

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук  
(повна назва)

Кафедра Штучного інтелекту  
(повна назва)

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

рівень вищої освіти другий (магістерський)

Розробка підсистеми розпізнавання дорожніх знаків автомобільного автопілота  
(тема)

Виконав:  
студент 2 курсу, групи СШМ-19-2  
Коновалов О.Ю.  
(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки  
(код і повна назва спеціальності)

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

Освітня програма Системи штучного інтелекту  
(повна назва спеціалізації)

Керівник доц. Вітько О.В.  
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри \_\_\_\_\_  
(підпис)

В.О. Філатов  
(прізвище, ініціали)

2021 р.

Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ Комп'ютерних наук  
(повна назва)  
Кафедра \_\_\_\_\_ Штучного інтелекту  
(повна назва)  
Рівень вищої освіти \_\_\_\_\_ другий (магістерський)  
Спеціальність \_\_\_\_\_ 122 Комп'ютерні науки  
(код і повна назва)  
Тип програми \_\_\_\_\_ освітньо-наукова  
(освітньо-професійна або освітньо-наукова)  
Освітня програма \_\_\_\_\_ Системи штучного інтелекту (СШІ)  
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

«\_\_\_\_\_» \_\_\_\_\_ 20 \_\_\_\_ р.

**ЗАВДАННЯ**  
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові \_\_\_\_\_ Коновалову Олександрю Юрійовичу  
(прізвище, ім'я, по батькові)

1. Тема роботи Розробка підсистеми розпізнавання дорожніх знаків автомобільного автопілота

затверджена наказом університету від 29 \_\_\_\_\_ 03 \_\_\_\_\_ 20 21 р. № 390Ст

2. Термін подання студентом роботи до екзаменаційної комісії \_\_\_\_\_ 20 \_\_\_\_ р.

3. Вихідні дані до роботи Науково-технічні публікації, електронні документації, патенти, дані Інтернет.

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

4. Перелік питань, що потрібно опрацювати в роботі \_\_\_\_\_

1) Аналіз предметної області

2) Розробка алгоритмів виявлення і розпізнавання дорожніх знаків на зображенні

3) Розробка програмного забезпечення для розпізнавання дорожніх знаків на зображенні

4) Оцінка якості та тестування роботи програмної системи

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Рисунок 1 – типи ознак для алгоритму Віоли-Джонса, Рисунок 2 – архітектура згорткової нейронної мережі, Рисунок 3 – приклад скелетизації символу, Рисунок 4 – приклад заборонних дорожніх знаків, Рисунок 5 – приклад попереджувальних дорожніх знаків, Рисунок 6 – приклад розпорядчих дорожніх знаків, Рисунок 7 – приклад знаків особливих розпоряджень, Рисунок 8 – приклад роботи Opel Eye, Рисунок 9 – приклад роботи Speed limit assist, Рисунок 10 – приклад роботи Road sign information, Рисунок 11 – функціональна схема комбінованого методу виявлення дорожніх знаків, Рисунок 12 – структура згорткової нейронної мережі, Рисунок 13 – порівняння точності розпізнавання.

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1 )

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

#### КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	29.03.2021	Виконано
2	Аналіз завдання та пошук літератури з теми	30.03.2021	Виконано
3	Аналіз об'єкту дослідження	01.04.2021	Виконано
4	Розробка алгоритму розпізнавання	08.04.2021	Виконано
5	Вибір програмних засобів для розробки додатку	09.04.2021	Виконано
6	Розробка програмного додатку	12.04.2021	Виконано
7	Аналіз отриманих результатів	20.04.2021	Виконано
8	Оформлення пояснювальної записки	22.04.2021	Виконано
9	Оформлення презентаційних матеріалів	23.04.2021	Виконано
10	Представлення на рецензування	10.05.2021	
11	Попередній захист	11.05.2021	
12	Захист перед ЕК		

Дата видачі завдання 29 03 2021 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_  
(підпис) \_\_\_\_\_  
(посада, прізвище, ініціали)

## РЕФЕРАТ

Пояснювальна записка: 82 с., 3 табл., 28 рис., 2 дод., 27 джерел.

АВТОМОБІЛЬНИЙ АВТОПЛОТ, АЛГОРИТМИ, ДОРОЖНІ ЗНАКИ, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, КОМП'ЮТЕРНИЙ ЗІР, РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ, C#, OPENCV, SVM.

Об'єкт дослідження – система обробки зображень, заснована на сверточное нейронної мережі і методі опорних векторів.

Предмет дослідження – застосування нейромережевих алгоритмів для виявлення і розпізнавання дорожніх знаків на зображеннях.

Мета роботи – розробка алгоритму виявлення і розпізнавання дорожніх знаків, що володіє високим ступенем спотворень.

Отримані результати і новизна – комбінований алгоритм, що складається з методу опорних векторів та згорткової нейронної мережі. На його основі було створено програмне рішення для розпізнавання дорожніх знаків, яке показало високий відсоток ідентифікації, а також стійкість до шумів і різних спотворень.

Сфера застосування – результати роботи можуть бути використані для розпізнавання дорожніх знаків і інших об'єктів в системах відеоспостереження, відеореєстраторах та інших сферах, пов'язаних із забезпеченням безпеки руху по дорогах.

Актуальність завдання – розпізнавання дорожніх знаків обумовлена підвищенням рівня безпеки на дорогах і надзвичайної важливості інформації, яку містять дорожні знаки. При використанні автоматизованої системи розпізнавання вкрай важливо точно і своєчасно ідентифікувати дорожні знаки при русі транспортного засобу, як в умовах міста, так і автостради.

## РЕФЕРАТ

Пояснительная записка: 82 с., 3 табл., 28 рис., 2 доп., 27 источников.

АВТОМОБИЛЬНЫЙ АВТОПИЛОТ, АЛГОРИТМЫ, ДОРОЖНЫЕ ЗНАКИ, КОМПЬЮТЕРНОЕ ЗРЕНИЕ, РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ, СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ, C#, OPENCV, SVM.

Объект исследования – система обработки изображений, основанная на сверточной нейронной сети и методе опорных векторов.

Предмет исследования – применение нейросетевых алгоритмов для обнаружения и распознавания дорожных знаков на изображениях.

Цель работы – разработка алгоритма обнаружения и распознавания дорожных знаков, обладающего высокой степенью искажений.

Полученные результаты и новизна – комбинированный алгоритм, состоящий из метода опорных векторов и сверточной нейронной сети. На его основе было создано программное решение для распознавания дорожных знаков, которое показало высокий процент идентификации, а также устойчивость к шумам и различным искажениям.

Сфера применения – результаты работы могут быть использованы для распознавания дорожных знаков и других объектов в системах видеонаблюдения, видеорегистраторах и других сферах, связанных с обеспечением безопасности движения по дорогам.

Актуальность задачи – распознавание дорожных знаков обусловлена повышением уровня безопасности на дорогах и чрезвычайной важности информации, которую содержат дорожные знаки. При использовании автоматизированной системы распознавания крайне важно точно и своевременно идентифицировать дорожные знаки при движении транспортного средства, как в условиях города, так и автострады.

## **ABSTRACT**

Thesis: 82 p., 3 tabl., 28 fig., 2 ann., 27 references.

ALGORITHMS, AUTOMOBILE AUTOPILOT, COMPUTER VISION, CONVOLVED NEURAL NETWORKS, C#, IMAGE RECOGNITION, OPENCV, ROAD TRAFFIC SIGNS, SVM.

The object of research – an image processing system based on a convolutional neural network and the method of reference vectors.

The subject of research is the application of neural network algorithms for detection and recognition of road signs on images with a complex background.

The aim of the work is to develop an algorithm for detecting and recognizing road signs with a high degree of invariance to affine and projection distortions.

The obtained results and novelty are a combined algorithm consisting of two methods, the first is the method of reference vectors, the second is a convolutional neural network. Based on this method, a software solution for road sign recognition was created, which showed a high percentage of identification, as well as resistance to noise and various distortions such as changes in the angle of inclination.

Scope – the results of the work can be used to recognize road signs and other objects in images in video surveillance and video surveillance systems, video recorders and other areas related to traffic safety on public roads and in general computer vision .

The relevance of the task – the recognition of road signs due to the increased level of safety on public roads and the extreme importance of the information contained in road signs. When using an automated recognition system, it is extremely important to accurately and timely identify road signs when a vehicle is moving, both in a city and on a highway.

## ЗМІСТ

Вступ.....	9
1 Аналіз предметної області.....	10
1.1 Огляд існуючих алгоритмів, що застосовуються для ідентифікації дорожніх знаків.....	10
1.1.1 Порівняння із шаблоном.....	12
1.1.2 Метод перетворення Хафа.....	12
1.1.3 Дескриптори локальних особливостей.....	14
1.1.4 Методи, які використовують гістограмний аналіз зображення.....	16
1.1.5 Метод головних компонент.....	17
1.1.6 Лінійний дискримінантний аналіз.....	19
1.1.7 Алгоритм AdaBoost.....	19
1.1.8 Алгоритм Віоли-Джонса.....	21
1.1.9 Згорткові нейронні мережі (ЗНМ).....	22
1.1.10 Структурні методи розпізнавання.....	25
1.1.11 Морфологічний аналіз.....	27
1.2 Дорожні знаки.....	27
1.3 Аналіз існуючих систем, що застосовуються для розпізнавання дорожніх знаків.....	31
1.3.1 Загальна архітектура.....	31
1.3.2 Opel Eye.....	31
1.3.3 Speed limit assist (Mercedes).....	33
1.3.4 Road sign information.....	34
1.3.5 Система розпізнавання дорожніх знаків лабораторії комп'ютерної графіки та мультимедіа МДУ ім. Ломоносова.....	35
1.3.6 Порівняння технічних характеристик існуючих систем розпізнавання дорожніх знаків.....	37

1.4	Постановка задачі.....	38
2	Розробка алгоритмів виявлення і розпізнавання дорожніх знаків на зображенні.....	40
2.1	Розробка методу розпізнавання об'єктів на зображенні .....	40
2.2	Метод опорних векторів.....	41
2.3	Опис алгоритму.....	45
2.4	Класифікація за допомогою згорткової нейронної мережі.....	47
3	Розробка програмного забезпечення для розпізнавання дорожніх знаків на зображенні.....	53
3.1	Загальні вимоги до програмного забезпечення, що розробляється...53	
3.2	Вибір засобів розробки програмного забезпечення.....	54
3.3	Інтерфейс програмного забезпечення.....	55
3.4	Оцінка якості роботи програмної системи.....	59
3.4.1	Оцінка якості розпізнавання дорожніх знаків на помилкових зображеннях.....	60
3.4.2	Оцінка якості розпізнавання дорожніх знаків при різних кутах відхилення.....	62
3.4.3	Порівняння технічних характеристик розроблених алгоритмів і програмних засобів з існуючими.....	64
	Висновки.....	67
	Перелік джерел посилання.....	69
	Додаток А Вихідний код програми.....	72
	Додаток Б Відомість кваліфікаційної роботи .....	82

## ВСТУП

За останні кілька десятиліть транспортні засоби настільки сильно інтегрувалися в наше повсякденне життя, що зараз важко уявити життєдіяльність людей без них.

В сучасному автомобілебудуванні все частіше зустрічаються технічні системи, спрямовані на оптимізацію водіння автомобіля, що включають в себе також системи автоматичного розпізнавання дорожніх знаків. Це полегшує завдання водія, дозволяє йому краще зосередитися на процесі водіння, підвищує безпеку всіх учасників дорожнього руху. На відміну від інших способів ідентифікації знаків (як, наприклад, радіометричні «маяки», здатні повідомити радіо-модулю автомобіля про те, що він увійшов в зону дії даного знаку) система безпосереднього розпізнавання знаків з зображень має переваги в плані надійності можливості її застосування до існуючої інфраструктури дорожніх знаків. Розпізнавання дорожніх знаків цікаве підвищенням рівня безпеки на дорогах загального користування та надзвичайної важливості інформації, яку містять дорожні знаки. При використанні автоматизованої системи розпізнавання вкрай важливо точно і своєчасно ідентифікувати дорожні знаки при русі транспортного засобу як в умовах міста, так і автостради.

На сьогоднішній день для виконання завдання розпізнавання розробляються і використовуються комерційні системи, які поставляються «в комплекті» з автомобілем. Такі системи є закритими, зокрема, компанія Opel з системою «Opel Eye», компанія Mercedes з системою «Speed limit assist», компанія Volvo з системою «Road sign information».

## 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

### 1.1 Огляд існуючих алгоритмів, що застосовуються для ідентифікації дорожніх знаків

Об'єкти на зображеннях є деякою впорядкованою сукупністю ознак. Сукупність об'єктів, що володіють деякими загальними характерними ознаками, називається класом. Машинне розпізнавання об'єктів на зображеннях полягає в здатності автоматично класифікувати вихідні дані і відносити їх до певного класу за допомогою виділення характерних ознак об'єкта. Проблема автоматичного розпізнавання об'єктів на зображеннях є відносно новою і отримала свій розвиток у другій половині 20 століття. У перших системах розпізнавання використовувалися прості геометричні моделі, і була потрібна участь адміністратора, який виробляв виділення ознак об'єкта на зображенні. Потім система виконувала чисельні вимірювання розмірів і відстаней виділених ознак щодо контрольних точок.

В даний час розпізнавання об'єктів на зображеннях здійснюється на основі складних математичних уявлень про існуючі процеси і є дуже затребуваною в областях комп'ютерного зору, обробки і аналізу зображень, біометрії, систем безпеки та відео-контролю. Сучасні методи, які вирішують задачу розпізнавання об'єктів, застосовуються для вирішення широкого кола завдань: розпізнавання осіб, відбитків пальців, сітківки ока, друкованих символів, автомобільних номерних знаків, маркування на поверхнях різних об'єктів і т.д.

На сьогоднішній день досягнуто значних успіхів при вирішенні задач розпізнавання об'єктів і символів на зображеннях, проте існує ряд складнощів, які істотно знижують надійність застосування сучасних методів:

– низький дозвіл зображень. В результаті чого, ознаки об'єктів

можуть бути погано помітні;

- наявність складної фонові структури на зображеннях. Це має на увазі наявність на зображеннях сторонніх об'єктів, які можуть мати візуальні ознаки, схожі з шуканим об'єктом;

- різні спотворення, отримані в процесі реєстрації зображень. Реєстрація зображень може проводитися при невдалих ракурсах, поганих погодних умовах, різних кутах і умов освітлення. В результаті цього, на зображеннях можуть бути різні шумові перешкоди, символи можуть бути схильні до афінним і проекційним спотворень;

- кількість шуканих об'єктів на зображенні заздалегідь не відомо.

Все це вимагає застосування різних алгоритмів попередньої обробки, що в свою чергу ускладнює процес розпізнавання, робить його більш громіздким, збільшує обсяг і час обчислювальних процесів.

Тому на сьогоднішній день до сих пір існує потреба в розробці методів і алгоритмів, які вирішують перераховані вище проблеми.

Для усунення перерахованих вище складнощів потрібно застосування різних додаткових підходів попередньої обробки, що збільшує обчислювальну складність системи. Таким чином, можна зробити висновок, що сьогодні існує необхідність в реалізації методів і алгоритмів, що дозволяють знизити вплив перерахованих вище складнощів на процес розпізнавання.

Розпізнавання дорожніх знаків на зображеннях можна логічно розділити на два основних етапи:

- виявлення області розташування дорожнього знака на зображенні;

- розпізнавання дорожнього знака.

Для реалізації кожного з цих етапів застосовуються свої індивідуальні методи і алгоритми. Тому далі в цій главі наводиться огляд методів і алгоритмів за трьома категоріями застосування.

### 1.1.1 Порівняння із шаблоном

Головним принципом роботи цих методів є порівняння кожної області зображення з заданим шаблоном, в наслідок чого визначається взаємна кореляція. Шаблони можуть задаватися вручну або визначатися будь-якої функцією, а також можуть відповідати цілій об'єкту або окремими складовими об'єкта. При перевищує заданий поріг коефіцієнті кореляції, область дослідження зображення буде позначатися як містить образ дорожнього знака. Так само виділяються різні поліпшення даного підходу, що використовують кілька шаблонів, які відповідають окремим компонентам об'єктів, що представляли собою області у вигляді ліній. Робота даних алгоритмів складалася з декількох етапів, на кожному з яких визначалася достовірність знайдених областей. Якщо область зображення відповідає одному з шаблонів, то вона відзначається як область, яка представляє інтерес для більш детального аналізу.

Таким чином, підвищується міра подібності областей зображення з заданим шаблоном. Далі виявлені області порівнювалися з іншими шаблонами, що дозволяло визначити наявність шуканого об'єкта. Дані методи використовувалися в основному в ранніх роботах по класифікації та розпізнавання об'єктів на зображеннях. Це були перші спроби сформулювати ознаки зображення об'єктів [1].

Сучасні модифікації шаблонних методів, показують високу точність розпізнавання. Недолік даних методів полягає в низькій швидкості роботи і високої чутливості спотворень об'єктів на зображеннях.

### 1.1.2 Метод перетворення Хафа

Даний метод був запропонований Р. Hough в 1962 році. Метод полягає в пошуку на зображенні об'єктів, що належать певному класу

геометричних фігур: прямі лінії, кола та т.д. Пошук здійснюється за допомогою процедури голосування, що застосовується до простору параметрів зображення. Шуканий об'єкт представляється у вигляді параметричного рівняння, параметри якого представляють так зване фазовий простір. Основне твердження перетворення Хафа полягає в тому, що будь-яка точка зображення може належати деякому набору ліній. Пряма лінія може здаватися такою формулою (1.1):

$$r = X \cos \theta + Y \sin \theta, \quad (1.1)$$

де  $r$  – довжина перпендикуляра, побудованого з початку координат на пряму;

$\theta$  – кут між цим вектором і віссю абсцис.

Синусоїдальні криві для кожної точки є унікальними. Дані криві перетнуться в точці  $(r, \theta)$  тільки в тому випадку, якщо породжують їх точки знаходяться на прямій, описуваної рівнянням (1.1). Функція  $A(r, \theta)$  називається акумуляторної функцією і її значення в точці  $(r, \theta)$  буде дорівнює кількості точок, розташованих на відповідній прямій. Щоб знайти прямі лінії на зображенні досить знайти локальні максимуми акумуляторної функції.

Для посилення просторових частот на початковому етапі даного методу виконується поліпшення контрасту, бінаризація або виділення кордонів на зображенні. Далі виконується послідовний аналіз кожної точки зображення і її сусідів. При цьому обчислюється вага кордону в даній точці. Якщо вага відповідає заданому граничному значенню, то обчислюються параметри прямої і збільшується значення у відповідній клітинці масиву. Далі виконується пошук максимальних значень в масиві, за які проголосувала більша кількість пікселів зображення. Таким чином, знаходяться параметри рівнянь шуканого об'єкта.

Алгоритм полягає в пошуку прямих ліній, що задовольняють наступним умовам:

- прямі лінії повинні мати приблизно одну довжину і збігаються координати початку і кінця;
- лінії повинні знаходитися на певній відстані один від одного.

Таким чином, на зображенні знаходяться пари ліній, які потім зіставляються. Якщо лінії перетинаються і утворюють прямокутник, то обчислюється співвідношення висоти і ширини граней знайденого прямокутника. Якщо співвідношення задовольняють умовам, то область позначається на можливий зміст дорожнього знака. Гідність даного методу полягає в тому, що він є інваріантним до яскравості і кольору зображення. Недолік даного методу полягає в тому, що він сприйнятливий до будь-якої області зображення, що має прямокутну форму. Це знижує ефективність його застосування на зображеннях зі складною фоновією структурою [2].

### 1.1.3 Дескриптори локальних особливостей

Для виділення ознак об'єктів на зображеннях часто застосовуються дескриптори локальних особливостей. Дескриптором називається сукупність параметрів, за допомогою яких можна описати характеристики зображення, наприклад, такі як колір, текстуру і т.д. Початковим етапом виявлення об'єктів на зображенні за допомогою дескрипторів є визначення характерних точок. Характерними точками називаються точки, що володіють високою локальною інформативністю параметри, яких не змінюються при різних фотометричних і геометричних перетвореннях зображення. Для знайденої безлічі точок на зображенні розраховуються дескриптори [3].

Одним з найбільш популярних алгоритмів, які включають дескриптор і детектор характерних точок зображення, є SIFT (Scale

Invariant Feature Transform) [25]. Даний алгоритм є локальною гистограмою напрямків градієнтів зображення. Принцип роботи алгоритму SIFT полягає в обчисленні згортки вихідного зображення з ядром Гаусса при змінному параметрі згладжування. Після цього відбувається перетворення зображень до одного розміру, і обчислюється їх різницю. Далі виконується порівняння кожного пікселя на зображенні з вісьмома сусідніми пікселями при тих же параметрах і масштабі, з дев'ятьма сусідніми пікселями в більшому масштабі і з дев'ятьма в меншому масштабі. Пікселі, в яких локальні екстремуми перевершують заданий поріг, вибираються як характерні точки. Для кожної обраної точки обчислюється певний локальний дескриптор, який характеризує напрямок градієнтів в даній околиці пікселів. У 2005 році дослідники N. Dalal і B. Triggs представили алгоритм HOG (Histogram of Oriented Gradients) [26]. Принцип роботи даного алгоритму полягає в тому, що зображення представляється у вигляді щільної сітки рівномірно розподілених осередків. Для пікселів всередині кожної комірки обчислюються гистограми напрямків градієнтів. На основі отриманих параметрів виконується побудова дескриптора. Для збільшення точності в даному алгоритмі застосовується нормалізація перекривати локального контрасту гістограм. Нормалізовані дескриптори мають підвищену стійкість до зміни інтенсивності освітлення.

У 2006 році T. Tuytelaars, H. Bay, L. Van Gool представили алгоритм SURF (Speeded Up Robust Features) [27], який включає в себе дескриптор і детектор характерних точок зображення. В даному алгоритмі при обчисленні характерних точок використовуються не гистограми зважених градієнтів, а цілочисельні прямокутні фільтри різного масштабу. Це забезпечує стійкість до поворотів об'єкта і до зміни масштабу. Пікселі, в яких локальні екстремуми перевершують заданий поріг, вибираються як характерні точки і в них обчислюються локальні дескриптори. Навколо точки будується квадратна область і ділиться на декілька підгалузей. У

кожній підгалузі обчислюються відгуки на два типи вейвлетів – горизонтально і вертикально спрямовані. Отримані відгуки зважуються Гауссіаном і підсумовуються. Перераховані дескриптори широко використовуються при детектуванні різних об'єктів, в тому числі і автомобільних номерних знаків. Дані дескриптори забезпечують високу ступінь інваріантності до геометричних перетворень і зміни масштабу зображення.

Недоліком застосування дескрипторів є низька стійкість роботи при різних умовах освітлення, при відображають поверхнях, а також при різних кутах реєстрації об'єктів [4].

#### 1.1.4 Методи, які використовують гістограмний аналіз зображення

Дані методи ґрунтуються на припущенні, що частотна характеристика області розташування об'єкта, що містить символічні образи, відрізняється від інших областей зображення і має більш високу інтенсивність пікселів.

Для посилення просторових частот і придушення шумів на початковому етапі даного підходу виконується поліпшення контрасту, бінаризація або виділення кордонів на зображенні. В результаті цього, межі об'єктів на зображенні стають більш контрастними, а фон затемнюється. Зазвичай в задачах подібних розпізнаванню маркувань технічних об'єктів і автомобільних номерних знаків, шуканий об'єкт являє собою сукупність символів і ліній темного кольору на більш світлому тлі, або ж навпаки. Таким чином, дана область зображення буде мати високий контраст. Ідея даних методів полягає в наступному: виконується сканування зображення і при цьому обчислюється середнє значення яскравості пікселів в кожному рядку зображення (іноді обчислюють за стовпцями). У тому місці, де розташований шуканий об'єкт середня

інтенсивність пікселів буде значно відрізнятися від інших областей зображення. Максимальне значення отриманої проекції може збігтися з розташуванням об'єкта.

Дані методи застосовувалися багатьма дослідниками з різними модифікаціями. У роботах даний метод застосовувався з різними видами фільтрації зображення для придушення шумів і більш яскравого виділення області розташування автомобільного номерного знака. В роботі для детектування автомобільного номерного знака даний метод застосовувався спільно з перетворенням Хафа. Дані алгоритми можуть показувати хороші результати роботи в тому випадку, якщо розмір зображення автомобіля або іншого технічного об'єкта можна порівняти з розмірами кадру.

Гідність даних методів полягає в простій реалізації та високій швидкості роботи. Недолік даних методів полягає в чутливості до будь-якої області зображення, що має параметри інтенсивності пікселів схожі з дорожнім знаком.

### 1.1.5 Метод головних компонент

Метод головних компонент (Principal Component Analysis, PCA) був запропонований К. Pearson в 1901 році. Даний метод застосовується в різних областях, таких як: обробка і аналіз зображень, розпізнавання образів, стиснення даних і т.д. Метод PCA спрямований на зменшення розмірності даних, при найменших втратах інформації, що буває дуже актуально в області аналізу і обробки зображень. В даному методі виконується лінійне ортогональне перетворення вхідного вектора  $X$ , що містить корельовані компоненти, в вектор меншого розміру  $Y$ , який містить некорельовані змінні. Некорельовані змінні називаються головними компонентами. Вектор  $X$  складається з усіх прикладів навчального набору зображень. Власні вектори виходять при вирішенні

наступного рівняння (1.2):

$$D = \Phi^T K \Phi, \quad (1.2)$$

де  $D$  – діагональна матриця власних чисел;

$K$  – матриця коваріації для вектору  $X$ ;

$\Phi$  – матриця власних векторів.

З матриці власних векторів вибираються  $M$  найбільших власних чисел, з яких формується підматриця. Таким чином, отримуємо вираз:  $y = \Phi_M^T \tilde{x}$ , де  $\tilde{x} = x - \bar{x}$  – нормалізований вектор з нульовим математичним очікуванням. При виборі перших  $M$  компонент векторний простір розбивається на власний простір, що містить головні компоненти. Спочатку обчислюються головні компоненти для навчального набору зображень об'єктів. Виконується фільтрація параметрів об'єктів, де відкидаються менш істотні з них і простір параметрів зменшується. Таким чином, об'єкти на зображенні можуть бути представлені у вигляді мінімального набору параметрів. Інші компоненти є другорядними і характеризують незначні відмінності між об'єктами і шумом [5].

При класифікації об'єктів на зображенні обчислюється Евклідова відстань між головними компонентами досліджуваного зображення і зваженої комбінацією головних компонент навчального набору. Якщо головні компоненти зображення мають найбільш близьке відповідність з головними компонентами якого-небудь об'єкта з навчального набору, то вважається, що на зображенні міститься даний об'єкт.

Для ефективного застосування даного методу зображення об'єктів повинні бути отримані при однакових умовах реєстрації. Гідність методу РСА полягає в забезпеченні зменшення розмірності даних, що є важливим чинником в роботі з великими базами зображень.

### 1.1.6 Лінійний дискримінантний аналіз

Лінійний дискримінантний аналіз (Linear Discriminant Analysis, LDA) спрямований на вирішення завдання класифікації об'єктів по декількох класах. На відміну від PCA в даному методі не ставиться мета перебування простору ознак найменшої розмірності, ефективно описує навчальний набір даних. Основним завданням даного методу є знаходження проєкції ознак в простір, в якому різні класи об'єктів мають максимальне відмінність один від одного. Для кожного об'єкта зображення, що належить певному класу, знаходяться характерні ознаки, які мають загальне схожість з ознаками інших об'єктів даного класу. Набір таких об'єктів формує навчальну вибірку, в якій відома приналежність кожного об'єкта до певного класу. Набір ознак визначається шляхом проєкції простору ознак в підпростір меншої розмірності. Причому вектори ознак різних класів об'єктів повинні мати максимально відмінні один від одного параметри, щоб звести до мінімуму їх внутрікласові відмінності і збільшити міжкласові. Таким чином, формуються максимально компактні кластери характерних ознак об'єктів, які відповідають різним класам, які мають мінімально можливі перетину один з одним [6].

Завдання класифікації зводиться до того, щоб виділити характерні ознаки вхідного зображення і провести їх кореляцію з наявними ознаками, сформованими в результаті навчання. Таким чином, відбувається визначення приналежності даного об'єкту зображення до певного класу.

### 1.1.7 Алгоритм AdaBoost

Алгоритм AdaBoost (Adaptive boosting – адаптоване поліпшення) був запропонований Йоавом Фройндом і Робертом Шапіре в 1999 році [28]. Даний алгоритм є алгоритмом машинного навчання і успішно

застосовується при вирішенні задач класифікації об'єктів на зображеннях і відео послідовностях. Алгоритм AdaBoost має адаптивний принцип роботи і являє собою каскадну структуру зі слабких класифікаторів, кожен з яких вчиться на помилках попереднього. Кожен слабкий класифікатор орієнтований на певний набір характерних ознак і може видавати два варіанти відповідей: «вірно», або «брехня». У кожному каскаді визначається сума значень слабких класифікаторів і порівнюється із заданим порогом. Таким чином, каскадний класифікатор прибирає з розгляду всі області зображення, в яких відсутні ознаки шуканого об'єкта. В результаті залишаються тільки області, які мають найбільшу ймовірність того, що вони містять шуканий об'єкт. У 2001 році Паул Віола і Майкл Джонс запропонували вдосконалений алгоритм – алгоритм Віоли-Джонса, що був заснований на AdaBoost. Даний алгоритм дозволяє виконувати детектування об'єктів на зображеннях на основі ознак Хаара. Ознаки Хаара є прямокутні області, що складаються з декількох суміжних частин. Різні класи об'єктів мають індивідуальні ознаки, які можна виразити у вигляді розподілу ознак Хаара. У кожній області зображення обчислюється кілька тисяч варіантів розташування ознак, залежно від їхнього економічного становища і масштабу. Відбувається підсумовування інтенсивностей пікселів в чорних і білих областях ознак Хаара і потім обчислюється різниця між сумами [7].

Найбільш часто алгоритм Віоли-Джонса використовується для детектування осіб, але також відомі його застосування для вирішення завдань виявлення автомобільних номерних знаків. Даний алгоритм показує досить високу точність детектування, а так само високу швидкість роботи. Недолік даного алгоритму полягає в відносно не високою мірою інваріантності до афінних і проекційним спотворень об'єктів на зображеннях і інтенсивності освітлення [8].

### 1.1.8 Алгоритм Віоли Джонса

Метод Віоли-Джонса (англ. Viola-Jones object detection) – найбільш популярний алгоритм, що дозволяє з високою швидкістю виявляти області розташування об'єктів на зображеннях. Був запропонований в 2001 році Полом Віолою і Майклом Джонсом. Незважаючи на те, що основне завдання алгоритму полягає в виявленні осіб, він може використовуватися для розпізнавання різних класів об'єктів. Існує безліч реалізацій, в тому числі в складі бібліотеки комп'ютерного зору OpenCV (функція `cvHaarDetectObjects()`) [9].

Основною ідеєю при створенні алгоритму Віоли-Джонса для розпізнавання осіб є виділення локальних особливостей (ознак) зображення і подальшого навчання алгоритму на них. Ознаки, використовувані алгоритмом, використовують підсумовування пікселів з прямокутних регіонів. Ознаки, використані Віолою і Джонсом, містять більше однієї прямокутної області. На рисунку 1.1 показано чотири різних типи ознак.

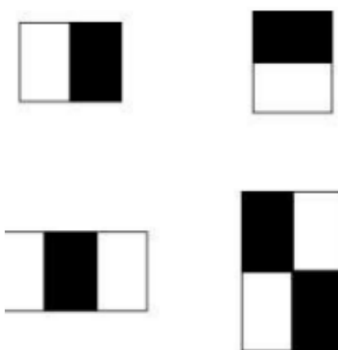


Рисунок 1.1 – Типи ознак для алгоритму Віоли-Джонса

Значення для кожної ознаки обчислюється як віднімання суми пікселів в білих прямокутниках з суми пікселів в чорних. Незважаючи на

чутливість прямокутних ознак до вертикальних і горизонтальних особливостей зображень, вони дуже примітивні, і результат їх пошуку грубий. Проте, за умови збереження зображення в інтегральному форматі (англ. Integral image, коли в кожному пікселі зображення записана сума всіх пікселів, що знаходяться вище і ліворуч), перевірка такої ознаки на конкретній позиції проводиться за константне час, що є перевагою в порівнянні з більше очними варіантами. Кожна прямокутна область в використовуваних ознаках завжди суміжно з іншим прямокутником, тому розрахунок ознаки з 2 прямокутниками складається з 6 звернень в інтегральний 19 масив, для ознаки з 3 прямокутниками – з 8, з 4 прямокутниками – з 9.

Базовий алгоритм Віоли-Джонса має ряд недоліків:

- тривалий час роботи алгоритму навчання. В ході навчання алгоритму необхідно проаналізувати велику кількість тестових зображень;
- велика кількість близько розташованих один до одного результатів через застосування різних масштабів і ковзного вікна.

### 1.1.9 Згорткові нейронні мережі (ЗНМ)

У 1998 році дослідники Y. LeCun, L. Bottou, Y. Bengio і P. Haffner запропонували вид нейронних мереж, що працюють за принципом зорової системи людини, які були названі згортковими нейронними мережами (ЗНМ). ЗНМ являє собою особливий клас багатошарового перцептрона, який володіє двовимірною структурою і добре підходить для обробки зображень з високим ступенем інваріантності до зсуву, поворотам, масштабування і іншим спотворень вхідних даних. Структура ЗНМ являє собою послідовність з двох типів шарів: згорткові і підвбіркові. Кожен шар складається з набору площин (карт характеристик), які в свою чергу складаються з нейронів. Кожен нейрон згорткового шару

має зв'язок з невеликою групою нейронів попереднього шару (локальне рецептивне поле). Локальні рецептивні поля нейронів згорткового шару частково накладаються один на одного за принципом черепиці. Значення нейронів з локального рецептивного поля множаться на матрицю синаптичних коефіцієнтів, а результат записується у відповідний нейрон згорткового шару [8].

На рисунку 1.2 можна побачити архітектуру згорткової нейронної мережі.



Рисунок 1.2 – Архітектура згорткової нейронної мережі

Слідом за згортковим шаром розташовується підвибірковий шар, який забезпечує часткову інваріантність нейронної мережі до зміни масштабу вхідного зображення. Кількість площин підвибіркового шару зазвичай таке ж, як і в попередньому шарі. У підвибірковому шарі локальні рецептивні поля не перетинаються один з одним і мають фіксований розмір  $2 \times 2$  нейрона. Кожен нейрон даного шару обчислює середнє значення своїх чотирьох входів, примножує їх на вагові коефіцієнти синапсів і отриманий результат передає через функцію активації. Таким чином, підвибірковий шар зменшує розмірність площин попереднього шару в два рази [10].

Послідовно чергуючись один за одним, розміри площин зменшуються, але їх кількість збільшується. Чергування шарів дозволяє формувати різні карти характеристик, що наділяє ЗНМ здатністю до ідентифікації більш складних ієрархічних ознак.

Поступово при проходженні декількох шарів карта ознак вироджується в вектор. Останні кілька шарів ЗНМ є класичний персептрон, складаються зі звичайних нейронів і призначені для класифікації виділених ознак.

Важливою властивістю ЗНМ є використання загальної матриці вагових коефіцієнтів для всіх нейронів в межах площині. Даний підхід дозволяє використовувати меншу кількість параметрів, що настраюються (вагових 33 коефіцієнтів) при великій кількості зв'язків і підвищити швидкість процесу навчання. Таким чином, площини ЗНМ є фільтри, кожен з яких здійснює пошук індивідуальних характерних ознак вхідного зображення. Це дозволяє сверточное нейронної мережі запам'ятовувати взаємозв'язок просторово залежних областей зображення. Характерні ознаки, які добуваються тій чи іншої площиною, визначаються в процесі навчання [11].

Якщо вхідне зображення спотворено або зміщене, то на виході площині буде аналогічно зміщений результат. Завдяки цій властивості забезпечується стійкість ЗНМ до спотворень вхідних даних. Недолік використання ЗНМ полягає в складності настройки оптимальних параметрів: кількість шарів, площин, нейронів, розмір рецептивного поля і т.д. Переваги використання ЗНМ для детектування об'єктів на зображеннях [12]:

- структура ЗНМ добре підходить для обробки двовимірних даних;
- нашаровуються один на одного рецептивні поля забезпечують взаємозв'язок просторово залежних областей зображення;
- підвищена стійкість до афінних і проєкційних спотворень вхідних

даних, шумів, зміни масштабу.

Завдяки своїм перевагам на сьогоднішній день ЗНМ активно використовуються для виявлення осіб та інших об'єктів на зображеннях і відео послідовностях. Як зазначено у вступі, здійснювати реалізацію та апробацію алгоритмів було вирішено на прикладі розпізнавання 27 автомобільних номерних знаків. На даний момент ЗНМ не використовувалися для детектування автомобільних номерних знаків і нічого не відомо про подібні експерименти, що створює додатковий інтерес щодо можливості їх застосування для вирішення даного завдання. На підставі проведеного аналізу для вирішення завдання детектування автомобільних номерних знаків в даній роботі було прийнято рішення використовувати алгоритми, засновані на згорткових нейронних мережах [11].

#### 1.1.10 Структурні методи розпізнавання

Структурні методи передбачають розпізнавання символів на основі аналізу їх топології, яка містить інформацію про взаємне розташування окремих складових частин символів. Кожен символ можна представити у вигляді набору характеристик, які можуть бути представлені у вигляді форми графа. На досліджуваному зображенні проводиться аналіз даних характеристик, на основі чого виконується розпізнавання символу на зображенні. На першому етапі дані методи вимагають здійснення попередньої обробки зображення символу, яка полягає в проведенні процедури скелетизації (рис. 1.3).

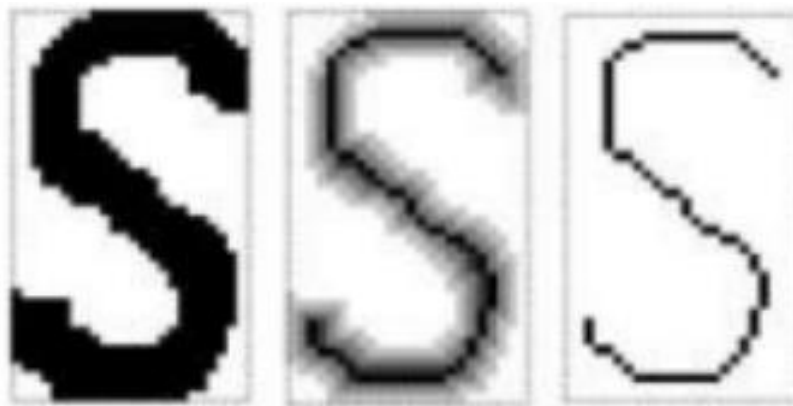


Рисунок 1.3 – Приклад скелетизації символу

Після процедури скелетизації кожен символ представляється у вигляді набору особливих точок, для яких обчислюються наступні параметри: координати особливих точок; довжина ребер графа; нормоване напрямком входів і виходів особливих точок; кривизна дуги, що з'єднує особливі точки.

Так само до структурних методів можна віднести метод, заснований на виділенні кліки максимального розміру в графі відповідності ознак. Клікою графа називається таке підмножина його вершин, в якому будь-які дві вершини з'єднані ребром. Ознаки, виділені на зображенні символу, формуються в граф і порівнюються з еталонним графом для кожного символу. Якщо ознаки сформованого графа відповідають еталонним ознаками, то вершини графа з'єднуються дугами. Після цього в сформованому графі виконується пошук найбільшою максимальною кліки. Символ на досліджуваному зображенні відносять до того класу, граф якого містить найбільшу кількість вершин. Перевагою структурних методів є інваріантність щодо розміру і стилю шрифтів. Недолік структурних методів полягає в низькій швидкості роботи і неможливості розпізнавання символів, схильних до різних спотворень і дефектів (розриви ліній, злиття сусідніх ліній) [13].

### 1.1.11 Морфологічний аналіз

Даний метод розроблений J. Serra і Ю.П. Питьєвим в 70-і роки ХХ століття і застосовується для розпізнавання символів на зображеннях, на основі морфологічного аналізу структурних характеристик кожного символу. Даний метод передбачає вивчення топології кожного символу і опис їх у вигляді індивідуального набору ознак: затоки, протоки, озера. Озером називається структурна частина символу, яка з усіх боків обмежена контуром. Затокою називається структурна частина символу, яка не обмежена контуром тільки, з одного боку. Протокою називається частина символу, яка не обмежена контуром з декількох сторін. Кількість даних ознак, їх розмір і взаємне розташування є унікальним для кожного символу. У роботах даний метод використовувався для розпізнавання символів на автомобільних номерних знаках. Структура кожного символу описується в вигляді еталонного шестивимірної вектора ознак, який містить інформацію про кількість заток, проток і озер [14]. Досліджуване зображення аналізується на виявлення ознак і порівнюється з еталонними векторами. Таким чином, відбувається класифікація символів на зображеннях. Гідність даного методу полягає в низькій чутливості до афінних і проєкційних спотворень символів. Недолік даного методу полягає в великій обчислювальній складності і високої чутливості до шумів на зображенні, в результаті чого структура символів може бути частково спотворена [15].

## 1.2 Дорожні знаки

На даний момент в Україні дорожні знаки розділені на вісім груп:

- попереджувальні знаки;
- знаки пріоритету;

- заборонні знаки;
- розпорядчі знаки;
- знаки особливих приписів;
- інформаційні знаки;
- знаки сервісу;
- знаки додаткової інформації;

Розглянемо основні з них.

Основними знаками в Україні є заборонні дорожні знаки. Заборонні дорожні знаки – забороняють водієві здійснити будь-яку дію, що пропонує на знак. Найчастіше аварійні ситуації трапляються через, недотримання цих знаків. Прикладом будуть знаки обмеження швидкості, заборони обгону, в'їзду, стоянки і зупинки. Як відомо, червоний колір відразу ж звертає на себе увагу, тому що забороняють знаки мають круглу форму і яскраво червоний колір, для того щоб водій зміг його побачити в будь-якому стані і погодних умовах.

На рисунку 1.4 можна побачити приклад заборонних дорожніх знаків:



Рисунок 1.4 – Приклад заборонних дорожніх знаків

Попереджувальні дорожні знаки (рисунок 1.5) – основною функцією цих знаків є попередження водіїв про небезпеку, і застерігати від

дорожньо-транспортних пригод. Такі знаки зазвичай встановлюють на небезпечних ділянках дороги. В основі знака в більшості випадків лежить трикутник з червоним обведенням.



Рисунок 1.5 – Приклад попереджувальних дорожніх знаків

Розпорядчі дорожні знаки – використовуються для того, щоб показувати водіям єдино дозволений дію, будь то проїзд тільки прямо, або тільки поворот наліво (але в цьому випадку дозволено і розворот). Такі знаки дозволяють виконувати тільки певні дії тільки певним учасникам дорожнього руху. На рисунку 1.6 можна побачити приклад розпорядчих знаків.



Рисунок 1.6 – Приклад розпорядчих дорожніх знаків

Знаки пріоритету – встановлюють черговість проїзду перехресть, перехрещень доріг або вузьких ділянок дороги.

Знаки особливих розпоряджень можуть поєднувати елементи як розпорядчих, так і заборонних знаків. Приклад знаків особливих приписів зображено на рисунку 1.7.



Рисунок 1.7 – Приклад знаків особливих розпоряджень

Розпізнавати буде найрозумніше саме ці групи знаків тому, що саме ці знаки мають істотний вплив на дорожню ситуацію. У той же час, з перших п'яти груп найбільшу важливість представляють забороняють дорожні знаки, як, наприклад, знаки обмеження швидкості або заборони обгону. Саме ігнорування заборонних дорожніх знаків, згідно зі статистикою, є найчастішою причиною виникнення ДТП.

### 1.3 Аналіз існуючих систем, що застосовуються для розпізнавання дорожніх знаків

#### 1.3.1 Загальна архітектура

Інформаційна система ідентифікації дорожніх знаків в загальному вигляді складається з наступних апаратних рішень і програмних модулів:

- камера для зйомки зображення;
- апаратний модуль захоплення;
- модуль виявлення;
- модуль класифікації;
- база даних.

Зображення з камери надходить на вхід системи. Потім за допомогою якогось алгоритму визначається положення дорожнього знака. Після цього знак розпізнається програмою розпізнавання. База даних, в залежності від поставлених перед системою завдань, може мати різну схему.

#### 1.3.2 Opel Eye

Система Opel Eye – це система розпізнавання дорожніх знаків. Система Opel Eye працює за наступним принципом: між лобовим склом автомобіля і дзеркалом заднього виду розташована відеокамера з високою роздільною здатністю і широким кутом огляду. Ця камера постійно отримує інформацію про те, що знаходиться попереду автомобіля. Відеокамера здатна передавати до тридцяти кадрів в секунду на два процесори обробки сигналів. У пам'яті закладена база дорожніх знаків. Якщо зображення, яке передається з камери, відповідає якомусь знаку, закладеному в базу даних, то система виводить на панель приладів відповідний знак. Це можуть бути, наприклад, знак із заборонаю обгону,

знак з обмеженням швидкості рух і т.д. На рисунку 1.8 представлений приклад роботи системи Opel Eye.



Рисунок 1.8 – Приклад роботи Opel Eye

Система Opel Eye допоможе водієві: якщо з якої-небудь причини водій не помітив обмеження швидкості (або інший знак дорожнього руху), то система Opel Eye цей знак не упустить і виведе його на панель приладів. Зараз багато автомобілів використовують навігаційну систему, в яку вже закладені повідомлення щодо обмеження швидкості на різних ділянках траси. Система Opel Eye принципово відрізняється від інших навігаційних систем, так як в режимі реального часу вона оцінює дійсну ситуацію на дорозі.

Також, варто відзначити, що система розпізнавання знаків Opel Eye має такі доповнення як FCA і LDW. FCA – це система попередження про фронтальному зіткненні. LDW – це система, яка контролює зміну смуги руху.

Таким чином, при появі небезпеки зіткнення з автомобілем, який рухається попереду або при перетині смуги руху – в салоні машини

прозвучить звуковий сигнал. Також разом зі звуковим сигналом на панель приладів висвітлиться відповідне повідомлення. Точність розпізнавання знаків різко знижується якщо він нахилений або забруднений загальна точність, приблизно 75 відсотків.

### 1.3.3 Speed limit assist (Mercedes)

У нових автомобілях S-, CL-, CLS- і E-класів існує нова система Speed Limit Assist – система допомоги водієві в дотриманні швидкісного режиму. Вона покликана нагадати водієві про обмеження швидкості на прохідному їм в цей момент ділянці дороги. На рисунку 1.9 показаний приклад працюючої системи Speed Limit Assist.



Рисунок 1.9 – Приклад роботи Speed limit assist

Остання версія Speed Limit Assist здатна не тільки розпізнавати дорожні знаки, а й аналізувати чи можна застосувати отриману

інформацію для вантажівок. Крім того, система розпізнає знаки закінчення обмеження швидкісного режиму. У той же час, якщо після знака обмеження швидкості через деякий час не було його скасування, навігаційна система покаже збережені у себе дані для цієї ділянки дороги. Система аналізує отриману з камери на лобовому склі інформацію в режимі реального часу. Місце розташування знака обмеження швидкості – збоку від дороги або вгорі – не має значення. Завдяки доопрацюванню технології розпізнавання знаків, що надходять від камери до системи Speed Limit Assist, аналіз інформації займає не більше секунди, і результат моментально передається на дисплей. На думку німецьких експертів, система допомоги водієві в дотриманні швидкісного режиму допоможе уникнути або, принаймні, знизити кількість дорожньо-транспортних пригод.

Статистика показує, що багато серйозних автомобільні аварії на дорогах викликані перевищенням допустимого швидкісного режиму. Точність дорожніх знаків істотно знижується при забрудненні знака і нахилі. Загальна точність визначення приблизно 70 відсотків.

#### 1.3.4 Road sign information

Система допомагає водієві пам'ятати про дорожні знаки, які він проїхав. Показує знаки обмеження швидкості і обгін заборонений одночасно. Може також показувати знак «Дорога для автомобілів» і «Автомагістраль». При виявленні знака обмеження швидкості, на шкалу швидкості ставиться червоний трикутник зі значенням обмеження. Знаки скидаються при виявленні знака «Кінець обмеження» і т.п. Бувають помилкові спрацьовування. Рисунок 1.10 демонструє роботу системи.



Рисунок 1.10 – Приклад роботи системи Road sign information

Програмний комплекс може розпізнавати такі знаки:

- обмеження швидкості;
- обгін заборонений;
- кінець всіх обмежень;
- автомагістраль / кінець автомагістралі;
- інформаційні таблички (англійською мовою).

Точність розпізнавання різко знижується при нахилі, часткове перекриття або забрудненості знака, його повороті.

### 1.3.5 Система розпізнавання дорожніх знаків лабораторії комп'ютерної графіки та мультимедіа МДУ ім. Ломоносова

Головними напрямками досліджень лабораторії є комп'ютерна графіка, комп'ютерне зір, а також обробка зображень і відео. Так само цією лабораторією була розроблена інформаційна система по ідентифікації знаків дорожнього руху по відео для мобільних платформ. У групі

лабораторії ведуться дослідження за кількома напрямками в рамках даного завдання:

- високошвидкісне виділення дорожніх знаків;
- розпізнавання знаків за допомогою багат шарових нейронних мереж;
- побудова синтетичних колекцій зображень для навчання алгоритмів.

На рисунку 1.11 можна побачити приклад результату розпізнавання цією системою.



Рисунок 1.11 – Приклад результату розпізнавання знака системою

Для виявлення знаків в розробленій системі використовується детектор Віюли-Джонса на основі каскаду класифікаторів, яких навчають за допомогою бустінга. Розробка була протестована на загальнодоступній базі бельгійських знаків, що складається більш ніж з 4000 зображень. Для забороняючих і розпорядчих знаків точність склала 96.5% і 95.5% відповідно, при 10-9 помилкових спрацьовувань на вікно детектора.

### 1.3.6 Порівняння технічних характеристик існуючих систем розпізнавання дорожніх знаків

У таблиці 1.1 представлена порівняльна характеристика існуючих систем розпізнавання.

Таблиця 1.1 – Порівняльна характеристика існуючих систем розпізнавання

Параметр	Opel Eye	Speed Limit Assist	Road sign information	Лабораторія комп. графіки та мультимедіа ім. Ломоносова
Заявлена точність розпізнавання	90%	95%	96%	95%
Освітленість знака не менше	50	50	50	50
Розпізнавання знаків обмеження швидкості	Да	Да	Да	Да
Розпізнавання інших заборонних знаків	Да	Нет	Да	Да
Інша інформація	Початок роботи від швидкості руху >60			Відкрита система, в основі алгоритм Віоли Джонса
Реальна точність розпізнавання	75%	70%	75%	95%

З таблиці можна зробити висновок, що окремо реалізовані інформаційні системи для розпізнавання дорожніх знаків дуже рідкісні, більшість ІС не є окремо встановленої функцією, а в основному вони входять до складу заводських опцій автомобіля. З цього випливає, що їх комерційна вартість занадто велика. Всі встановлені системи показують невеликий відсоток ідентифікації дорожніх знаків, відсоток розпізнавання падає, якщо дорожній знак знаходиться не в ідеальних умовах (нахил, часткове перекриття, забруднення знака) так як зображення терплять проектні спотворення. Під час випробувань надійно і точно розпізнаються лише чисті знаки з контрастним контуром. Дізнатися які алгоритми лежать в основі цих ІС не представляється можливим так як у вільному доступі ця інформація не поширюється.

Розробка методу ідентифікації дорожніх знаків стійких до таких перетворень, підвищить точність розпізнавання, що є актуальним при:

- нахиленому дорожньому знаку;
- маневруванні автомобіля в зоні ідентифікації;
- високій швидкості руху автомобіля.

#### 1.4 Постановка задачі

Підсумком магістерського проекту винна стати система для розпізнавання дорожніх знаків. Метою створення системи є створення програмного комплексу, який дає наступні можливості:

- виконання попередньої обробки цифрових зображень для здійснення подальших операцій з ними;
- можливість настройки класифікаторів, на основі навчальних даних, які зберігаються в зовнішніх файлах;
- здійснення нормалізації зображень символів;
- розпізнавання символів на зображеннях.

Виходячи із сучасних вимог до якості програмного забезпечення, що розробляється програмна система повинна володіти такими характеристиками:

- коректна робота програмної системи в вирішенні поставлених завдань.

- стійкість програмної системи до різних помилок, збоїв та дій, що виходять за рамки її функціонального призначення.

- невелика обчислювальна складність використовуваних алгоритмів.

- можливість розширювати, змінювати і доповнювати функціональність розроблених бібліотек. для забезпечення цього пункту необхідно слідувати принципам простоти побудови програми і децентралізації.

- програмна система повинна володіти модульною архітектурою, де кожен модуль призначений для вирішення певних завдань

- можливість повторного використання програмних модулів. Забезпечує скорочення часу розробки програмної системи і скорочення розміру програмного коду.

Для цього необхідно провести повний аналіз існуючих методів і алгоритмів для ідентифікації об'єктів на зображеннях зі складним фоном. Необхідно описати характеристики предметної області (дорожні знаки) дослідження. Також важливим моментом є характеристика існуючих рішень в області розпізнавання дорожніх знаків на зображенні. Після цього буде запропоновано власний метод для розпізнавання зображень, який буде представляти із себе синтез двох існуючих методів, таких як метод опорних векторів і згорткової нейронної мережі.

## 2 РОЗРОБКА АЛГОРИТМІВ ВИЯВЛЕННЯ І РОЗПІЗНАВАННЯ ДОРОЖНІХ ЗНАКІВ НА ЗОБРАЖЕННІ

### 2.1 Розробка методу розпізнавання об'єктів на зображенні

При вирішенні складних завдань класифікації виникає необхідність в реалізації алгоритмів, які ефективно вирішують поставлені завдання. Однак якість реалізованих алгоритмів часто залишає бажати кращого. У подібних ситуаціях має сенс реалізувати алгоритмічну композицію (boosting – поліпшення), що складається з декількох алгоритмів, кожен з яких намагається компенсувати помилки класифікації, отримані на попередніх ітераціях. У багатьох дослідженнях спостерігалось, що при нарощуванні алгоритмічних композицій відбувається інтенсивне зменшення частоти помилок на незалежних тестових наборах даних, а також збільшення швидкості роботи системи. Дані факти змінили склалося уявлення про необхідність обмеження складності алгоритмів. В результаті багатьма дослідниками був зроблений висновок, що алгоритмічні композиції не знижують ефективність роботи системи, а лише згладжує похибка базових алгоритмів.

На сьогоднішній день побудова алгоритмічних композицій, в яких різні алгоритми компенсують недоліки один одного, є одним з найбільш перспективних напрямків машинного навчання. Переваги такого підходу полягають в універсальності, гнучкості, високої узагальнюючої здатності і можливості отримання якісного навчання, недосяжного для окремих базових алгоритмів.

Детектування області розташування об'єктів, в даному випадку дорожніх знаків є досить складним завданням в зв'язку з наявністю на зображеннях складної фонові структури та шумових перешкод. При вирішенні даного завдання виникає необхідність реалізації алгоритму, що

виконує кілька етапів класифікації і працює за принципом послідовної класифікації. На підставі цього, було прийнято рішення розробити алгоритм детектування автомобільних номерних знаків у вигляді алгоритмічної композиції, що складається з методу опорних векторів і згорткової нейронної мережі.

Алгоритмічна композиція будувалася за принципом послідовної класифікації, де спочатку з допомогою методу опорних векторів відбувається перше розпізнавання, яке передається далі на нейронну мережу. Така модель розпізнавання значно збільшує швидкість ідентифікації, а також підвищує точність розпізнавання.

## 2.2 Метод опорних векторів

Метод опорних векторів (Support Vector Machine, SVM) був запропонований російськими вченими В. Вапніком і А. Червоненкісом. Даний метод є лінійним класифікатором і застосовується для вирішення завдань класифікації та регресійного аналізу. Застосування методу SVM до вирішення задачі класифікації об'єктів на зображеннях можна описати таким чином. Дано безліч точок в просторі, які згруповані за двома класами. Для того, щоб розділити безліч точок на класи можна провести розділяє гіперплощина. З точки зору класифікації необхідно знайти оптимальну гіперплощина, відстань від якої до кожного класу буде максимально. Чим більше відстань від розділяє гіперплощину до кожного класу, тим менше помилка класифікації. Вектори, розташовані в просторі найближче до гіперплощини, називаються опорними векторами [16]. Необхідно знайти класифікуючу функцію  $f(x)$ , яка приймає різні значення для векторів різних класів. Для визначення даної функції використовується навчальний набір даних  $(x_1, y_1), \dots, (x_t, y_t)$ . Функція, при якій очікувана помилка класифікації приймає мінімальне значення, є найбільш

оптимальною функцією. Класифікує функція має вигляд (2.1):

$$f(x) = \text{sign}((w, x) + b), \quad (2.1)$$

де  $x$  – об'єкт, що належить простору;

$w$  – вектор нормалі до гіперплощини;

$b$  – допоміжний параметр.

Об'єкти, для яких функція  $f(x)$  приймає значення рівні 1 або -1, відносяться до різних класів. Необхідно вибрати параметри  $w$  і  $b$  таким чином, щоб відстань до кожного класу було максимальним. На практиці часто виникають випадки, коли дані неможливо розділити лінійним способом.

Для вирішення такої ситуації виконується перетворення початкового простору  $x$  в простір більш високої розмірності з допомогою спеціального відображення  $\varphi: R^n \rightarrow X$ . Відображення  $\varphi$  вибирається таким чином, щоб в отриманому просторі безліч даних було лінійно нероздільні. Основна складність полягає в тому, що зі збільшенням розмірності простору, збільшується складність процесу класифікації [17].

На практиці випадки, де дані в задачі розпізнавання можуть бути розділені лінійно, досить рідкісні. У таких випадках застосовується метод використання ядер, запропонований Б. Босером, І. Гийон і В. Вапніка, і полягає в тому, що елементи, що навчається вибірки вкладаються в простір  $x'$  більш високої розмірності з допомогою спеціального відображення  $\varphi: R^n \rightarrow x'$ . При цьому відображення вибирається так, щоб в просторі  $x'$  вибірка була роздільна лінійно. Ядром класифікатора називається вираз  $k(x, x') = (\varphi(x), \varphi(x'))$ , що задає відображення вибірки в новий простір, і його роль, як правило, може виконувати будь-яка позитивно певна симетрична функція двох змінних. На практиці зустрічаються такі ядра: радіальна базисна функція, гауссова базисна функція, сигмоїда [18].

Навчання моделі, що використовує метод опорних векторів проводиться методами квадратичного програмування, такими як послідовна мінімальна оптимізація. Метод опорних векторів має деякі переваги і недоліки по відношенню до використання багат шарових перцептронів:

- багат шаровий перцептрон є модель з безліччю прихованих параметрів, що залежать від числа нейронів мережі. Параметризована модель потенційно здатна до інкапсуляції більш складних, високорівневих функцій, але при цьому вимагає більше часу і обчислювальних ресурсів для навчання та налаштування параметрів. Метод опорних векторів використовує вектори, відібрані з навчальної вибірки, при цьому кількість параметрів обмежена зверху розміром вибірки, а на практиці може бути проріджені за рахунок використання інженерії ознак;

- на відміну від навчання нейронної мережі, яке здійснюється за допомогою методу градієнтного спуску (і його варіацій) і оцінки помилки мережі, навчання моделі опорних векторів включає в себе не тільки оцінку помилки, але і метрику складності отриманої гіперплощини. Пошук оптимального значення нейронної мережі вразливий до наявності локального мінімуму, здатного зупинити процес градієнтного спуску, при цьому метод опорних векторів при коректному виборі метепараметрів гарантує знаходження глобального рішення;

- навчена нейронна мережа вимагає мінімальних обчислювальних ресурсів для роботи в режимі розпізнавання (передбачення категорій). Метод опорних векторів в деяких випадках, коли число векторів велике в порівнянні з розміром вибірки, буде передбачення істотно повільніше;

- в порівнянні з нелінійним (використовують ядра) методом опорних векторів, нейронна мережа демонструє розширені можливості до онлайн-навчання, коли розмір вибірки не фіксований і поповнюється за рахунок надходження нових даних.

У більшості сучасних додатків алгоритмів розпізнавання і машинного навчання зараз віддано перевагу методу опорних векторів за рахунок скорочення часу навчання та стійкості до локального мінімуму.

Метод опорних векторів також широко використовується для розпізнавання зображень, таких як людські особи, демонструючи високу точність розпізнавання (80-85% успішно розпізнаних зображень) для вирівнюється вибірки. Особливість завдання розпізнавання зображень полягає в тому, що дані, що представляють собою візуальні сигнали, демонструють вкрай низьку інформаційну ємність – тобто, більша частина точок растрового зображення (наприклад, відповідні ділянкам однотонного або рівномірно розподіленого фону) не містить інформації, що впливає на розпізнавання .

При цьому розмірність зображень, що використовуються в системах обробки інформації, як правило, досить велика – сучасні засоби мультимедіа, графічні дисплеї і сенсори забезпечують масове поширення зображень (фотографій, кадрів відео, комп'ютерної графіки) з високою роздільною здатністю, розмірність яких вимірюється мільйонами точок.

Для класичних методів розпізнавання образів характерна пряма залежність між розмірністю (числом параметрів) даних навчальної вибірки і часом навчання, а також показниками збіжності при оптимізації моделі. Наявність великої кількості параметрів, основна частина яких не містить істотної для розпізнавання інформації, негативно впливає на продуктивність моделі, і крім вимоги значно більш високих обчислювальних ресурсів веде до появи так званої проблеми перенавчання, коли функція розпізнавання, апроксимована моделлю, задовільно класифікує навчальну вибірку, але є при цьому не генералізованою і демонструє низьку точність в тестовій вибірці. Для вирішення цієї проблеми використовується підхід пошуку компактного представлення зображення – виділення обмеженого числа генералізованих ознак, що

містять основну інформацію, необхідну для розпізнавання.

Даний метод активно застосовується для детектування різних об'єктів, в тому числі і дорожніх знаків. Недолік даного методу полягає в тому, що в процесі класифікації враховується не всі безліч даних, а тільки лише частина, що знаходиться на кордонах безлічі. Перевага методу SVM полягає в тому, що для класифікації досить невеликого набору даних [19].

### 2.3 Опис алгоритму

Перевагою методу опорних векторів в порівнянні з іншими методами ідентифікації об'єктів на зображенні є високий показник розпізнавання на зображеннях з шумами і спотвореннями, що передбачає високу точність розпізнавання. Головним недоліком даного методу є те, що параметри моделі складно інтерпретувати.

Запропонований метод ідентифікації дорожніх знаків на зображенні являє собою комбінацію двох методів: методу опорних векторів і згорткових нейронних мереж. Комбінація будується наступним чином: метод SVM забезпечує прийняття рішення, чи знаходиться на зображенні дорожній знак чи ні, в разі його присутності зображення надходить далі на вхід згорткової нейронної мережі, яка формує остаточне рішення про наявність дорожнього знака на зображенні. На рисунку 2.1 представлена функціональна блок–схема комбінованого методу.

Синтез цих двох методів дозволяє розраховувати на високий відсоток правдивих розпізнавань, а також на ідентифікацію при різних кутах спотворення, а також погані погодні умови.

Алгоритм для розпізнавання дорожніх знаків складається з двох етапів: навчання та розпізнавання.

Першим етапом алгоритму буде навчання якої складається з декількох кроків:

- виділення компонентів зображення дорожнього знака в навчальній вибірці за допомогою методу опорних векторів;
- зменшення розмірності векторів виділених ознак методом опорних векторів;
- навчання згорткової нейронної мережі на отриманих векторах ознак.

На другому етапі згорткова нейронна мережа повторно виконує перевірку на присутність дорожнього знака на зображенні, тим самим дозволяючи прибрати помилки розпізнавання [20].

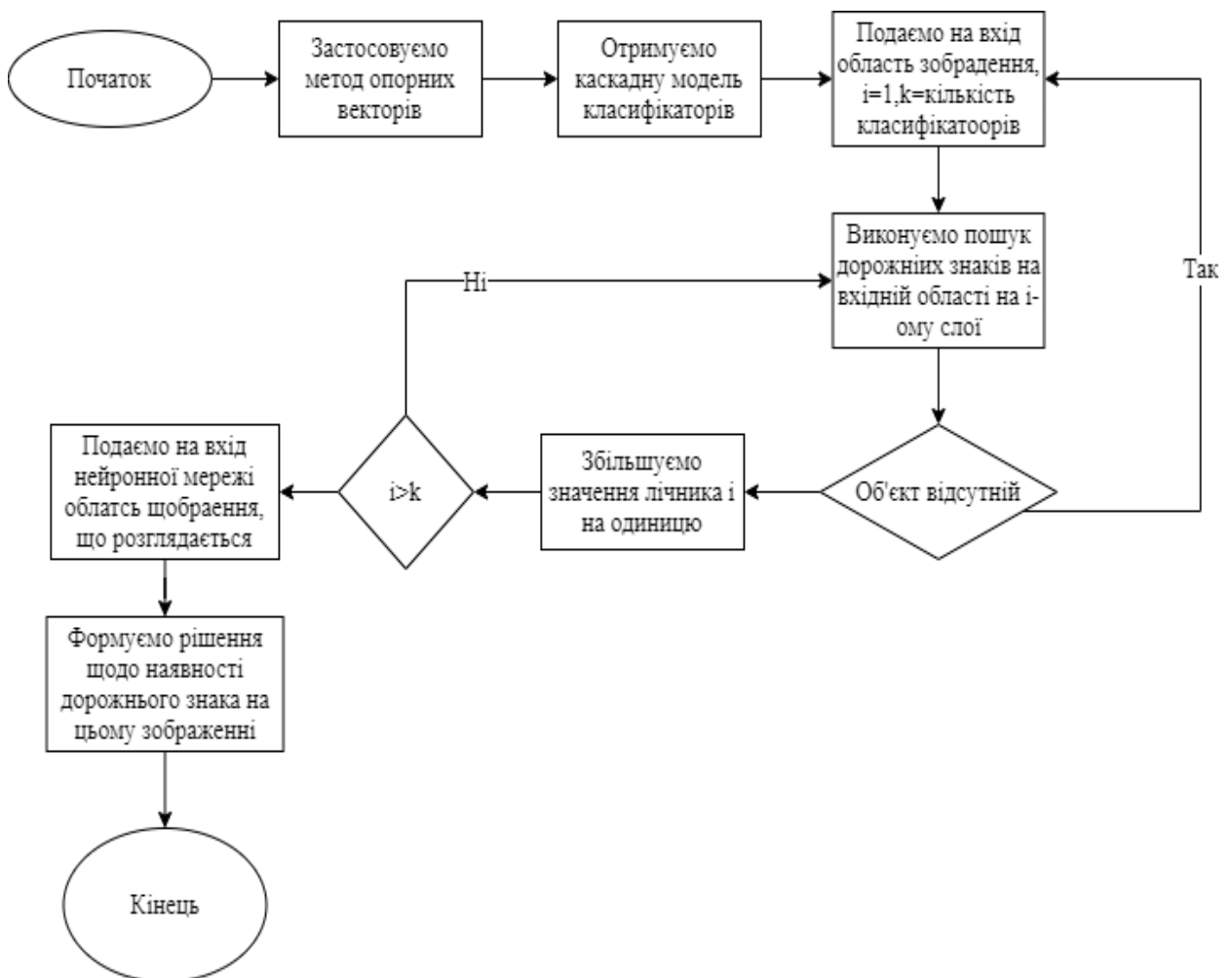


Рисунок 2.1 – Функціональна схема комбінованого методу виявлення дорожніх знаків

## 2.4 Класифікація за допомогою згорткової нейронної мережі

Для повторної класифікації дорожніх знаків на зображеннях розроблена мережа. Дана нейронна мережа має структуру: меншу кількість площин, нейронів і зв'язків. На рисунку 2.2 показана структура використаної в роботі згорткової нейронної мережі.

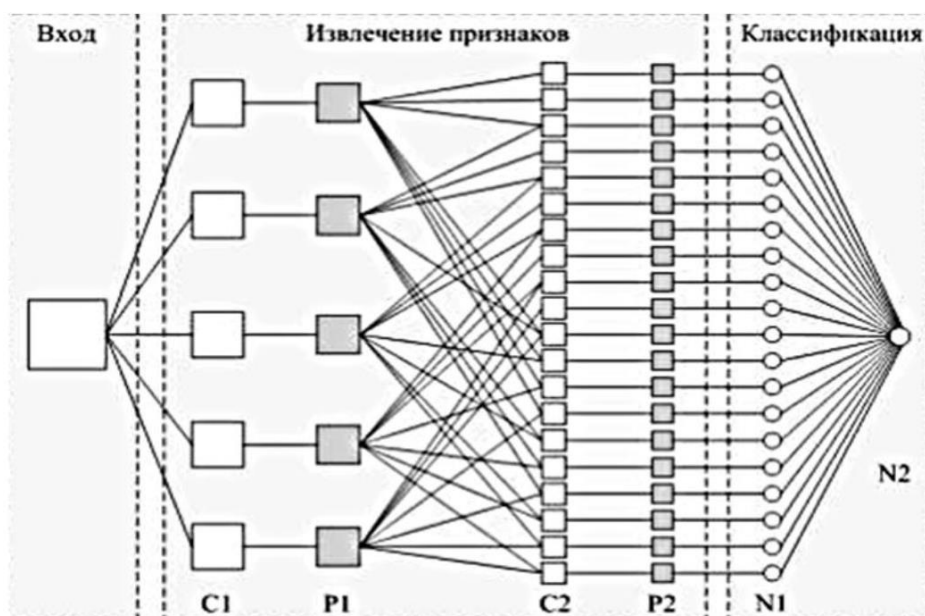


Рисунок 2.2 – Структура згорткової нейронної мережі

Перший шар згорткової нейронної мережі є шаром згортки, складається з 32 згорткових площин розміром 32x32 нейронів. Кожна площина даного шару має власну матрицю синаптичних коефіцієнтів і таким чином виконує шістнадцять згорток вхідного зображення. Значення нейронів поточного шару обчислюються при послідовному скануванні вхідного шару рецептивних полем і множенні значень назв нейронів на матрицю вагових коефіцієнтів. Розмір локального рецептивного поля дорівнює 5x5 нейронів. Кожній площині відповідає  $5 * 5 = 25$  вагових коефіцієнтів і нейронів зміщення.

Розмір площин згорткового шару визначається наступними формулами (2.2):

$$s_k = s_p - N + 1, h_k = h_p - M + 1, \quad (2.2)$$

де  $h_k$  – параметр висоти площині згорткового шару;

$s_k$  – параметр ширини площині згорткового шару;

$h_p$  – параметр висоти площині попереднього шару;

$s_p$  – параметр ширини площині попереднього шару;

$M, N$  – параметри висоти і ширини локального рецептивного поля.

Функціонування нейрона згорткового шару визначається наступною формулою (2.3):

$$y_q^{(i,j)} = a_q + \sum_{s=1}^N \sum_{t=1}^M w_{q,s,t} x_{q-1}^{(i,j)}, \quad (2.3)$$

де  $y_q^{(i,j)}$  – нейрон згорткового шару,  $q$ -ої площині;

$a_q$  – нейрони зміщення  $q$ -ої площині;

$N, M$  – ширина і висота локального рецептивного поля;

$x_{q-1}^{(i,j)}$  – вихідні значення нейронів попереднього шару;

$w_{q,s,t}$  – вагові коефіцієнти синапсів [21].

Другий шар є шаром підвибірки і складається з 16 під вибірових площин. Кожна з цих площин має зв'язок тільки з однією відповідної їй площиною попереднього шару нейронної мережі. Даний шар забезпечує локальне усереднення площин попереднього шару, тому розмір площин даного шару вдвічі менше, ніж в попередньому шарі і дорівнює 16x16 нейрона. Рецептивні поля суміжних нейронів не перетинаються і мають розмір 2x2 нейрона, що забезпечує часткову інваріативність нейронної мережі до зміни масштабу вхідного зображення. Кожній площині

відповідає один синаптичний коефіцієнт. Кожен нейрон даного шару обчислює максимальне значення нейронів з відповідною йому області нейронів попереднього шару. Після здійснення операції підвибірки підвищується інваріативність нейронної мережі до зміщення, зміни масштабу вхідного зображення і точне розташування ознак об'єктів на зображенні стає менш важливим. Розмір площин під вибіркового шару визначається наступними формулами (2.4):

$$S_k = \frac{S_p}{N}, h_k = \frac{h_p}{M}, \quad (2.4)$$

де  $h_k$  – параметр висоти площині згорткового шару;

$S_k$  – параметр ширини площині згорткового шару;

$h_p$  – параметр висоти площині попереднього шару;

$S_p$  – параметр ширини площині попереднього шару;

$M, N$  – параметри висоти і ширини локального рецептивного поля.

Функціонування нейрона підвибіркового шару визначається наступною формулою (2.5):

$$y_q^{(i,j)} = a_q + \frac{1}{NM} w_q \sum_{s=1}^N \sum_{t=1}^M x_{q-1}^{(i,j)}, \quad (2.5)$$

де,  $y_q^{(i,j)}$  – нейрон підвибіркового шару, q-ої площині;

$a_q$  – нейрони зміщення q-ої площині;

$N, M$  – ширина і висота локального рецептивного поля;

$x_{q-1}^{(i,j)}$  – синаптичний коефіцієнт.

Третій шар є шаром згортки і складається з 64 згорткових площин розміром 16x16 нейронів. Таким чином, здійснюється згортка кожної з 32 площин попереднього шару. Розмір локального рецептивного поля дорівнює 3x3 нейрона. Кожній площині відповідає  $3 * 3 = 9$  вагових

коефіцієнтів і нейронів зміщення. Таким чином, даний шар містить 16384 нейронів і  $(9 + 1) * 64 = 640$  вагових коефіцієнтів [22].

Четвертий шар є шаром підвибірки, складається з 64 підвибіркових площин. Розмір площин даного шару в два рази менше розміру площин попереднього шару і дорівнює 16x16 нейронів. Кожна площина має зв'язок тільки з однією відповідної їй площиною попереднього шару. Рецептивні поля суміжних нейронів не перетинаються і мають розмір 2x2 нейрона. Кожній площині відповідає один синаптичних коефіцієнт і нейронів зміщення. Даний шар міститься 16384 нейронів і 640 синаптичних вагових коефіцієнтів.

П'ятий шар є шаром згортки і складається з 128 згорткових площин розміром 8x8 нейронів. Таким чином, здійснюється згортка кожної з 64 площин попереднього шару. Даний шар не є повнозв'язним через те, що площини повинні отримувати різні характерні ознаки і отримувати на вхід різні набори даних. В результаті виконання цих операцій забезпечується здатність нейронної мережі об'єднувати різні види характеристик, які в меншій мірі залежать від спотворень вхідного зображення. Розмір локального рецептивного поля дорівнює 3x3 нейрона. Кожній площині відповідає  $3 * 3 = 9$  вагових коефіцієнтів і нейронів зміщення. Таким чином, даний шар містить 8192 нейронів і  $(9 + 1) * 128 = 1280$  вагових коефіцієнтів [22].

Шостий шар є шаром підвибірки, складається з 128 підвибіркових площин. Розмір площин даного шару в два рази менше розміру площин попереднього шару і дорівнює 8x8 нейронів. Кожна площина має зв'язок тільки з однією відповідної їй площиною попереднього шару. Рецептивні поля суміжних нейронів не перетинаються і мають розмір 2x2 нейрона. Кожній площині відповідає один синаптичних коефіцієнт і нейронів зміщення. Даний шар міститься 8192 нейронів і 1280 синаптичних вагових коефіцієнтів.

Шість перших шарів призначені для вилучення ознак зображення. Наступні два шари є шарами класифікації. Сьомий шар складається з 128 нейронів. Восьмий шар є вихідним шаром і складається з 43 нейронів. Роль даного шару полягає в обчисленні остаточного результату класифікації. Вихідні значення цього нейрона знаходяться в діапазоні від 1 до 43, що означає приналежність знайденого на зображенні знака до одного з класів дорожніх знаків.

Завдяки використанню загальних синаптичних вагових коефіцієнтів в площинах згорткових нейронних мереж досягається зменшення кількості параметрів, що настраюються. На рисунку 2.3 представлена конфігурація згорткової нейронної мережі.

```

Model: "sequential"
-----
Layer (type)                Output Shape          Param #
-----
conv2d (Conv2D)              (None, 32, 32, 32)   896
-----
max_pooling2d (MaxPooling2D) (None, 16, 16, 32)   0
-----
conv2d_1 (Conv2D)            (None, 16, 16, 64)   18496
-----
max_pooling2d_1 (MaxPooling2 (None, 8, 8, 64)     0
-----
conv2d_2 (Conv2D)            (None, 8, 8, 128)    73856
-----
max_pooling2d_2 (MaxPooling2 (None, 4, 4, 128)    0
-----
flatten (Flatten)            (None, 2048)          0
-----
dense (Dense)                 (None, 128)           262272
-----
dense_1 (Dense)               (None, 43)            5547
-----
Total params: 361,067
Trainable params: 361,067
Non-trainable params: 0

```

Рисунок 2.3 – Конфігурація згорткової нейронної мережі

При використанні класичних нейронних мереж кожному ділянці зображення розраховується окремо від усіх інших ділянок, це призводить до великих обсягів обчислень і зниження швидкості роботи системи. У згорткових нейронних мережах даний недолік усувається використанням загальних синаптичних коефіцієнтів і так званого методу загальних обчислень. Локальні рецептивні поля згорткових верств частково нашаровуються один на одного і мають загальні значення нейронів. Тому годі й обчислювати значення всіх нейронів локального рецептивного поля, а використовувати значення, обчислені на попередніх ітераціях. Використання цього методу дозволяє значно знизити час і обсяг обчислень.

## **3 РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ РОЗПІЗНАВАННЯ ДОРОЖНІХ ЗНАКІВ НА ЗОБРАЖЕННІ**

### **3.1 Загальні вимоги до програмного забезпечення, що розробляється**

Виходячи із сучасних вимог до якості програмного забезпечення, що розробляється, програмна система повинна володіти такими характеристиками:

- коректна робота програмної системи в вирішенні поставлених завдань;
- стійкість програмної системи до різних помилок, збоїв та дій, що виходять за рамки її функціонального призначення;
- невелика обчислювальна складність використовуваних алгоритмів;
- можливість розширювати, змінювати і доповнювати функціональність розроблених бібліотек для забезпечення цього пункту необхідно слідувати принципам простоти побудови програми і децентралізації;
- програмна система повинна володіти модульною архітектурою, де кожен модуль призначений для вирішення певних завдань;
- можливість повторного використання програмних модулів забезпечує скорочення часу розробки програмної системи і скорочення розміру програмного коду.

У відповідності з поставленими завданнями до розроблюваної програмної системи пред'являються наступні функціональні вимоги:

- виконання попередньої обробки цифрових зображень для здійснення подальших операцій з ними;

- можливість налаштування класифікаторів, на основі навчальних даних, які зберігаються в зовнішніх файлах;
- здійснення нормалізації зображень символів;
- розпізнавання символів на зображеннях.

### 3.2 Вибір засобів розробки програмного забезпечення

В даний час існує величезна кількість мов програмування, а також програмних бібліотек для роботи з штучними нейронними мережами та генетичними алгоритмами, реалізованих на різних мовах програмування, таких як C ++, C #, Python і ін. В якості інструменту для вирішення задач, і реалізації алгоритмів вибрана мова програмування Python. Python – високорівнева мова програмування загального призначення, орієнтований на підвищення продуктивності розробника і читання коду. Синтаксис ядра Python мінімалістичний. У той же час стандартна бібліотека включає великий обсяг корисних функцій. Python підтримує структурний, об'єктно-орієнтоване, функціональне, імперативне і аспектно-орієнтоване програмування. Основні архітектурні риси – динамічна типізація, автоматичне керування пам'яттю, повна інтроспекція, механізм обробки виключень, підтримка багатопоточних обчислень, високорівневі структури даних. Підтримується розбиття програм на модулі, які, в свою чергу, можуть об'єднуватися в пакети. Виходячи з цього, в якості мови програмування була обрана мова програмування Python, як середовище програмування було обрано програмне забезпечення Microsoft Visual Studio Code.

Для вирішення завдань машинного навчання існує безліч бібліотек і систем, що підтримують інтерфейс мови Python. Деякі з них: Apache Singa, Caffe, Keras, Microsoft Cognitive Toolkit, MXNet, TensorFlow, Theano, PyTorch. Для реалізації розпізнавання обрана бібліотека Keras – відкрита

нейромережева бібліотека, написана на мові Python – вона націлена на оперативну роботу з мережами глибинного навчання, при цьому спроектована так, щоб бути компактною, модульною та розширюється та OpenCV – бібліотека функцій та алгоритмів комп'ютерного зору, обробки зображень і чисельних алгоритмів загального призначення з відкритим кодом.

Щоб досягти максимальних результатів від навчання мережі для навчальної вибірки був обраний пакет даних German Traffic Sign Recognition Benchmark (GTSRB), що складається з близько 51 839 зображень, з різними розмірами від 15x15 до 250x250 пікселів. Всі вони містять один з 43 видів дорожніх знаків. Зображення вже розділені в співвідношенні 1: 3 на тестові і тренувальні дані відповідно. Тренувальні дані вже відсортовані на основі класу, зображених на них дорожніх знаків, і знаходяться в 43 папках [23]. GTSRB містить в собі безліч зразків знаків одного виду, які відрізняються один від одного розмірами, інтенсивністю кольору, нахилом і кутом повороту знака на зображенні. Продемонстровані в GTSRB зображення відносяться до так званих позитивних прецедентів, це означає, що вони містять зображення дорожнього знака, а негативними прецедентами вони будуть, якщо вони не містять зображень з дорожніми знаками. У GTSRB існують кадри, які не містять дорожні знаки, причому вони відібрані так, що є шанс прийняти їх за шуканий дорожній знак. Їх використовують з метою, розходження позитивних від негативних прецедентів [24].

### 3.3 Інтерфейс програмного забезпечення

При запуску програмної системи відкривається форма головного інтерфейсу, на рисунку 3.1 показаний знімок екрана з головним вікном програми.

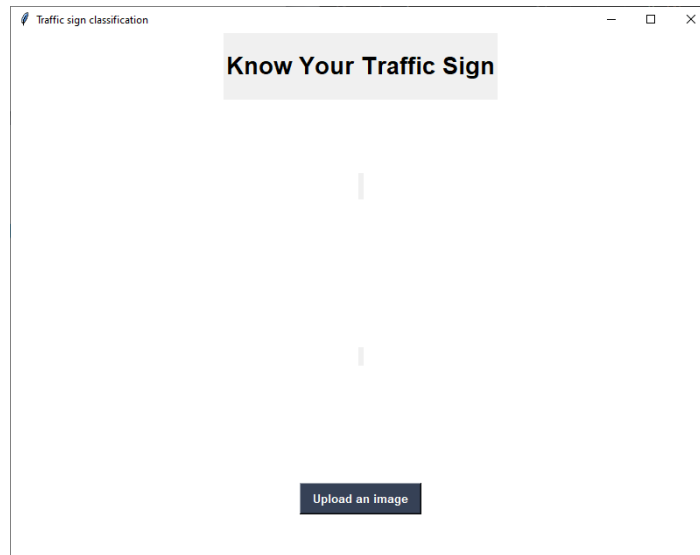


Рисунок 3.1 – Головне вікно системи розпізнавання дорожніх знаків

У правому верхньому куті вікна програми знаходяться стандартні кнопки для кожного програмного продукту:

- згорнути;
- розгорнути в повне вікно;
- закрити програму.

Також є заголовок «Traffic sign Classification» програми (рис.3.2).

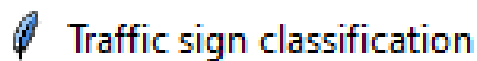


Рисунок 3.2 – Заголовок програми

Кнопка «Upload an image» відкриває вікно імпорту тестових зображень для перевірки працездатності програми. На рисунку 3.3 показана тестова вибірка зображень.

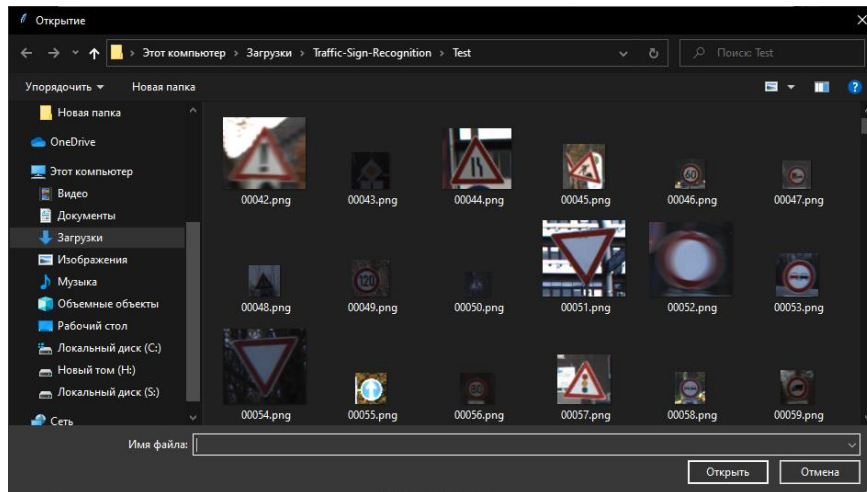


Рисунок 3.3 – Імпорт зображень в програму

Після вибору зображення воно відображається посередині вікна програми та з'являється кнопка «Classify Sign» – рисунок 3.4.



Рисунок 3.4 – Вікно програми після імпорту тестового зображення

Після натискання кнопки «Classify Sign» і вибору тестового зображення, відбувається розпізнавання дорожнього знака на зображенні.

За випадку успішного розпізнавання назва класу дорожнього знака з'являється над зображенням (рис. 3.5).



Рисунок 3.5 – Вікно програми з результатом розпізнавання дорожнього знака

Отримані результати точності створеної моделі представлені на рисунку 3.6.

Accuracy score:  
0.9238321456848773

Рисунок 3.6 – Точність створеної моделі

Технічні характеристики ПК, використаного для обчислювальних елементів:

Таблиця 3.1 – Характеристики ПК

Назва параметра	Характеристика
Модель	ASUS UX501V
ОС	Microsoft Windows 10
Тип системи	64-розрядна ОС, процесор x64
Процесор	Intel Core i7-6700HQ
Частоти процесору	2.6 ГГц – 3,5 ГГц
Тип оперативної пам'яті	DDR4
Об'єм оперативної пам'яті	15,9 GB
Частота оперативної пам'яті	2133 МГц
Відеокарта	NVIDIA GeForce GX 960M
Тип пам'яті відеокарти	GDDR5

#### 3.4 Оцінка якості роботи програмної системи

У даному розділі представлені результати тестування розробленої програми. Експериментальна оцінка якості роботи програмної системи здійснювалася в чотири етапи:

- оцінка якості розпізнавання дорожніх знаків на зображеннях де такі відсутні;
- оцінка якості розпізнавання дорожніх знаків при різних кутах відхилення, щодо реєструючого пристрою;
- оцінка якості розпізнавання декількох дорожніх знаків в одному кадрі.

### 3.4.1 Оцінка якості розпізнавання дорожніх знаків на помилкових зображеннях

Для оцінки даного параметра використовувався набір із зображень, які не містять дорожні знаки, а також зображення, які містять дорожні знаки, але які не були в навчальній вибірці додатки.

Всього є 30 помилкових зображень для перевірки роботи програми, представлені на рисунку 3.7.

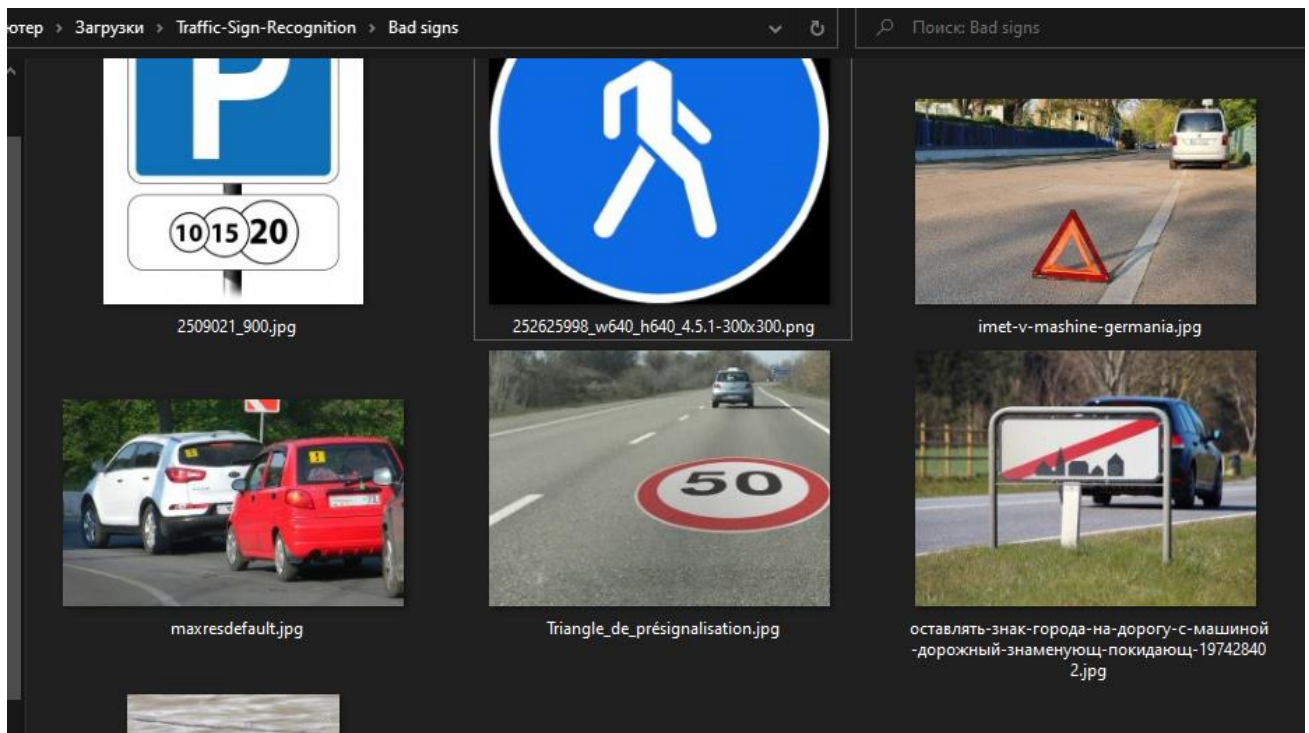


Рисунок 3.7 – Помилкові зображення

На рисунку 3.8 видно, що програма не знаходить на зображенні дорожній знак так як його немає в навчальній вибірці, але за параметрами (форма і колір) він підходить під опис дорожнього знака.



Рисунок 3.8 – Не розпізнаний програмою дорожній знак

Незважаючи на високий відсоток «правильного» розпізнавання об'єктів іноді програма виводить помилкові ідентифікації, що видно з рисунка 3.9.

Програма визначила дорожній знак, але це не є вірним результатом розпізнавання. Це сталося через схожість кольору і форми і буде розглядатися як помилкове розпізнавання.



Рисунок 3.9 – Приклад помилкової ідентифікації програми

В цілому програма добре себе показала в помилковій спрацьовування з 30 зображень, помилково ідентифікували всього три, що дає результат в 90% правдивих розпізнань.

### 3.4.2 Оцінка якості розпізнавання дорожніх знаків при різних кутах відхилення

Щоб оцінити на скільки ПО справляється із завданням розпізнавання дорожніх знаків при різних кутах відхилення, була використана тестова вибірка, що складається з 40 зображень на якій дорожні знаки в різному ступені відхилені щодо реєструючого пристрою, приклад див. на рис. 3.10.

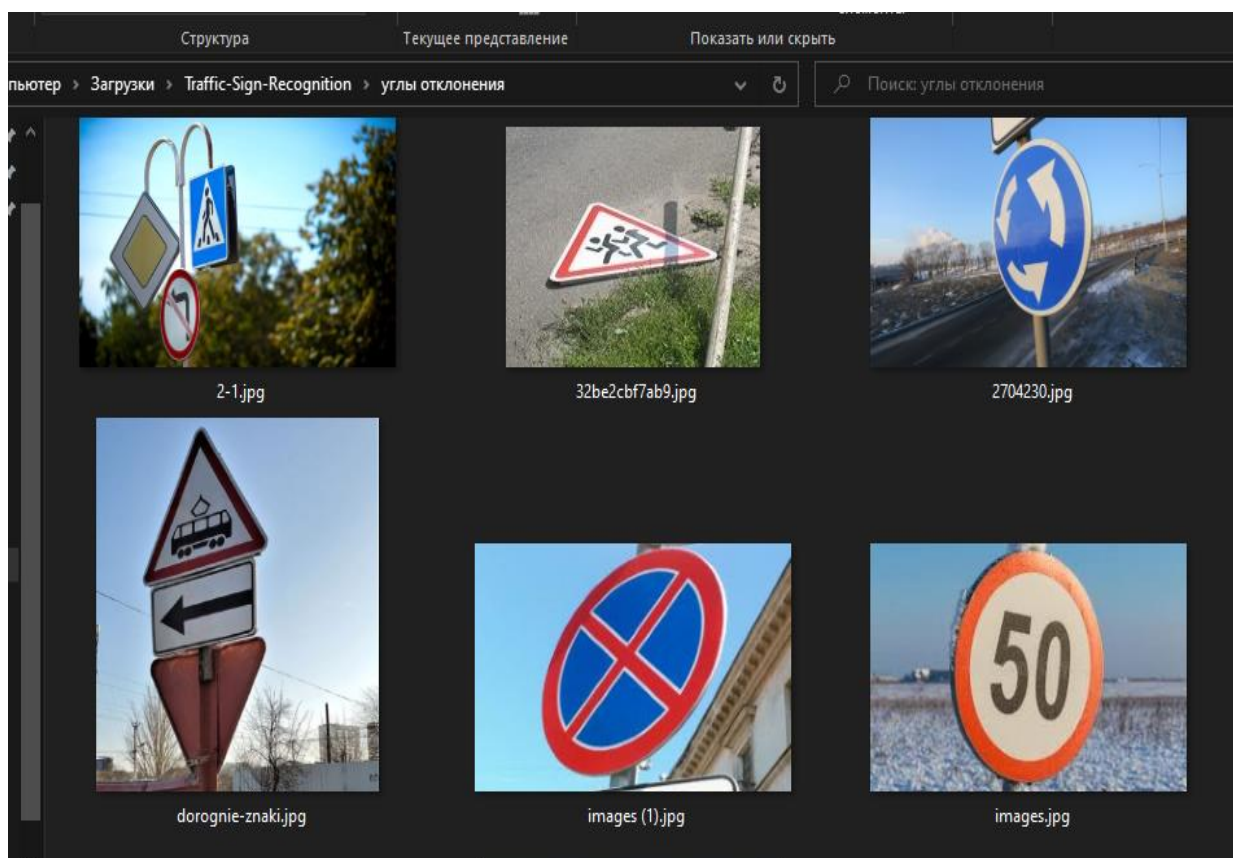


Рисунок 3.10 – Тестова вибірка зображень з відхиленням

На рисунках 3.11, 3.12, 3.13 зображені дорожні знаки під значним кутом нахилу, система без проблем розпізнала їх, це результат роботи методу SVM – опорних векторів.



Рисунок 3.11 – Розпізнавання дорожнього знака під кутом нахилу

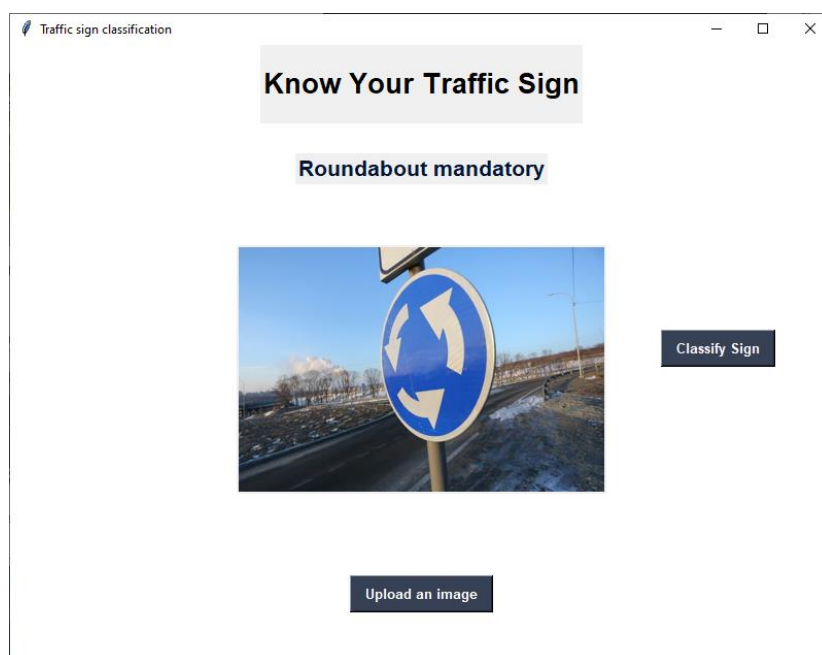


Рисунок 3.12 – Розпізнавання дорожнього знака під кутом нахилу



Рисунок 3.13 – Розпізнавання дорожнього знака під кутом нахилу

Результати виявилися наступними з 40 зображень 4 зображення не розпізнались, що дає точність розпізнавання в 90%. З цього випливає, що алгоритм і метод розпізнавання стійкий до зміни кута нахилу.

#### 3.4.3 Порівняння технічних характеристик розроблених алгоритмів і програмних засобів з існуючими

Для порівняння створеного програмного засобу будуть запропоновані програмні рішення, які були описані раніше в першому розділі:

- Opel Eye (opel);
- Speed limit assists (Mercedes);
- Road sign information;
- система розпізнавання дорожніх знаків лабораторії комп'ютерної графіки та мультимедіа МДУ ім. Ломоносова.

Дані для порівняння наведені в таблиці 3.2.

Таблиця 3.2 – Порівняння створеного ПО із схожими рішеннями

Параметр	Розроблена система	Opel Eye	Speed limit assists	Road sign information	Лабораторія комп. графіки та мультимедіа ім. Ломоносова
Точність розпізнавання	92,83	75	70	75	95
Робота в реальному часі	-	+	+	+	+
Відсоток точності з різними кутами відхилення	92,5	67	63	65	95
Метод розпізнавання	Метод опорних векторів та згортова нейронна мережа	Невідомо	Невідомо	Невідомо	Згортова нейронна мережа, що базується на алгоритмі Віюлі-Джонса

На рисунку 3.14 представлено порівняння точності розпізнавання програмних рішень.

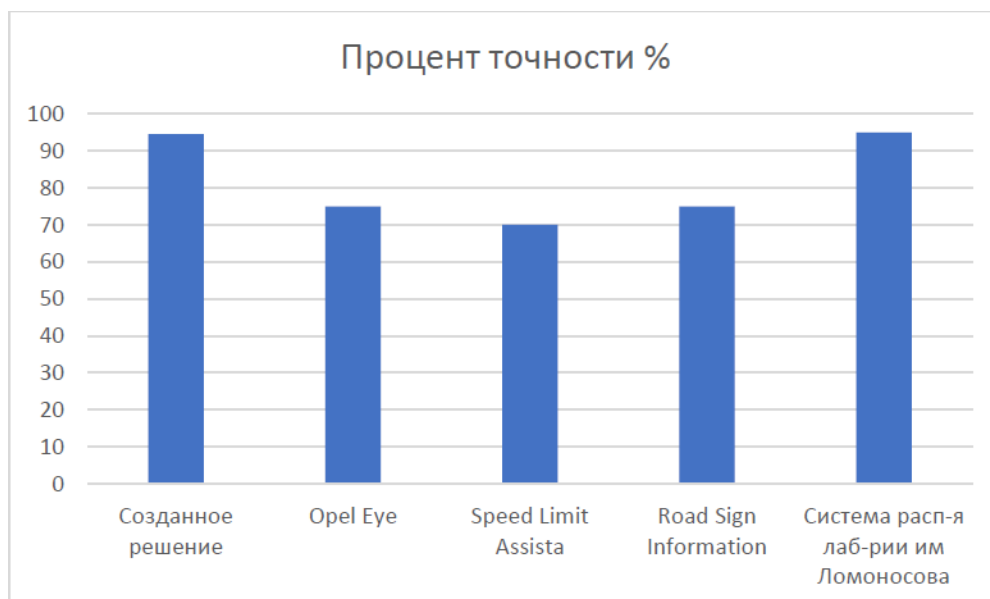


Рисунок 3.14 – Порівняння точності розпізнавання

Створений додаток показав гарний результат в розпізнаванні дорожніх знаків, завдяки комбінованому методу ідентифікації, що є поєднанням методу опорних векторів і згорткової нейронної мережі, вдалося домогтися високого відсотка розпізнавання навіть при різних кутах нахилу, цьому посприяла згорткова нейронна мережа, яка є стійкою до різних видів спотворень. Незважаючи на це, система дає і помилкові результати на зображеннях з об'єктами схожими за кольором і формою з дорожнім знаком, але відсоток помилкових розпізнавань невеликий, що можна вважати прийнятним для нашого застосування. Подальший напрям роботи бачиться в поліпшенні швидкості розпізнавання, доопрацювання алгоритму для зниження відсотку помилкових розпізнавань і додавання функції ідентифікації в реальному часі за допомогою відеокамери.

## ВИСНОВКИ

В ході виконання кваліфікаційної роботи були отримані наступні наукові і практичні результати:

– в роботі проведено повний аналіз існуючих методів і алгоритмів для ідентифікації об'єктів на зображенні складним фоном. Були описані характеристики предметної області (дорожні знаки) дослідження. Подано порівняльну характеристику існуючих рішень в області розпізнавання дорожніх знаків на зображенні. Обрані найбільш підходящі методи для реалізації власного алгоритму;

– був запропонований власний метод для розпізнавання зображень, цей метод представляє з себе синтез двох існуючих методів, такі як метод опорних векторів і згорткові нейронні мережі. Даний метод працює за принципом подвійного розпізнавання, спочатку зображення проходить перший етап ідентифікації за допомогою методу опорних векторів, який відшукує на зображеннях дорожній знак по його формі і, якщо він там знаходиться передає його згортковій нейронній мережі, яка вже приймає остаточне рішення про присутність шуканого об'єкта на зображенні. Завдяки цьому методу можна поліпшити швидкодію системи, швидкість її навчання, а також свідчення точності при різних спотвореннях, шуми і зміни кутів нахилу;

– було створено програмне рішення на основі методів і алгоритмів, запропонованих у другому розділі кваліфікаційної роботи. Програмної областю для реалізації було прийнято рішення вибрати Microsoft Visual Studio Code, відкриту бібліотеку комп'ютерного зору і алгоритмів OpenCV та Keras для роботи із згортковою нейронною мережею. Розроблена програма на основі нового алгоритму показала відмінні результати у всіх основних вимогах для системи розпізнавання, таких як розпізнавання об'єкта на помилкових зображеннях, ідентифікація

під різними кутами нахилу об'єкта, розпізнавання на великих відстанях (від 50 і до 100 м), а також захоплення і розпізнавання відразу декількох об'єктів в кадрі. Проведена порівняльна характеристика програми з існуючими аналогами.

Плюси розробленого програмного рішення в щодо високій швидкості навчання мережі, великий відсоток точних розпізнавань при різних кутах нахилу, також до плюсів можна віднести здатність розпізнавання відразу декількох дорожніх знаків. Мінусом програми є відсутність функції розпізнавання в реальному часі. Надалі планується вести роботу в поліпшенні алгоритму розпізнавання, а також додавання в програму функції ідентифікації в реальному часі. Результати експериментальних досліджень, наведені в роботі, становлять практичний інтерес при проектуванні систем управління та обробки інформації в області комп'ютерного зору і розпізнавання зображень.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Каковкин П.А. Применение алгоритмов глубокого обучения для локализации и распознавания дорожных знаков на изображениях. *Высокие технологии в современной науке и технике: сборник научных трудов IV Международной научно-технической конференции молодых ученых, аспирантов и студентов*. Томск: ТПУ, 2015. С. 360–364.
2. Вапник В.Н. Теория распознавания образов (статистические проблемы обучения). Москва: Наука, 1974. 416 с.
3. Якимов П.Ю. Отслеживание дорожных знаков в видеопоследовательности с использованием скорости. *Компьютерная оптика*. 2015. № 5. С. 795–800 .
4. Айзерман М.А. Метод потенциальных функций в теории обучения. Москва: Наука, 1970. 384 с.
5. Ciresan U. Meier, J. Masci, L. M. Gambardella, J. Schmidhuber Flexible, High Performance Convolutional Neural Networks for Image Classification. *Proceedings of the Twenty-Second international joint conference on Artificial Intelligence*. 2011. P. 1237–1242.
6. Крючин О.В. Параллельные алгоритмы обучения искусственной нейронной сети Quickprop и Rprop. *Вестник Тамбовского университета*. 2012. № 1. С. 34–41.
7. Митчелл Б.Ф. Нахождение ближайшей к началу координат точки *Вестник ЛГУ*. 1971. № 19. С. 38–45.
8. Солдатова О.П. Применение сверточной нейронной сети для распознавания рукописных цифр. *Компьютерная оптика*. 2010. № 2. С. 252–259.
9. Open CV Library. URL: <http://opencv.org/downloads.html> (дата звернення: 31.03.2021)

10. Чернухин Ю.В. Цифровая нейροкомпьютерная модель зрительного анализатора. *Многопроцессорные вычислительные структуры*. 2011. № 13. 49 с.
11. Королёв Б.Е. Разработка информационной системы распознавания дорожных знаков. URL: <http://epostulat.ru/index.php/Postulat/article/view/644/66> (дата звернення: 31.03.2021)
12. Шапошников Д.Г. Применение поведенческой модели зрения для инвариантного распознавания лиц и дорожных. *Нейрокомпьютеры: разработка и применение*. 2002. №7. С. 21–33.
13. Друки А.А. Система поиска, выделения и распознавания лиц на изображениях. *Известия Томского политехнического университета*. 2011. № 5 С. 64.
14. Нейрокомпьютеры архитектура и реализация. URL: <http://www.chipinfo.m/literature/chipnews/200101/24.html> (дата звернення: 01.04.2021)
15. Хайкин С. Нейронные сети: полный курс. 2-е издание. Москва: Вильямс, 2006. 1103 с.
16. Viola P. Robust real-time face detection. *International Journal of Computer Vision*. 2004. Vol. 57, No. 2. P.137–145.
17. Журавлев Ю.И. Распознавание. Математические методы. Программная система. Практические применения. Москва: ФАЗИС, 2006. 176 с.
18. Алпатов Б.А. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. Москва: Радиотехника, 2010. 176 с.
19. Houben S. Detection of Traffic Signs in Real-World Images: The German Traffic Sign Detection Benchmark. *International Joint Conference on Neural Networks*. 2013. С. 58–60.

20. Вежнев А. Выделение связных областей в цветных и полутоновых изображениях. *Компьютерная графика и мультимедиа*. 2003. № 1. С. 56.

21. Макаренко А.А. Методика локализации изображения лица для систем видеоконтроля на основе нейронной сети. *Известия Томского политехнического университета*. 2010. № 8. С. 113–118.

22. Садыков С.С. Скелетизация бинарных изображений. *Зарубежная радиоэлектроника*. 1985. № 11. С. 30–37.

23. Jongmin Y., Bongnam K., Kim D., Car License Plate Detection under Large Variations Using Covariance and HOG Descriptors. 8th International Symposium on Visual Computing (ISVC). 2012. P. 636–647.

24. Машина опорных векторов. URL: <http://www.intuit.ru/studies/courses/10621/1105/lecture/17990>. (дата звернения: 01.04.2021)

25. Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image: patent No. US6711293B1 Appl. No. 09/519,893; Filed 06.03.2000; publ. 23.03.2004, 20 p.

26. Dalal N., Triggs B., Histograms of oriented gradients for human detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Montbonnot, France 2005. P. 886-893

27. Bay H., Ess A., Tuytelaars T., Luc Van Gool, Speeded Up Robust Features, *Computer Vision and Image Understanding*, 2008, doi: 10.1016/J.CVIU.2007.09.014

28. Schapire R. E., Singer Y., Improved Boosting Algorithms Using Confidence-rated Predictions. *Machine Learning*. Netherlands: Kluwer Academic Publishers, 1999, P. 297–336.