям або інших дефектів. Сучасні підходи, засновані на глибоких нейронних мережах, дозволяють досягти більш високої точності розпізнавання. Вони навчаються на великих обсягах даних і можуть автоматично визначати різні типи дорожнього покриття, а також виявляти пошкодження. Такі моделі, як Convolutional Neural Networks (CNN) [7, 8], є особливо ефективними у завданнях класифікації зображень [9]. Класифікація та ідентифікація покриття. Після аналізу зображень алгоритми класифікують тип дорожнього покриття (асфальт, бетон, гравій тощо) та визначають його стан. Ця інформація може бути використана для різних цілей, таких як автоматичне керування транспортним засобом, оцінка стану дороги для ремонтних робіт або попередження водія про небезпечні ділянки дороги. Використання даних. Нарешті, класифіковані дані можуть бути використані різними системами транспортного засобу. Наприклад, системи автономного керування можуть

адаптувати свою поведінку залежно від типу дорожнього покриття, а системи безпеки можуть попереджати водія про можливі небезпеки. У підсумку, візуальний аналіз із використанням камер і алгоритмів комп'ютерного зору є ефективним підходом для розпізнавання типу дорожнього покриття. Ці технології допомагають підвищити безпеку на дорогах та сприяють розвитку автономних транспортних засобів.

Наступним варіантом є використання різних сенсорів, таких як акселерометри та гіроскопи. Ці сенсори вимірюють вібрації та інші фізичні властивості дорожнього покриття, що дозволяє системі визначати його тип.

Акселерометри вимірюють прискорення автомобіля в трьох напрямках: вперед-назад, вгору-вниз та з боку в бік. Коли автомобіль рухається по різних типах дорожнього покриття, його прискорення змінюється через різні рівні вібрацій та шорсткість поверхні. Наприклад, рух по гладкому асфальту буде створювати менші вібрації, ніж рух по гравійній дорозі. Аналізуючи ці вібрації, система може ідентифікувати тип покриття.

Гіроскопи вимірюють кутову швидкість обертання автомобіля навколо трьох осей. Вони корисні для виявлення нахилів і поворотів, що можуть бути важливими при аналізі дорожнього покриття. Наприклад, гіроскоп може допомогти розпізнати круті повороти або нерівності на дорозі.

Зазвичай, дані з акселерометрів та гіроскопів комбінуються для отримання більш точної картини стану дорожнього покриття. Цей процес включає в себе складні алгоритми обробки сигналів та машинного навчання. Спочатку дані з сенсорів збираються та обробляються для виділення характерних ознак, таких як амплітуда та частота вібрацій. Потім ці ознаки використовуються для навчання моделі, яка може розпізнавати різні типи покриття на основі зібраних даних.

Машинне навчання є ключовим елементом у цьому процесі. Воно дозволяє системі самостійно вчитися на основі великих обсягів даних та постійно покращувати можливості до розпізнавання дорожнього покриття. Типовий підхід включає використання нейронних мереж (рис. 1.1) або інших

алгоритмів машинного навчання, які навчаються на даних з реальних доріг. Після навчання, модель може швидко і точно класифікувати тип дорожнього покриття на основі нових даних.

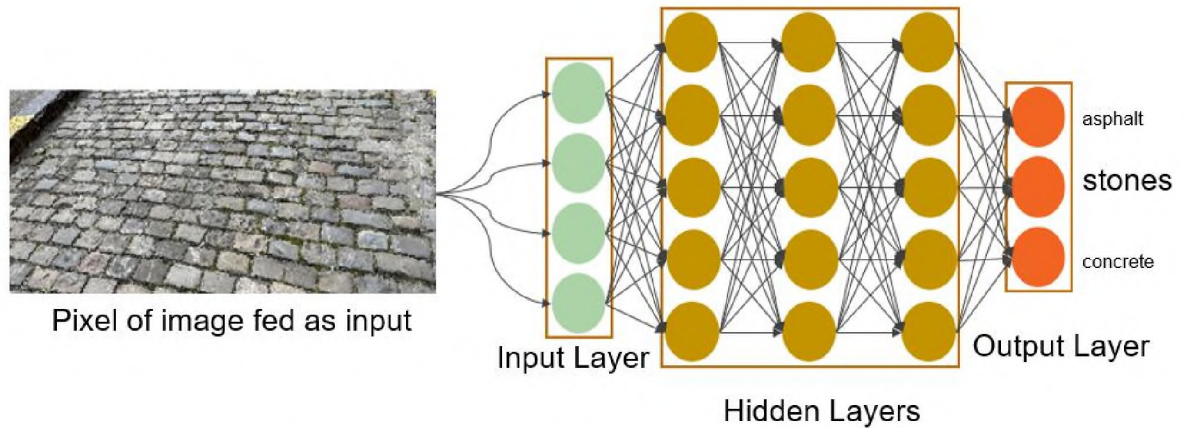


Рисунок 1.1 – Використання нейронних мереж для розпізнавання дорожнього покриття

Одним із методів, які набувають популярності, є використання акустичних методів, що базуються на аналізі звуків, які видаються транспортними засобами під час руху по різних типах покриття.

Акустичні методи розпізнавання типу дорожнього покриття базуються на використанні мікрофонів для збору звуків, які виникають під час взаємодії шин автомобіля з дорогою. Звуки, що виникають при їзді по асфальту, бетону, гравію або інших типах покриття, мають різні акустичні характеристики, які можна аналізувати і класифікувати. Мікрофони встановлюються на транспортному засобі або вздовж дороги для запису звуків. Звукові дані можуть бути зібрані в реальному часі з рухомо автомобіля. Шумозаглушення: видалення небажаних шумів, таких як вітер або інші транспортні засоби.

Нормалізація звукових сигналів для забезпечення стабільності аналізу. Використання методів спектрального аналізу, таких як швидке перетворення Фур'є (FFT), для розбиття звукових сигналів на частотні компоненти. Визначення характерних частотних діапазонів для кожного типу покриття. Застосування алгоритмів машинного навчання для класифікації типу

дорожнього покриття на основі аналізованих звукових даних. Використання тренувальних наборів даних для навчання моделей розпізнавання.

В цілому, такий підхід має кілька переваг: мікрофони є відносно недорогими та простими у встановленні; можливість отримання даних і аналізу їх у реальному часі; акустичні методи можуть бути менш чутливими до погодних умов порівняно з візуальними методами.

До недоліків слід віднести необхідність ефективного шумозаглушення для отримання точних даних; різні транспортні засоби можуть видавати різні звуки навіть на одному типі покриття, що ускладнює класифікацію, а також необхідність ретельного калібрування системи для різних умов і транспортних засобів.

Наведені підходи можуть забезпечувати високу точність у розпізнаванні типів покриття тільки за рахунок впровадження машинного навчання, що дозволяє постійно вдосконалювати відповідні системи. Тобто машинне навчання та штучний інтелект також відіграють ключову роль у розвитку цих технологій, дозволяючи точніше і ефективніше розпізнавання типів дорожнього покриття на основі накопичених даних. Це може бути особливо корисним у складних умовах, таких як змінні погодні умови або погано освітлені дороги.

Машинне навчання є частиною штучного інтелекту, що зосереджується на створенні алгоритмів, які дозволяють комп'ютерам навчатися на основі даних. У контексті розпізнавання дорожнього покриття ML використовується для аналізу зображень та інших даних, щоб класифікувати типи покриття.

Збір даних: Перший крок - це збір великої кількості зображень та інших даних про різні типи дорожнього покриття (асфальт, бетон, гравій, бруківка тощо). Ці дані можуть надходити від камер, датчиків лідарів, радарів тощо, встановлених на транспортних засобах. Анотація даних: Зібрані дані потрібно анотовано, тобто кожне зображення або зразок даних має бути позначено відповідним типом дорожнього покриття. Це робиться вручну або напіваавтоматично. Тренування моделей: Далі ці анотовані дані

використовуються для тренування моделей машинного навчання. Найчастіше використовуються нейронні мережі, зокрема CNN, які добре підходять для обробки зображень (рис. 1.2).

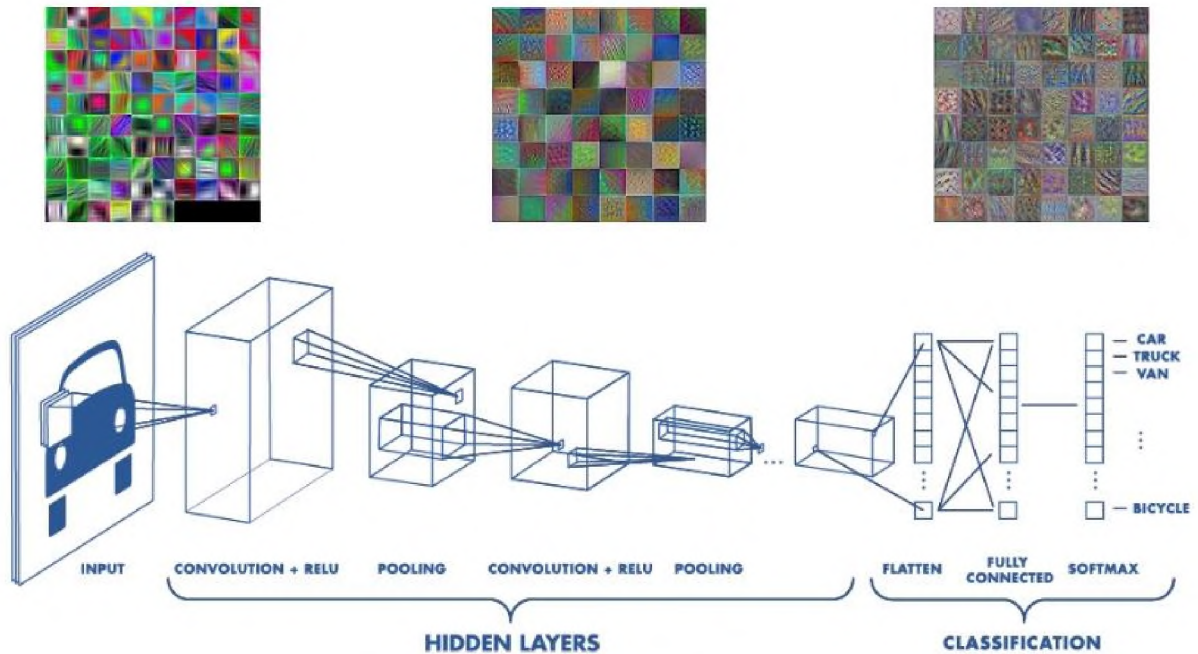


Рисунок 1.2 – Архітектура згорткової мережі

Валідація та тестування: Після тренування моделей, їх перевіряють на нових, не анотованих даних, щоб оцінити точність і надійність класифікації.

Розгортання: Коли модель проходить усі тести, її інтегрують у систему, яка працює в реальному часі на транспортних засобах, забезпечуючи розпізнавання типу дорожнього покриття під час руху.

Штучний інтелект охоплює ширший спектр технологій, які включають в себе машинне навчання, обробку природної мови, робототехніку тощо. У розпізнаванні типу дорожнього покриття AI може використовуватися для інтеграції даних з різних сенсорів, прийняття рішень і навіть прогнозування змін у дорожніх умовах. Основні аспекти включають:

Мультимодальне розпізнавання [10]. AI може об'єднувати дані з різних джерел, таких як камери, лідара та радарів, для більш точного розпізнавання. Наприклад, зображення з камери можуть доповнюватися

даними про відбиття світла від лідара, що дозволяє точніше визначити тип покриття.

Аналіз в реальному часі: AI-системи здатні обробляти великі обсяги даних в реальному часі, що є критично важливим для безпілотних транспортних засобів. Вони можуть не тільки розпізнавати тип покриття, але й оцінювати його стан (наприклад, наявність ям чи вибоїн).

Адаптація та навчання на ходу: Завдяки технологіям глибокого навчання, AI-системи можуть постійно вдосконалюватися, отримуючи нові дані під час роботи. Це дозволяє системі адаптуватися до нових типів покриття та умов, які не були передбачені під час первинного тренування.

Таким чином, наведені технології дозволяють наземним дронам та автопілотам автомобілів краще адаптуватися до умов дороги, що підвищує безпеку та комфорт для водія і пасажирів.

### **1.3 Особливості розпізнавання типу дорожнього покриття за допомогою нейронної мережі**

Розпізнавання типу дорожнього покриття є важливою задачею в контексті розвитку автономних транспортних систем, інтелектуального управління дорожнім рухом та покращення якості обслуговування доріг. Використання нейронних мереж для цієї мети має кілька переваг та практичну доцільність:

Нейронні мережі можуть автоматизувати процес розпізнавання типу дорожнього покриття, зменшуючи потребу у ручній праці та знижуючи ймовірність помилок, пов'язаних з людським фактором.

Сучасні нейронні мережі, такі як глибокі CNN, показують високу точність у розпізнаванні образів, що робить їх ідеальними для задач, пов'язаних з класифікацією дорожнього покриття. Нейронні мережі можуть працювати в реальному часі, що дозволяє використовувати їх у системах

автоматичного керування транспортними засобами та системах моніторингу дорожнього стану.

Нейронні мережі можуть бути навчені на великому різноманітті даних, що дозволяє їм ефективно розпізнавати різні типи дорожнього покриття в різних умовах (наприклад, мокра дорога, сніг, бруд тощо).

Результати розпізнавання можуть бути легко інтегровані з іншими системами, такими як навігаційні системи, системи прогнозування трафіку та управління дорожнім рухом.

Хоча початкові витрати на розробку та навчання нейронної мережі можуть бути високими, в довгостроковій перспективі автоматизовані системи розпізнавання можуть значно знизити експлуатаційні витрати, пов'язані з моніторингом та обслуговуванням доріг.

Використання нейронних мереж для розпізнавання типу дорожнього покриття є перспективним напрямком, який має потенціал значно покращити ефективність і надійність сучасних транспортних систем. Це дозволить підвищити рівень безпеки дорожнього руху, знизити витрати на обслуговування доріг та сприяти розвитку інтелектуальних транспортних систем.

На даний час, почали широко застосовувати великі візуальні та мовні моделі. Великі мовні моделі (LLM) [11], такі як GPT-4o [12], є потужними інструментами для обробки та аналізу текстової інформації. Проте, вони також можуть бути інтегровані з іншими технологіями, такими як комп'ютерний зір, для розширення їх можливостей [13]. LLM (Large Language Models) – це моделі машинного навчання, які призначені для обробки та генерації тексту. Вони використовують архітектури нейронних мереж, такі як трансформери, для навчання на великих обсягах текстових даних. Приклади LLM включають GPT-4o [12], BERT, RoBERTa тощо [14-17]. Зазвичай, вони призначені для розуміння, генерування та аналізу тексту. Використовують величезні обсяги текстових даних для навчання.

Використовуються в чат-ботах, системах рекомендацій, автоматичному перекладі, аналізі настроїв та інших завданнях, що вимагають обробки тексту. Моделі зору-мови (Vision-Language Models, VLM) – це моделі, які об'єднують обробку візуальної та текстової інформації (рис. 1.3) [18]. Вони здатні аналізувати зображення та текст одночасно, що дозволяє їм виконувати завдання, які вимагають інтеграції цих двох типів даних.

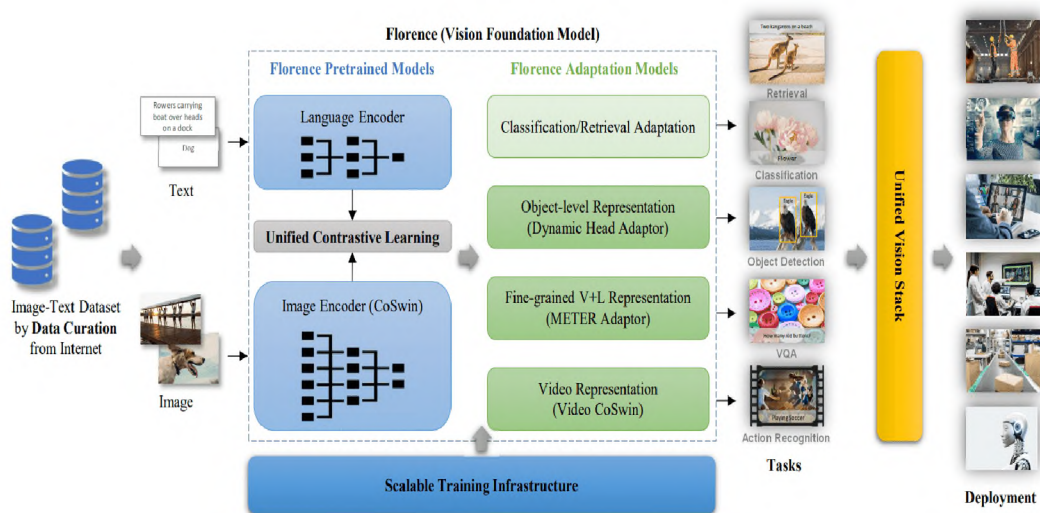


Рисунок 1.3. – Сутність реалізації VLM

Прикладом VLM є модель CLIP [19], розроблена OpenAI, яка може розуміти та співставляти зображення і текст. VLM об'єднують обробку зображень та тексту в одній моделі. Вони використовують дані, які містять як текст, так і зображення. Зазвичай, VLM застосовуються для завдань, таких як опис зображень, відповіді на запитання по зображенням, пошук зображень за текстовими запитами та інші мультимодальні задачі.

Основні відмінності між VLM та LLM полягають в тому, що LLM працюють виключно з текстовими даними і використовуються в задачах обробки природної мови (NLP), а VLM – як з текстовими, так і з візуальними даними в задачах, які вимагають інтеграції тексту та зображень (мультимодальні задачі). LLM навчаються на великих корпусах тексту, а VLM навчаються на даних, що містять як зображення, так і відповідний текст.

Дорожнє покриття може бути різним: асфальт, бетон, гравій, бруківка тощо. Кожен з цих типів має свої характеристики, які можуть впливати на зчеплення коліс, комфорт руху та загальну безпеку. Тому вміння точно визначати тип покриття є важливим для систем автономного водіння, навігаційних систем та управління транспортними потоками.

Обидва типи моделей мають свої унікальні переваги та застосування. LLM спеціалізуються на обробці та генерації тексту, тоді як VLM можуть інтегрувати та аналізувати як візуальну, так і текстову інформацію, що робить їх корисними для ширшого спектру завдань.

## РОЗДІЛ 2

# СИНТЕЗ МОДЕЛІ РОЗПІЗНАВАННЯ ТИПУ ДОРОЖНЬОГО ПОКРИТТЯ НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ

### 2.1 Визначення етапів розпізнавання типу дорожнього покриття за допомогою нейронної мережі

Нейронна мережа може розрізняти різні типи дорожнього покриття, вивчаючи характерні закономірності даних датчиків, зібраних з транспортних засобів, що рухаються по різних поверхнях. Зазвичай, для виконання таких завдань використовується CNN. Вона вчиться розпізнавати унікальні сигнатури вібрації та руху кожного типу поверхні за даними акселерометра та гіроскопа. Наприклад, гравій може виробляти високочастотні вібрації, в той час як пісок може мати чіткий профіль руху з низькою швидкістю. При цьому може здійснюватися оцінка коефіцієнта тертя для контролю динаміки автомобіля; визначитись дорожні умови для інтелектуальних транспортних систем; здійснюватися моніторинг дорожньої інфраструктури шляхом виявлення дефектів поверхні.

Поєднуючи дані з декількох датчиків, таких як акселерометри, гіроскопи та датчики швидкості, CNN може створити надійне уявлення про кожен тип поверхні. У процесі навчання мережа вчиться зіставляти дані датчика з правильним класом поверхні. Їх в управлінні динамікою транспортних засобів, моніторингу транспортної інфраструктури та автономному керуванні.

Навчання нейронної мережі для класифікації дорожніх покриттів пов'язане з кількома проблемами, які дослідники вирішили у цій галузі. Отримання різноманітного та репрезентативного набору даних датчиків дорожнього покриття може бути складним завданням. Точне маркування цих даних різних типів поверхонь має вирішальне значення для ефективного навчання нейронної мережі. Вилучення релевантних ознак даних датчиків,

які ефективно розрізняють різні дорожні покриття, має важливе значення. Розробка алгоритмів для попередньої обробки та вилучення значущих ознак з необроблених даних датчиків може бути складним завданням.

Розробка архітектури нейронної мережі, здатної ефективно вивчати складні закономірності та варіації даних про дорожнє покриття, є складним завданням. Балансування складності моделі, щоб уникнути перенавчання або недонавчання, має вирішальне значення для точної класифікації.

Інтеграція даних з кількох датчиків, таких як акселерометри, гіроскопи та камери для отримання повного представлення характеристик дорожнього покриття може бути складним завданням. Ключовим моментом є ефективне об'єднання даних датчиків підвищення точності класифікації.

Забезпечення того, щоб навчена нейронна мережа могла добре узагальнювати невидимі умови та варіації дорожнього покриття, є серйозною проблемою. Стійкість до різних умов навколишнього середовища, освітлення та шуму має вирішальне значення для реальних застосувань.

Розуміння того, як нейронна мережа приймає рішення, та здатність інтерпретувати її класифікації важливі, особливо в критично важливих для безпеки додатках, таких як автономне керування. Забезпечення прозорості та інтерпретованості рішень моделі є складним завданням.

Вирішення цих проблем за допомогою інноваційних методів збору даних, методів конструювання ознак, архітектур моделей та надійних стратегій навчання має важливе значення для успішного розгортання нейронних мереж для точної та надійної класифікації дорожніх покриттів.

Згідно п. 1.2, на даний час, найбільшого поширення отримав візуальний спосіб розпізнавання [20]. Тому варто зупинитись на ньому більш детально.

Тип дорожнього покриття може бути ефективно розпізнаний за допомогою глибоких нейронних мереж. Ці методи машинного навчання були успішно застосовані для класифікації дорожніх покриттів за категоріями, такими як гладкі, вибоїсті, шорсткі, асфальт, бетон та ін. На основі проведених досліджень можна визначити, що процес розпізнавання типу

дорожнього покриття за допомогою нейронної мережі включає кілька ключових етапів: збирання та підготовка даних, розробка та навчання моделі, тестування та валідація, а також розгортання та використання моделі.

Збирання та підготовка даних. Для ефективного навчання нейронної мережі потрібен великий набір даних зображень різних типів дорожнього покриття. Збирання даних включає зйомку зображень та анотацію даних. Для першого кроку використовуються камери, встановлені на транспортних засобах або дронах та інших безпілотних засобах, для зйомки дорожнього покриття. Другий крок – зібрані зображення розмічаються вручну або за допомогою напівавтоматичних інструментів, щоб вказати тип дорожнього покриття на кожному зображенні (наприклад, асфальт, бетон, гравій тощо).

Попередня обробка даних. Зображення повинні бути попередньо оброблені перед тим, як подаватися на вхід нейронної мережі:

- зміна розміру – зображення змінюються до стандартного розміру, який підходить для обраної моделі нейронної мережі;
- нормалізація – піксельні значення зображень нормалізуються для прискорення процесу навчання;
- аугментація даних – для збільшення різноманітності тренувального набору зображень використовуються техніки аугментації, такі як обертання, зміна яскравості, контрасту тощо (для цього можна застосовувати певні техніки та інструменти, що дозволяють автоматизувати цей процес).

Розробка та навчання моделі. Обирається відповідна архітектура нейронної мережі, найчастіше використовуються CNN [21], які добре підходять для задач розпізнавання образів. Процес навчання включає також кілька наступних кроків.

По-перше, набір даних розділяється на тренувальний, валідаційний і тестовий набори – розподіл даних.

По-друге, модель навчається на тренувальному наборі даних шляхом оптимізації вагових коефіцієнтів за допомогою алгоритму зворотного поширення помилки (backpropagation) – навчання моделі.

По-третє, під час навчання модель перевіряється на валідаційному наборі даних для оцінки її продуктивності та запобігання перенавчанню – валідація.

Тестування та валідація. Після навчання модель тестується на незалежному тестовому наборі даних для оцінки її продуктивності. Використовуються різні метрики, такі як повнота (recall), точність (precision) та F1-міра.

На решті, розгортання та використання моделі. Коли модель успішно навчена та протестована, вона може бути розгорнута в реальних умовах. Тобто, модель інтегрується з камерами та іншими сенсорами, встановленими на транспортних засобах або дорожніх інфраструктурах. Також необхідна відповідність вимогам щодо роботи в реальному часі, аналізуючи зображення дорожнього покриття та класифікуючи їх відповідно до типу покриття.

Таким чином, розпізнавання типу дорожнього покриття за допомогою нейронних мереж є складним, але ефективним процесом, який включає кілька етапів від збору та підготовки даних до розгортання та використання моделі. Використання сучасних технік глибокого навчання дозволяє досягти високої точності та надійності в розпізнаванні різних типів дорожнього покриття, що сприяє покращенню безпеки та ефективності дорожнього руху.

Розглянутий підхід має перспективне застосування в керуванні транспортними засобами, транспортною інфраструктурою та автономному керуванні.

## **2.2 Підготовка даних та створення датасету**

На даний час, декілька наборів даних можна використовувати для навчання нейронних мереж класифікації дорожнього покриття.

1. Набір даних KITTI [22]. Він містить зображення різних дорожніх умов, включаючи сухі дороги, мокрі дороги та снігові дороги, які

використовуються для навчання моделей глибокого навчання для класифікації стану дорожнього покриття.

2. Oxford RobotCar Dataset [23] включає зображення та дані датчиків транспортного засобу, що рухається різними дорожніми покриттями, які можна використовувати для навчання нейронних мереж класифікації дорожнього покриття.

3. Набір даних Vision Meets Robotics [24] містить зображення різних дорожніх умов, включаючи сухі дороги, мокрі дороги та снігові дороги, які використовуються для навчання моделей глибокого навчання для класифікації стану дорожнього покриття.

4. Набір даних класифікації стану дорожнього покриття включає зображення сухих, мокрих доріг та снігових доріг, які використовуються для навчання CNN для наскрізного розпізнавання стану дорожнього покриття.

5. Набір даних про дорожнє покриття гравію та асфальту містить зображення гравійних та асфальтових дорожніх покриттів, які використовуються для навчання CNN класифікації дорожніх покриттів.

Ці набори даних з відкритим вихідним кодом є гарною відправною точкою для дослідників та розробників для навчання нейронних мереж для завдань класифікації дорожнього покриття. Вони охоплюють цілий ряд типів та умов дорожнього покриття, що дозволяє розробляти надійні та точні моделі. Ці набори даних використовуються для навчання нейронних мереж розпізнавати закономірності даних про дорожнє покриття і класифікувати їх за різними категоріями, такими як асфальт, бетон, гравій, пісок та ін. Акустичні дані можуть значно підвищити точність класифікації дорожнього покриття за рахунок використання унікальних патернів та характеристик звукових хвиль, що генеруються різними дорожніми покриттями. Акустичні сигнали, що уловлюються звуковими датчиками, можуть виявити взаємодію між шиною та дорожнім покриттям, що має вирішальне значення для розрізнення різних дорожніх умов. На цю взаємодію впливають такі фактори, як текстура дорожнього покриття, вологість та температура, що робить його

цінним індикатором для класифікації дорожнього покриття. Аналіз частотного складу акустичних сигналів допоможе виявити конкретні закономірності, пов'язані з різними дорожніми покриттями. Наприклад, мокра дорога, як правило, виробляє більш високочастотні сигнали через підвищене тертя між шиною та дорогою, тоді як суха дорога виробляє сигнали нижчої частоти.

Вивчення характеристик акустичних сигналів у часовій області може надати додаткову інформацію про стан дорожнього покриття. Наприклад, тривалість і амплітуда звукових імпульсів можуть свідчити про шорсткість або гладкість дорожнього покриття.

Об'єднання акустичних даних з іншими даними датчиків, такими як зображення з камер або дані акселерометра з транспортних засобів, може підвищити загальну точність класифікації дорожнього покриття. Таке злиття може допомогти отримати більш повне уявлення про дорожнє покриття, включаючи візуальні та тактильні характеристики.

Акустичні дані можна використовувати як вхідні дані для моделей нейронної мережі, для класифікації дорожнього покриття. Ці моделі можуть навчитися розпізнавати закономірності в акустичних сигналах і прогнозувати тип дорожнього покриття, підвищуючи точність класифікації. Акустичні дані можуть оброблятися у режимі реального часу, що дозволяє миттєво виявляти та класифікувати стан дорожнього покриття. Це особливо важливо для таких програм, як автономні транспортні засоби, де своєчасна та точна інформація про дорожні умови має вирішальне значення для безпечної навігації. Використовуючи акустичні дані в системах класифікації дорожнього покриття, дослідники та розробники можуть створювати більш надійні та точні моделі, які краще відображають складність стану дорожнього покриття.

Аналіз наявних джерел, свідчить немає явної згадки про набори даних, які включають як візуальні, так і акустичні дані для дорожніх покриттів. Згадані набори даних в основному зосереджені на візуальних даних, таких як зображення дорожніх покриттів, знятих у різних умовах, та докладні

інструкції, пов'язані з властивостями матеріалу, тертя та нерівностей. Для наборів даних, які поєднують візуальні та акустичні дані про дорожнє покриття, може знадобитися подальше вивчення за межами наданих джерел для визначення конкретних наборів даних, що охоплюють обидва типи інформації. Для вирішення завдання розпізнавання типу дорожнього покриття було підготовлено спеціальний датасет. Він містить набір зображень дорожнього покриття, знятих у різних країнах у різний час року та пора доби. Всю множину зображень було поділено на 2 класи – бруківка та асфальт. Для збільшення кількості знімків датасета було використано генерацію зображень на основі текстових описів за допомогою нейромережі Open AI DALL-E [25]. Для цього використовувалася програма, написана мовою Python [26] у Google CoLab [27]. Її код та деякі результати запуску наведено на рис. 2.1 та 2.2, відповідно.

```

%pip install openai
import os
import openai
openai.api_key = "sk-YE..."
response = openai.Image.create(
    prompt="Площадь покрыта очень плотно уложенными и гладкими камнями",
    n=1,
    size="1024x1024"
)
image_url = response['data'][0]['url']
print(image_url)

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: openai in /usr/local/lib/python3.8/dist-packages (0.26.1)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.8/dist-packages (from openai) (3.8.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.8/dist-packages (from openai) (4.64.1)
Requirement already satisfied: requests>=2.20 in /usr/local/lib/python3.8/dist-packages (from openai) (2.25.1)
Requirement already satisfied: chardet<5,>=3.0.2 in /usr/local/lib/python3.8/dist-packages (from requests>=2.20->openai) (4.0.0)
Requirement already satisfied: certifi=2017.4.17 in /usr/local/lib/python3.8/dist-packages (from requests>=2.20->openai) (2022.12.7)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /usr/local/lib/python3.8/dist-packages (from requests>=2.20->openai) (1.24.3)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.8/dist-packages (from requests>=2.20->openai) (2.10)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->openai) (22.2.0)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.8/dist-packages (from aiohttp->openai) (6.0.4)
Requirement already satisfied: charset-normalizer<3.0,>=2.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->openai) (2.1.1)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.8/dist-packages (from aiohttp->openai) (1.3.3)
Requirement already satisfied: async-timeout<5.0,>=4.0.0a3 in /usr/local/lib/python3.8/dist-packages (from aiohttp->openai) (4.0.2)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.8/dist-packages (from aiohttp->openai) (1.3.1)
Requirement already satisfied: yarl<2.0,>=1.0 in /usr/local/lib/python3.8/dist-packages (from aiohttp->openai) (1.8.2)
https://oaidalleapiprodscus.blob.core.windows.net/private/org-hxP9YwYk4VKMEhbFYUg7rws6/user-SvETL0QfZnLa449fnW1yS4J/img-psiXimV4AQFLWe9GV5O2BUC
  
```

9 сек. выполнено в 11:48

Рисунок 2.1 – Код програми генерації зображень через API OpenAI

Для генерації зображень використовувався prompt: «Площа покрита дуже щільно покладеним і гладким камінням».

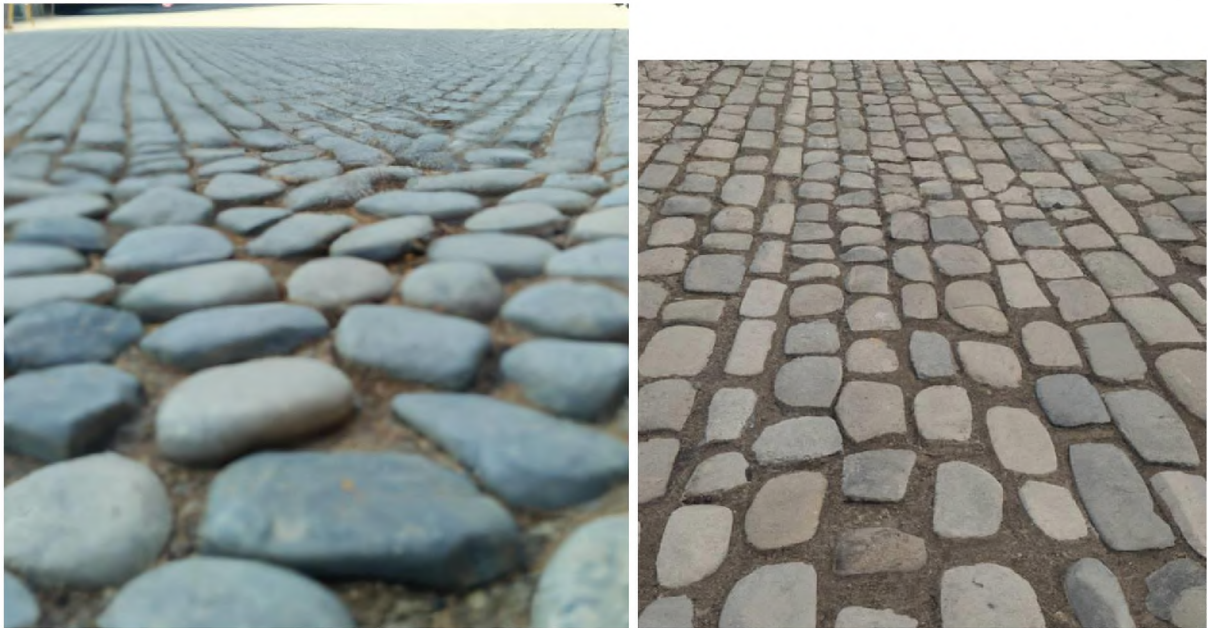


Рисунок 2.2 – Приклади зображень бруківки, що згенеровані за допомогою нейронної мережі DALL-E.

Надалі для поповнення датасету застосовувалася генерація зображень за допомогою DALL-E3, доступ до якої надавалася через платну версію LLM GPT-4. При цьому в остаточному варіанті датасета клас Asphalt містив 101 фото, з яких для тесту було виділено 10 файлів. Розмір класу Stones складає 114 фотографій, з яких для тесту виділено 11 файлів.

Таким чином, загальний розмір бази становив 215 фотознімків, з яких виділено для навчання 194 знімки, а для тесту – 21.

На початковому етапі реалізації нейронної мережі для розпізнавання типу дорожнього покриття було вибрано деякі значення параметрів та гіперпараметрів. Так, для навчання нейромережі було встановлено батч розміром 8 зображень, кількість епох навчання – 100. В якості базового оптимізатору виступає Adam [28]. Усі зображення датасету були перетворені на однаковий розмір. При цьому на першому етапі досліджень він був встановлений  $680 \times 440$  пікселів.

### 2.3 Синтез нейронної мережі

Вся архітектура моделі та процес навчання націлені на розпізнавання типу дорожнього покриття з використанням глибокої CNN [29]. Модель використовує кілька шарів згортки, нормалізації та відсіву для ефективного навчання та класифікації зображень. Враховуючі існуючу загальну практику, розробка та пошук оптимальної архітектури та параметрів нейронної мережі може здійснювати у наступній послідовності:

- імпорт бібліотек – імпортуються всі необхідні бібліотеки для роботи з нейронною мережею та обробки зображень (рис. 2.3);

```
import numpy as np
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Conv2D, MaxPooling2D,
Flatten, Dropout, BatchNormalization
from tensorflow.keras.optimizers import Adam
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
import gdown
import os
import matplotlib.pyplot as plt
%matplotlib inline
```

Рисунок 2.3 – Бібліотеки, що забезпечують основні функції для роботи з масивами, обробки зображень, побудови та навчання нейронних мереж, відображення результатів

- гіперпараметри – встановлюються значення гіперпараметрів для аугментації даних, навчання та моделі;
- підготовка даних – налаштовується генератор зображень для тренувального та валідаційного наборів даних з аугментацією;
- визначення моделі – створюється CNN (або інша нейронна мережа)) з кількома шарами;
- компіляція та навчання – модель компілюється та навчається на тренувальних даних;

– графік точності – будується графік точності навчання та валідації для візуалізації продуктивності моделі.

На першому етапі необхідно завантажити потрібні бібліотеки для роботи з даними у форматі DataFrame, з числовими масивами та математичними операціями, створення графіків, візуалізацій даних, створення послідовної моделі нейронної мережі, генерації зображень у процесі навчання, а також імпорту шарів нейронної мережі (щільні – Dense, згорткові – Conv2D, максимального пулінгу – MaxPooling2D, вирівнювання – Flatten, відсічення – Dropout).

Наступним етапом при створенні коду необхідно налаштувати параметри нейронної мережі. Ці параметри визначають шляхи до даних, розміри зображень, параметри аугментації та параметри навчання (рис. 2.4).

```

TRAIN_PATH = '/content/cars'
TEST_PATH = '/content/cars_test'
TEST_SPLIT = 0.1
VAL_SPLIT = 0.2
IMG_WIDTH = 320
IMG_HEIGHT = 240
IMG_CHANNELS = 3

ROTATION_RANGE = 8
WIDTH_SHIFT_RANGE = 0.15
HEIGHT_SHIFT_RANGE = 0.15
ZOOM_RANGE = 0.15
BRIGHTNESS_RANGE = (0.7, 1.3)
HORIZONTAL_FLIP = True

EPOCHS = 100
BATCH_SIZE = 8

```

Рисунок 2.4 – Код для налаштування параметрів нейронної мережі

На рис. 2.5 наведений фрагмент коду, що налаштовує генератор зображень для тренувального та валідаційного наборів даних, що використовує аугментацію. За рахунок штучного розширення наявного набору даних зображень шляхом створення модифікованих версій існуючих зображень можна сформувати більший та різноманітніший набір даних для

тренування моделі, що може покращити її точність та здатність до узагальнення на нові дані. Це досягається шляхом застосування різних трансформацій, таких як:

- обертання – поворот зображення на різні кути (наприклад, 90°, 180°, 270°);
- зміна розміру – зменшення або збільшення розміру зображення;
- обрізка – відрізання частин зображення (наприклад, обрізка країв);
- перевертання – дзеркальне відображення по горизонталі або вертикалі;
- зміна яскравості та контрасту – регулювання рівня яскравості та контрасту зображення.

```

datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=ROTATION_RANGE,
    width_shift_range=WIDTH_SHIFT_RANGE,
    height_shift_range=HEIGHT_SHIFT_RANGE,
    zoom_range=ZOOM_RANGE,
    brightness_range=BRIGHTNESS_RANGE,
    horizontal_flip=HORIZONTAL_FLIP,
    validation_split=VAL_SPLIT
)
train_generator = datagen.flow_from_directory(
    TRAIN_PATH,
    target_size=(IMG_HEIGHT, IMG_WIDTH),
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='training'
)
validation_generator = datagen.flow_from_directory(
    TRAIN_PATH,
    target_size=(IMG_HEIGHT, IMG_WIDTH),
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='validation'
)

```

Рисунок 2.5 – Підготовка даних за допомогою генератора зображень

Надалі розробляється код, що створює послідовну модель з кількома шарами CNN, шарами нормалізації, шарами відсіву (Dropout) та повнозв'язними (Dense) шарами (рис. 2.6).

```

model_conv = Sequential()
model_conv.add(Conv2D(256, (3, 3), padding='same', activation='relu',
input_shape=(IMG_HEIGHT, IMG_WIDTH, IMG_CHANNELS)))
model_conv.add(BatchNormalization())
model_conv.add(Conv2D(256, (3, 3), padding='same', activation='relu'))
model_conv.add(MaxPooling2D(pool_size=(3, 3)))
model_conv.add(Conv2D(256, (3, 3), padding='same', activation='relu'))
model_conv.add(BatchNormalization())
model_conv.add(Dropout(0.2))
model_conv.add(Conv2D(256, (3, 3), padding='same', activation='relu'))
model_conv.add(MaxPooling2D(pool_size=(3, 3)))
model_conv.add(Dropout(0.2))
model_conv.add(Conv2D(512, (3, 3), padding='same', activation='relu'))
model_conv.add(BatchNormalization())
model_conv.add(Conv2D(1024, (3, 3), padding='same', activation='relu'))
model_conv.add(MaxPooling2D(pool_size=(3, 3)))
model_conv.add(Dropout(0.2))
model_conv.add(Flatten())
model_conv.add(Dense(2048, activation='relu'))
model_conv.add(Dense(240, activation='relu'))
model_conv.add(Dense(CLASS_COUNT, activation='softmax'))

```

Рисунок 2.6 – Створення послідовної моделі CNN

Надалі формується код, що компілює модель з використанням оптимізатора Adam і функції втрат `categorical_crossentropy`, навчає модель на тренувальних даних та оцінює її на тестових даних.

```

model_conv.compile(optimizer=Adam(), loss='categorical_crossentropy',
metrics=['accuracy'])

history = model_conv.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // BATCH_SIZE,
    validation_data=validation_generator,
    validation_steps=validation_generator.samples // BATCH_SIZE,
    epochs=EPOCHS
)
model_conv.evaluate(test_generator)

```

Рисунок 2.7 – Компіляція, навчання та оцінка моделі

Варіант наведеного коду призначений для виконання у хмарному сервісі Google Collab. Якщо існує необхідність його реалізації у локальному середовищі, потрібно встановити бібліотеку TensorFlow [30], наприклад, використовуючи пакет `pip`, виконати команду:

```
pip install tensorflow
```

Крім цього потрібно оновити шляхи до даних. Для треба вказати правильні шляхи до ваших тренувальних та тестових даних у змінних TRAIN\_PATH та TEST\_PATH.

Таким чином, вся архітектура моделі та процес навчання націлені на розпізнавання типу дорожнього покриття з використанням глибокої CNN. Модель використовує кілька шарів згортки, нормалізації та відсіву для ефективного навчання та класифікації зображень.

## РОЗДІЛ 3

### РЕКОМЕНДАЦІ ЩОДО РЕАЛІЗАЦІЇ РОЗПІЗНАВАННЯ ТИПУ ДОРОЖНЬОГО ПОКРИТТЯ ЗА ДОПОМОГОЮ НЕЙРОННОЇ МЕРЕЖІ

#### 3.1 Оцінка точності розпізнавання типу дорожнього покриття нейронною мережею

Використовуючи певну модифікацію синтезованого код, що наведений у Додатку А, в хмарному сервісі Google Colab виконаний його запуск. При цьому було досягнуто точності класифікації 100 % на тестовій вибірці та 76 % – на валідаційній. При цьому, використовувалася платформа Google Colab Pro+. На відміну від безкоштовної версії, це дозволило підключити для розрахунків графічний прискорювач (GPU) від NVIDIA A100 з 40 ГБ графічної пам'яті [31]. Крім того, для проведення обчислень представлялося 166,8 ГБ дискового простору та 83,5 ГБ ОЗП процесорного модуля. Структура моделі CNN (рис. 3.1) для обробки зображень формату  $680 \times 440$  пікселів включає 836070658 (3.11 GB) параметрів, з яких навчанням піддавались 836068610 (3.11 GB) і тільки 2048 параметрів не підлягали зміні (8.00 KB) через використання шару Dropout.

```

Model: "sequential"
-----
Layer (type)                                     Output Shape                                     Param #
-----
conv2d (Conv2D)                                  (None, 480, 640, 256)                          7168
batch_normalization (Batch Normalization)        (None, 480, 640, 256)                          1024
conv2d_1 (Conv2D)                                 (None, 480, 640, 256)                          590080
max_pooling2d (MaxPooling2D)                    (None, 160, 213, 256)                           0
conv2d_2 (Conv2D)                                 (None, 160, 213, 256)                          590080
batch_normalization_1 (Batch Normalization)      (None, 160, 213, 256)                          1024
dropout (Dropout)                                (None, 160, 213, 256)                           0
conv2d_3 (Conv2D)                                 (None, 160, 213, 256)                          590080
max_pooling2d_1 (MaxPooling2D)                  (None, 53, 71, 256)                             0
dropout_1 (Dropout)                              (None, 53, 71, 256)                             0
conv2d_4 (Conv2D)                                 (None, 53, 71, 512)                             1180160
batch_normalization_2 (Batch Normalization)     (None, 53, 71, 512)                             2048
conv2d_5 (Conv2D)                                 (None, 53, 71, 1024)                            4719616
max_pooling2d_2 (MaxPoolin g2D)                 (None, 17, 23, 1024)                            0
dropout_2 (Dropout)                              (None, 17, 23, 1024)                            0
flatten (Flatten)                                (None, 400384)                                    0
dense (Dense)                                    (None, 2048)                                      819988480
dense_1 (Dense)                                  (None, 4096)                                      8392704
dense_2 (Dense)                                  (None, 2)                                         8194
-----
Total params: 836070658 (3.11 GB)
Trainable params: 836068610 (3.11 GB)
Non-trainable params: 2048 (8.00 KB)

```

Рисунок 3.1 – Структура моделі CNN для обробки зображень  
формату  $680 \times 440$  пікселів

Для навчання було задано 10 епох, що дозволило одержати на тестовій вибірці точність класифікації 73,68 % (рис. 3.2). На рис. 3.3 наведено графік помилок навчання.

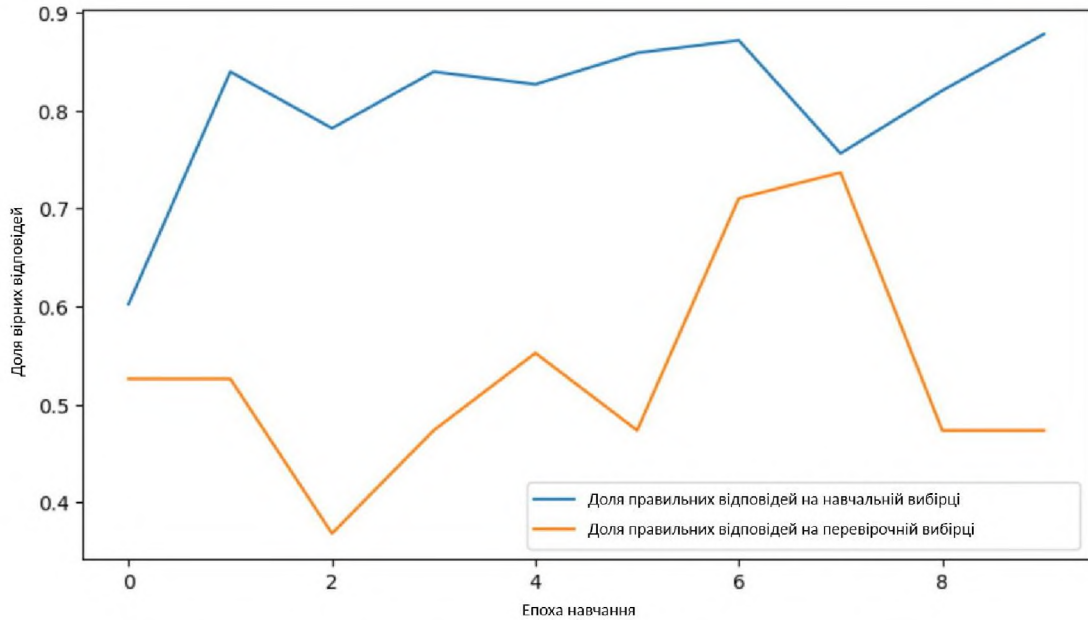


Рисунок 3.2 – Графіки точності навчання на 10 епохах

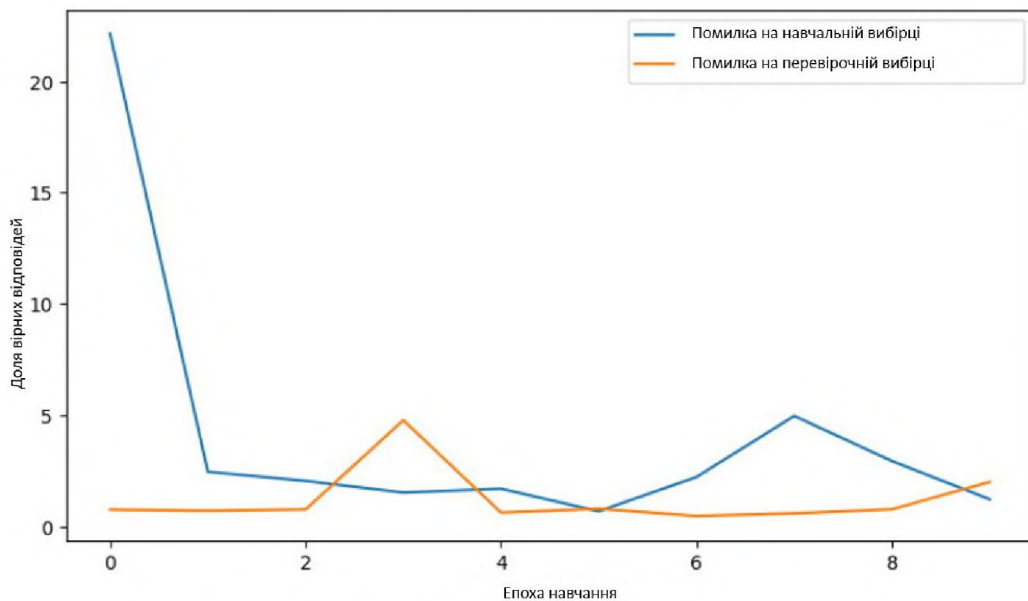


Рисунок 3.3 – Графіки помилок навчання

У процесі навчання нейронної мережі було задіяно 32,6 ГБ із 40 ГБ ОЗП графічної карти, 27.2 ГБ із 166.8 ГБ дискового простору, а також 5 ГБ із 83.5 ГБ ОЗП процесорного модуля.

На другому етапі досліджень усі знімки датасету були приведені до єдиного розміру зображення  $340 \times 220$  пікселів. Крім того, були внесені деякі зміни до структури моделі нейронної мережі, що спрямовані на підвищення її ефективності. Зокрема, в передостанньому шарі Dense кількість нейронів була зменшена з 4096 до 240. Це дозволило знизити вимоги до обчислювальних ресурсів, які необхідні для навчання та функціонування нейронної мережі. Таким чином, результуюча структура моделі нейронної мережі містить 192724946 параметрів і має розмір 735.19 МВ (рис. 3.4).

```

Model: "sequential_1"
Layer (type)                               Output Shape                                Param #
=====
conv2d_6 (Conv2D)                           (None, 240, 320, 256)                    7168
batch_normalization_3 (BatchNormalization)  (None, 240, 320, 256)                    1024
conv2d_7 (Conv2D)                           (None, 240, 320, 256)                    590080
max_pooling2d_3 (MaxPooling2D)             (None, 80, 106, 256)                    0
conv2d_8 (Conv2D)                           (None, 80, 106, 256)                    590080
batch_normalization_4 (BatchNormalization)  (None, 80, 106, 256)                    1024
dropout_3 (Dropout)                        (None, 80, 106, 256)                    0
conv2d_9 (Conv2D)                           (None, 80, 106, 256)                    590080
max_pooling2d_4 (MaxPooling2D)             (None, 26, 35, 256)                    0
dropout_4 (Dropout)                        (None, 26, 35, 256)                    0
conv2d_10 (Conv2D)                         (None, 26, 35, 512)                    1180160
batch_normalization_5 (BatchNormalization)  (None, 26, 35, 512)                    2048
conv2d_11 (Conv2D)                         (None, 26, 35, 1024)                   4719616
max_pooling2d_5 (MaxPooling2D)             (None, 8, 11, 1024)                    0
dropout_5 (Dropout)                        (None, 8, 11, 1024)                    0
flatten_1 (Flatten)                        (None, 90112)                            0
dense_3 (Dense)                            (None, 2048)                            184551424
dense_4 (Dense)                            (None, 240)                              491760
dense_5 (Dense)                            (None, 2)                                482
=====
Total params: 192724946 (735.19 MB)
Trainable params: 192722898 (735.18 MB)
Non-trainable params: 2048 (8.00 KB)

```

Рисунок 3.4 – Модифікація структура моделі CNN для обробки зображень формату  $320 \times 220$  пікселів

Таким чином, скорочення розмірів зображення у 4 рази ( $2 \times 2$ ) дозволило зменшити розмір моделі нейронної мережі більш ніж у 4 рази. При цьому, час навчання однієї епохи при тому ж батчі 8 скоротилося приблизно в 2 рази (з 28 до 13 с), тобто пропорційно квадратному кореню з величини зміни розмірів нейронної мережі.

При переході до зменшеного формату зображення обсяг затребуваної навчання оперативної пам'яті графічного процесора скоротився до 16,6 ГБ. При цьому потрібні ресурси дискового простору Google Colab Pro+ і ОЗП процесора залишилися в тих же обсягах, що і при розмірі фотографії  $640 \times 480$  пікселів.

Через малий обсяг датасету (215 фото) у процесі навчання було досягнуто точність 100 % на тестовій вибірці. Результати використання навченої нейронної мережі за валідаційною вибіркою представлені нормалізованою матрицею помилок на рис. 3.5. Графік зміни точності за тренувальною та тестовою вибірками наведено на рис. 3.6, а на рис. 3.7 – графік помилок навчання.

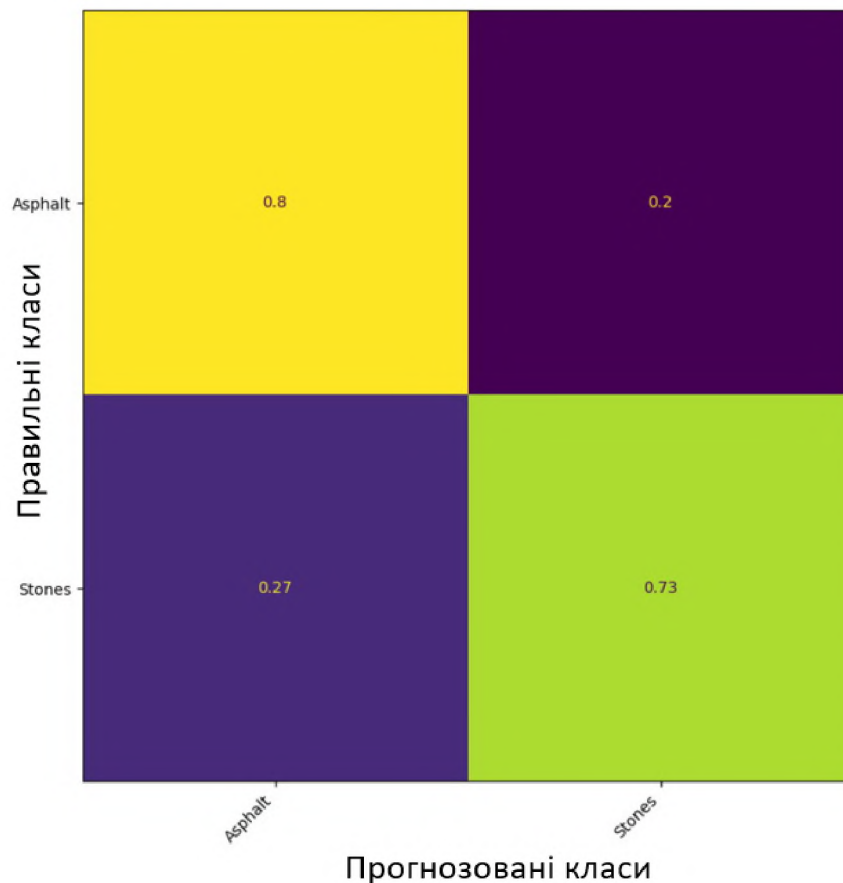


Рисунок 3.5 – Нормалізована матриця помилок

На наступному етапі досліджень було проведено вдосконалення датасету шляхом збільшення кількості фотознімків для обох класів, а також

обрізки вмісту деяких кадрів з метою видалення непотрібних об'єктів. Як приклад на рис. 3.8 представлений вихідний фотознімок та його модифікація, використана в новому датасеті.

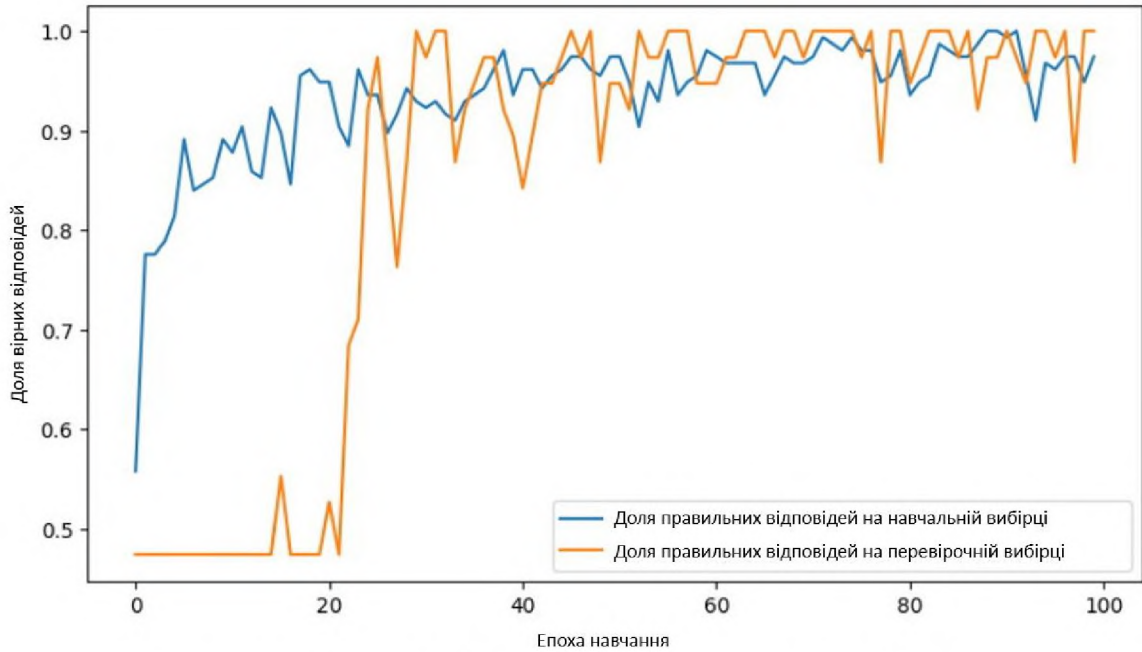


Рисунок 3.6 – Графік точностей навчання за 100 епох

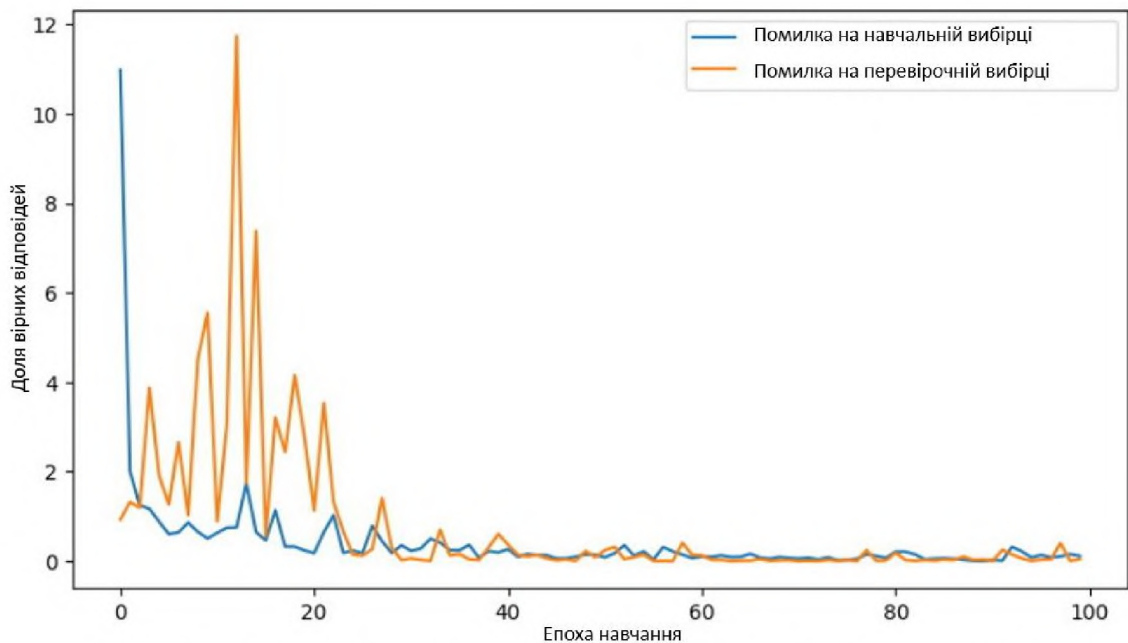


Рисунок 3.7 – Помилки навчання, що відповідають рис. 3.6

У модифікованому датасеті розмір класу Asphalt складає 189 фото, з яких для тесту було виділено 18 зображень, а клас Stones включає 185 фото, з

такою ж кількістю тестових файлів – 18. Загальний розмір бази зображень зріс до 374, при цьому для навчання було виділено 338 знімків дорожнього покриття, а для тесту – 36. Приклад аугментації вихідних зображень датасета наведений на рис. 3.9.



Рисунок 3.8 – Приклад модифікації зображень датасета шляхом їх обрізання для видалення непотрібних об'єктів

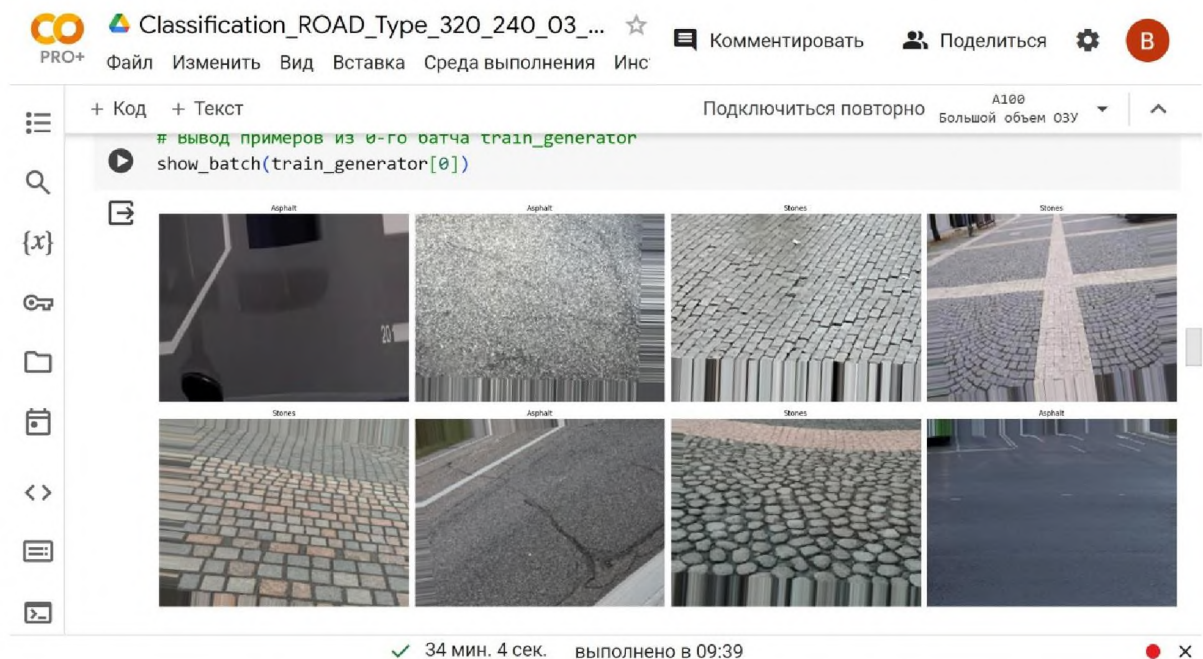


Рисунок 3.9 – Приклад аугментації вихідних зображень датасету спеціальним алгоритмом

Збільшення розміру датасета призвело до зростання часу навчання, яке при тому ж батчі 8 зросло з 13 до 20 секунд на одну епоху. Водночас, розмір датасету незначно позначився на обсязі задіяних обчислювальних ресурсів. У

процесі навчання було задіяно 27.7 GB дискового простору, 5.8 GB ОЗП процесора і 32.6 GB ОЗП GPU. У процесі навчання на 100 епохах було досягнуто максимальної точності на тестовій вибірці 97,01 %. Перевірка на тестовій вибірці, приклад якої показано на рис. 3.10 показала середню точність розпізнавання класів 92 %. Відповідну матрицю помилок представлено на рис. 3.11. При цьому найкраща точність була отримана на асфальтовому дорожньому покриття: неймерережа правильно віднесла до класу Asphalt 100 % відповідних представлених фотографій.

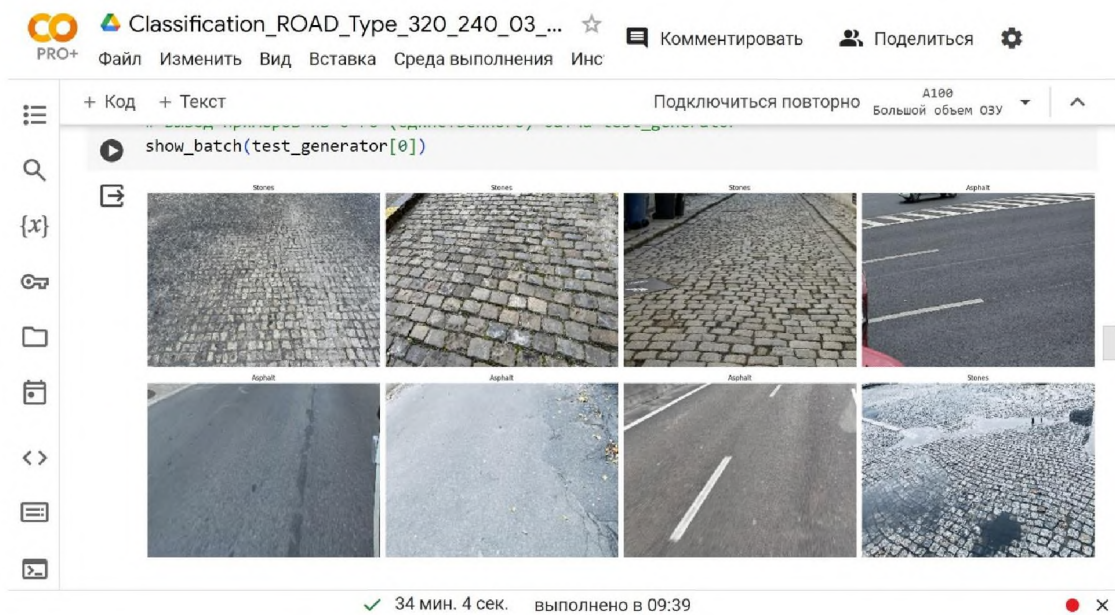


Рисунок 3.10 – Приклад зображень перевірконої вибірки

Аналогічно, для класу Stones було отримано 83 % правильних відповідей, у своїй 17 % зображень асфальту сприйняли нейронною мережею як бруківка.

Характерно, що збільшення датасету та тестової вибірки дозволило уникнути викидів точності навчання до 100-відсоткового показника та зробило перебіг кривих точності навчання та відповідних помилок більш характерним для об'ємних датасетів (рис. 3.12 і 3.13).

Збільшення розміру датасета дозволило провести експерименти щодо дослідження впливу розміру батча на точність навчання.

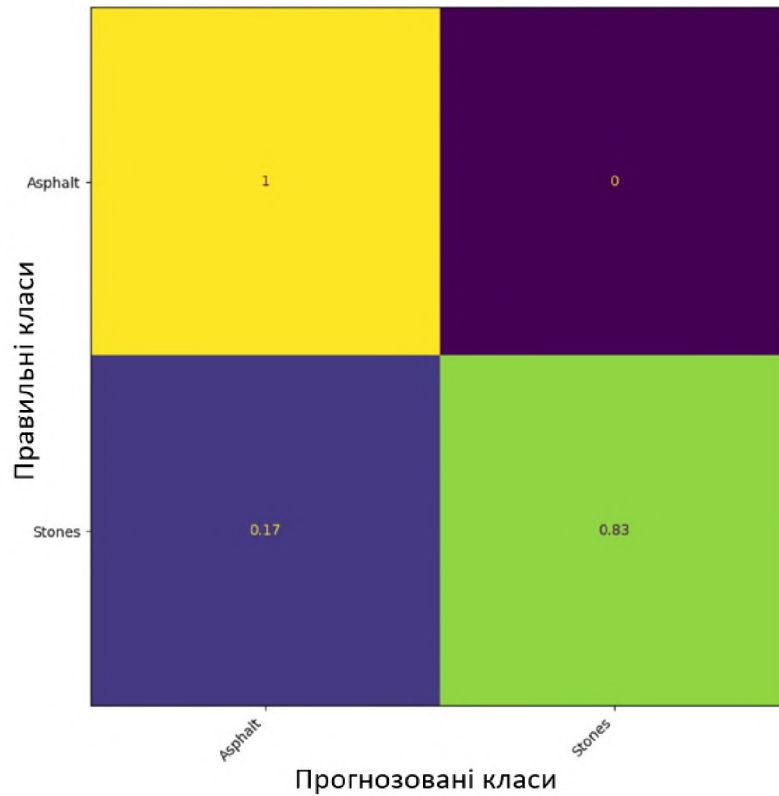


Рисунок 3.11 – Матриця помилок тестування на модифікованому датасеті

Зокрема, було досліджено навчання нейромережі на батчі 16 зображень. При цьому середовищем Google Colab Pro+ було надано серверний прискорювач Python 3 на базі Google Compute Engine (GPU) V100.

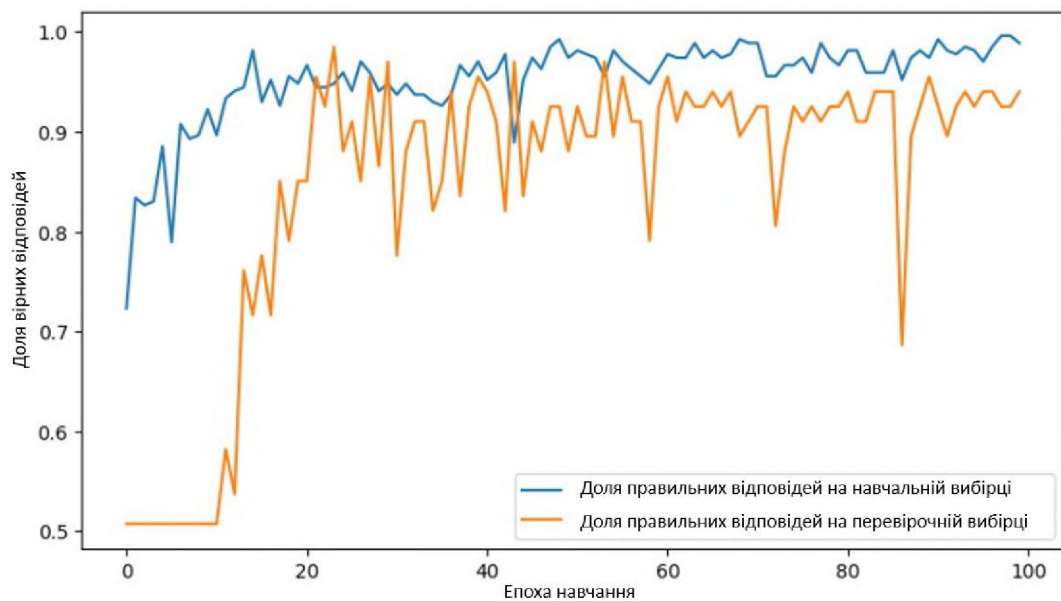


Рисунок 3.12 – Графік точності навчання на модифікованому датасеті

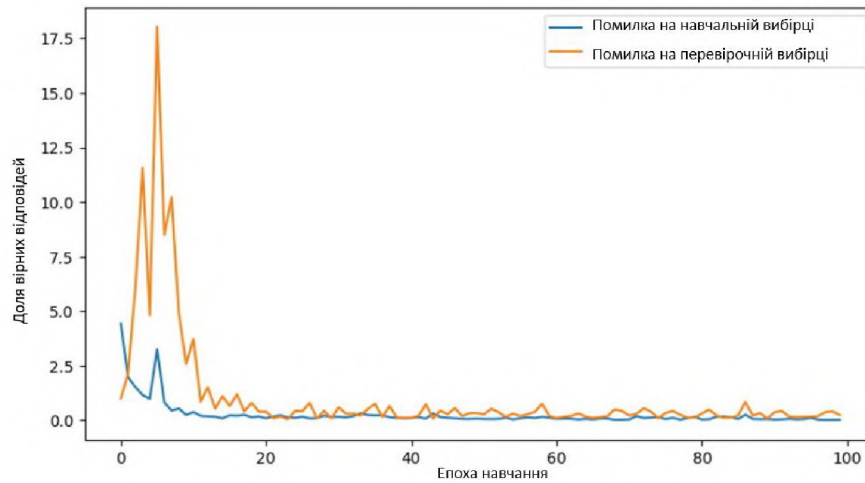


Рисунок 3.13 – Графік помилок навчання на модифікованому датасеті

Процес навчання характеризувався споживанням ресурсів оперативної пам'яті системи 4.9 GB з наданих 51.0 GB, оперативної пам'яті графічного процесора 14.8 з 16.0 GB у розпорядженні, а також 27.7 з 166.8 GB дискового простору. Наслідком збільшення розміру батча стало затягування часу початку інтенсивного навчання нейронної мережі до 29 епох. При цьому з урахуванням переходу на використання менш продуктивних обчислювальних засобів, час виконання однієї епохи навчання зріс до 21-22 секунд.

На рис. 3.14 наведено матрицю помилок по тестовій вибірці. Як видно, відсоток помилок ідентифікації асфальтового покриття як бруківка знизився до рівня 11 %.

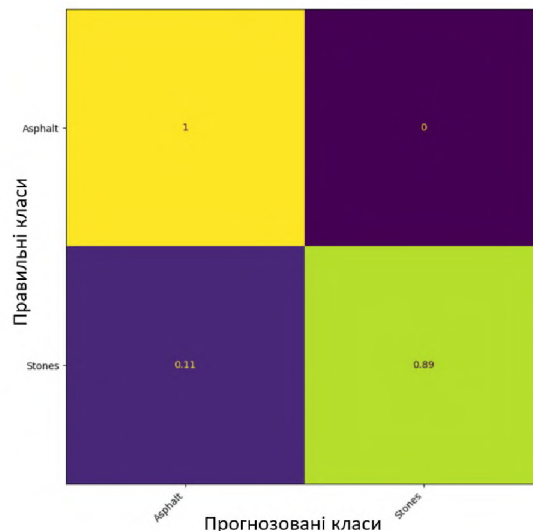


Рисунок 3.14 – Матриця помилок при батчі 16

Графіки, що ілюструють параметри процесу навчання, наведено на рис. 3.15 і 3.16.

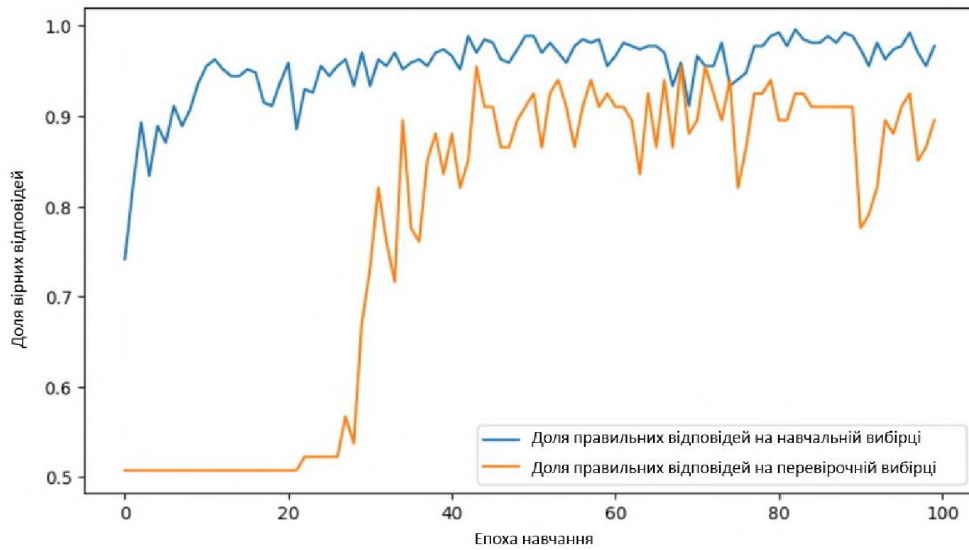


Рисунок 3.15 – Зміни показника точності у процесі навчання на 100 епохах при батчі 16

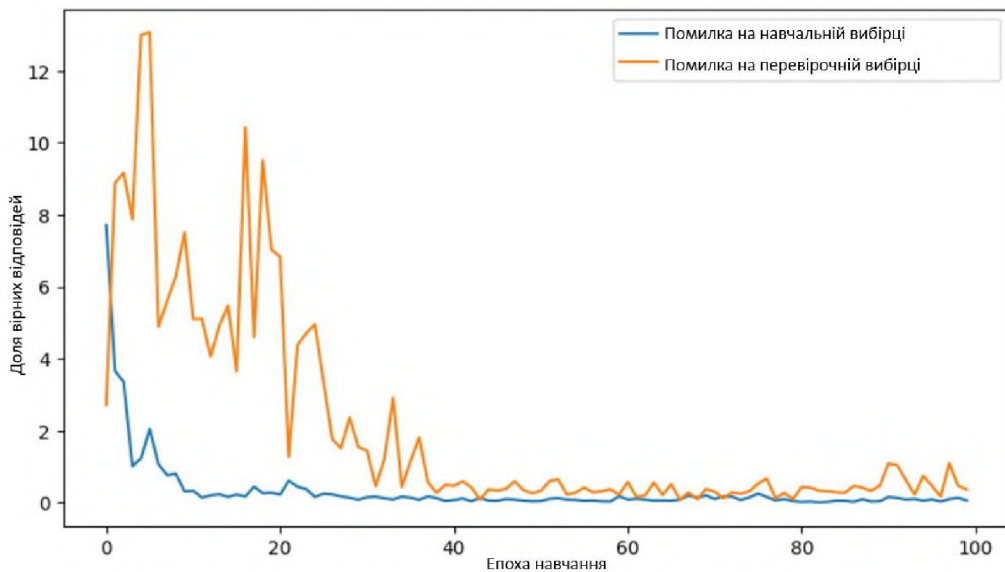


Рисунок 3.16 – Графіки помилок навчання на 100 епохах при батчі 16

Середня точність розпізнавання класифікатора на основі CNN при батчі 16 зростає до 94 %. При цьому, як і раніше, клас Asphalt розпізнавався правильно з точністю 100 %, а точність правильного зарахування зображення до класу Stones з'явилася до позначки 89 %.

### 3.2 Використання ChatGPT для класифікації дорожнього покриття

В якості альтернативи до використання CNN можна застосовувати ChatGPT для класифікації дорожнього покриття. Користувач завантажує зображення дорожнього покриття до системи. Потім система використовує можливості аналізу зображень для визначення типу покриття. На основі аналізу зображення, ChatGPT ідентифікує тип покриття (асфальт, бруківка тощо). Система повертає користувачу відповідь з класифікацією. Уточнення результатів можливе через додаткові питання або надання додаткових зображень. Приклад такого підходу наведений на рис. 3.17.

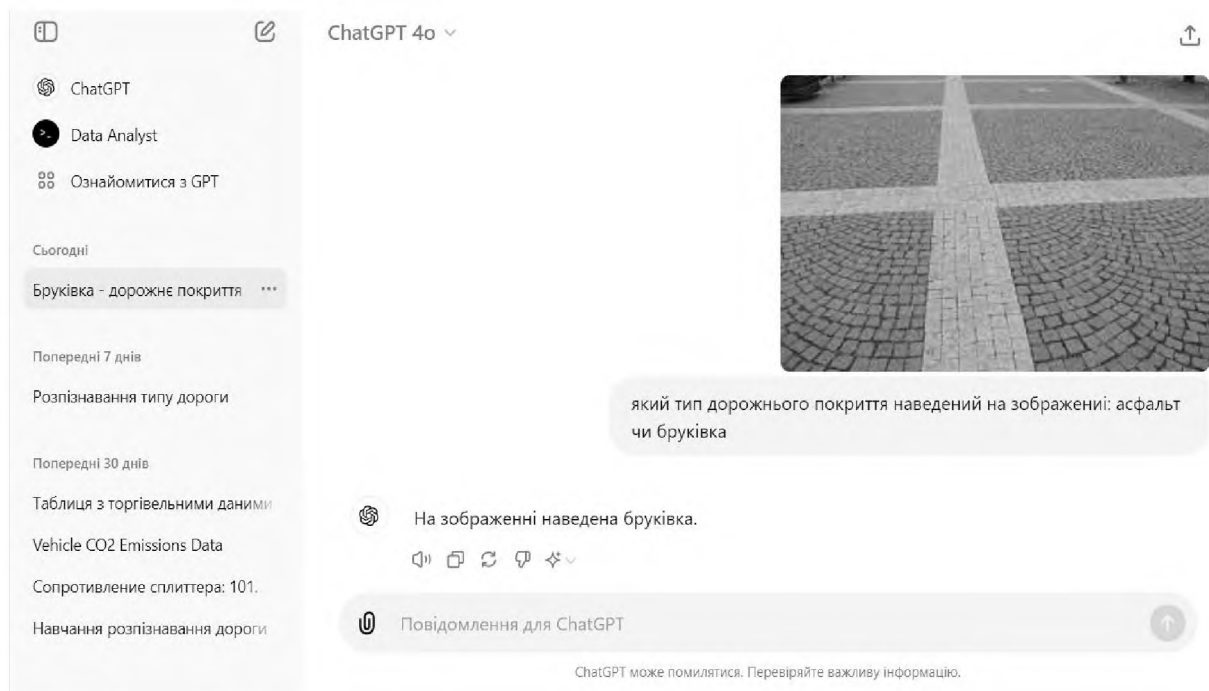


Рисунок 3.17 – Використання ChatGPT для класифікації дорожнього покриття (асфальт/бруківка)

Таким чином, ChatGPT може бути ефективним інструментом для класифікації дорожнього покриття на основі зображень, що надаються користувачем. В ході проведених досліджень встановлено, що перевагами використання ChatGPT для класифікації дорожнього покриття порівняно з використанням звичайної CNN слід вважати наступні положення:

– ChatGPT може обробляти як текстову, так і візуальну інформацію, що дозволяє надавати більш детальні пояснення та зворотний зв'язок, ніж звичайна CNN;

– крім класифікації зображень, ChatGPT може виконувати інші завдання, такі як відповідь на запитання, надання контекстної інформації та підтримка в інтерактивних діалогах;

– ChatGPT може вести діалог з користувачем, уточнювати деталі, отримувати додаткову інформацію та відповідно коригувати свої відповіді, що підвищує точність і корисність результатів;

– використання ChatGPT може бути менш ресурсомістким, ніж розробка та впровадження спеціалізованої CNN моделі, оскільки ChatGPT вже має вбудовані можливості обробки зображень та текстів;

– ChatGPT як сервіс може бути швидше інтегрований у різні додатки без необхідності тривалого навчання та налаштування моделі з нуля, як у випадку зі CNN;

– ChatGPT може легко адаптуватися до нових типів даних і завдань без необхідності повної переучення, що може бути більш гнучким у динамічних умовах;

Однак, CNN має свої переваги:

– Спеціалізовані CNN моделі, навчальні на великих наборах даних, зазвичай мають високу точність у задачах класифікації зображень;

– CNN можна спеціально налаштовувати та оптимізувати для конкретних типів завдань і даних, що може забезпечити кращу продуктивність в окремих випадках;

– CNN широко використовуються в задачах комп'ютерного зору і мають багато добре досліджених архітектур і методів покращення.

Таким чином, вибір між ChatGPT і CNN залежить від конкретних потреб і обмежень проекту.

### 3.3 Економічне обґрунтування прийнятих рішень

Розробка коду моделі розпізнавання типу дорожнього покриття на основі нейронної мережі реалізується на мові Python. Вважатимемо, що за одну години програміст здатен розробити 20 рядків коду. Для практичної реалізації розглянутої моделі глибокого навчання нейронної мережі (1500 рядків) трудомісткість розробки коду складає 75 год.

Крім цього, необхідно до 15 годин на тестування програмного продукту до 8 год для процедури DevOps. При цьому, вважаємо, що набір даних для формування датасету зібраний. Таким чином, загальна трудомісткість складає:

$$T = 75 + 15 + 8 = 98. \quad (3.1)$$

Відповідно, перерахунок трудомісткості у кількість робочих днів можна виконати з виразом:

$$M = \frac{T}{8} = 12,25 \text{ днів або } \approx 0,4 \text{ місяця}. \quad (3.2)$$

Враховуючи обсяг зазначених робіт, доцільно найняти програміста-фрілансера на мові Python з відповідною класифікацією. Згідно [32], в середньому по Україні, його заробітна плата складає  $C = 50000$  грн. Як наслідок, витрати на заробітну плату програміста-фрілансера дорівнюють:

$$V = M \cdot C = 0,4 \cdot 50000 = 20000 \text{ [грн]}. \quad (3.3)$$

Згідно, п. 3.1, для реалізації процесу навчання нейронної мережі використовується хмарний сервіс Google Colab. Щоб оптимізувати пошук оптимальних архітектури та гіперпараметрів синтезованої нейронної мережі доцільно отримати тарифний план Google Colab Pro+ [33]. Він забезпечує підключення серверних відеокарт (наприклад, NVIDIA A100) та/або TPU. За 2 місяця витрати на цей тарифний план складають 4560 грн. Таким чином, орієнтовні витрати на реалізацію запропонованої моделі складають  $\approx 24600$  грн. З іншого боку, використання API від ChatGPT передбачає абонплату  $\approx 800$  грн/місяць. При цьому, перший варіант дозволяє роботу без

доступу глобальної мережі Інтернет. Щоб зменшити витрати можна використовувати локальну версію LLM. Наприклад, для цього можна скористатись платформою LM Studio (рис. 3.18) [34].

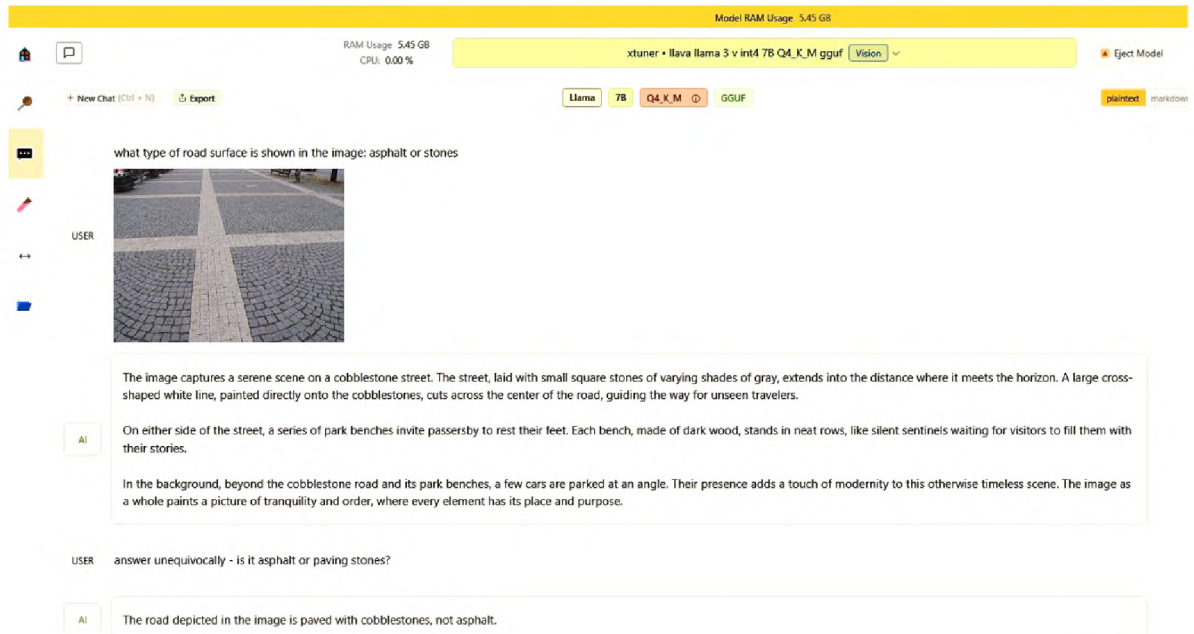


Рисунок 3.18 – Використання ПЗ LM Studio 0.2.24 з квантованою (int4) VLM Llava-Llama-3-8b-v1\_1 [35, 36] для розпізнавання типу дорожнього покриття

В даному випадку, перша відповідь виглядає дуже обширно: «На зображенні зображено спокійну сцену на брукованій вулиці. Вулиця, викладена невеликими квадратними каміннями різних відтінків сірого, простягається в далечінь, де зустрічається з горизонтом. Велика хрестоподібна біла лінія, намальована прямо на бруківці, прорізає по центру дороги, вказуючи шлях невидимим мандрівникам. Поруч вулиці ряд паркових лавок запрошує перехожих відпочити. Кожна лавка, зроблена з темного дерева, стоїть акуратними рядами, як мовчазні вартові, чекаючи, коли відвідувачі заповнять їх своїми історіями. На задньому плані, за брукованою дорогою та її парковими лавками, під кутом припарковано кілька автомобілів. Їхня присутність додає відтінок сучасності цій позачасовій сцені. Зображення в цілому малює картину спокою та порядку, де кожен елемент має своє місце та призначення». Як наслідок, для VLM потрібно додавати додаткові prompts

[37-40] для отримання більш лаконічної відповіді: «відповідай однозначно – це асфальт або бруківка». Після цього можна отримати найбільш оптимальний варіант відповіді: «Дорога, зображена на зображенні, вимощена бруківкою, а не асфальтом».

Таким чином, застосування штучного інтелекту на основі нейронних мереж для вирішення завдань розпізнавання типу дорожнього покриття дозволить значно підвищити функціональність систем автопілотування, дронів та ін.

## ВИСНОВКИ

Розпізнавання типу дорожнього покриття дозволяє транспортним засобам оптимізувати витрати палива та електроенергії, покращувати навігацію та планування маршрутів, а також підвищувати загальну безпеку.

Для розпізнавання типу дорожнього покриття використовуються візуальні методи, сенсори та машинне навчання. Особлива увага приділяється штучному інтелекту на основі нейронної мережі. Їх використання дозволяє системам постійно вдосконалювати свої здатності до розпізнавання різних типів покриття.

Нейронні мережі, такі як CNN ефективні у задачах класифікації зображень і можуть працювати в реальному часі. Результати їх роботи інтегруються з іншими системами, такими як навігаційні та системи прогнозування трафіку, що дозволяє покращити управління дорожнім рухом та знизити витрати на обслуговування доріг.

Розвиток технологій розпізнавання дорожнього покриття також включає використання LLM та VLM. Це дозволяє виконувати складні мультимодальні завдання, такі як опис зображень та пошук зображень за текстовими запитамі, що робить ці моделі корисними для широкого спектру застосувань у сфері автономних транспортних систем.

Синтез моделі розпізнавання типу дорожнього покриття за допомогою нейронної мережі базується на розробці програмного коду на мові Python. Процес включає кілька етапів, починаючи зі збору даних за допомогою датчиків та камер, обробки цих даних, розробки моделі, навчання, тестування та валідації моделі, а також її розгортання. Навчання моделі проводиться за допомогою великих наборів даних, що розділяються на тренувальний, валідаційний та тестовий набори. Після навчання модель тестується для оцінки її точності та надійності. На основі проведених досліджень сформовано рекомендації щодо реалізації розпізнавання типу дорожнього покриття за допомогою нейронної мережі та оцінку точності цього процесу.

За пропонована модель нейронної мережі розпізнавання типу дорожнього покриття дозволяє досягти середньої точності класифікації 92 %.

Також досліджено альтернативний підхід до класифікації типу дорожнього покриття з використанням ChatGPT. Цей підхід дозволяє обробляти як текстову, так і візуальну інформацію, надаючи детальні пояснення та зворотний зв'язок. Використання ChatGPT може бути менш ресурсомістким та швидше інтегрованим у різні додатки. Однак CNN забезпечує високу точність у задачах класифікації зображень і може бути спеціально налаштованою для конкретних типів завдань.

Економічне обґрунтування показало, що розробка моделі на основі нейронних мереж потребує певних витрат на заробітну плату програміста та використання хмарних сервісів. Загальні витрати на розробку та реалізацію моделі оцінюються у 24600 грн. Застосування штучного інтелекту для розпізнавання типу дорожнього покриття дозволяє підвищити функціональність систем автопілотування дронів та інших автономних транспортних засобів.

Таким чином, результатами роботи є створена модель розпізнавання типу дорожнього покриття на основі нейронної мережі; порівняльній оцінка точності нейронних мереж розпізнавання типу дорожнього покриття; рекомендації щодо використання моделі глибокого навчання нейронної мережі розпізнавання типу дорожнього покриття. Вони можуть бути використані для подальших досліджень за даною тематикою та при проектуванні реалізації систем автопілотування безпілотних транспортних засобів.