

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМ. В. ДАЛЯ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ЕЛЕКТРОНІКИ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК ТА ІНЖЕНЕРІЇ

До захисту допускається
Завідувач кафедри
_____ Скарга-Бандурова І.С.
« ____ » _____ 20__ р.

ДИПЛОМНИЙ ПРОЕКТ (РОБОТА) БАКАЛАВРА
ПОЯСНЮВАЛЬНА ЗАПИСКА

НА ТЕМУ:

Інформаційна система виявлення та розпізнавання тексту з відеопотоку

Освітньо-кваліфікаційний рівень “бакалавр”
Напрямок підготовки 6.050102 – “комп’ютерна інженерія”

Керівник проекту:

(підпис)

Барбарук Л.В.

(ініціали, прізвище)

Консультант з охорони праці:

(підпис)

Критська Я.О.

(ініціали, прізвище)

Здобувач вищої освіти:

(підпис)

Грошкова О.О.

(ініціали, прізвище)

Група:

КІ-14з

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВОЛОДИМИРА ДАЛЯ

Факультет Інформаційних технологій та електроніки
Кафедра Комп'ютерних наук та інженерії
Освітньо-кваліфікаційний рівень бакалавр
Напрямок підготовки 6.050102 – “комп'ютерна інженерія”
(шифр і назва)
Спеціальність _____
(шифр і назва)

ЗАТВЕРДЖУЮ:

Завідувач кафедри КНІ
_____ І.С. Скарга-Бандурова
« _____ » _____ 20__ р.

**З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЕКТ (РОБОТУ) БАКАЛАВРА**

Грошківій Олени Олександрівни
(прізвище, ім'я, по батькові)

1. Тема роботи Інформаційна система виявлення та розпізнавання тексту з відеопотоку

керівник проекту (роботи) Барбарук Ліна Вікторівна
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом вищого навчального закладу від " _____ " _____ 201_ р. № _____

2. Термін подання студентом роботи 16.06.2018

3. Вихідні дані до роботи текстові зображення у форматі RGB

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) Аналіз предметної області та постановка задачі, виявлення текстових областей і виділення символів на зображенні, розпізнавання символів за допомогою нейронних мереж, програмна реалізація алгоритмів, охорона праці

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень) Електронні плакати

6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	ст. викл. Критська Я.О.		

7. Дата видачі завдання 30.04.2018

Керівник

_____ (підпис)

Завдання прийняв до виконання

_____ (підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Аналіз завдання та робота з літературою	05.05.2018 - 13.05.2018	
2	Розробка алгоритмів	14.05.2018 - 22.05.2018	
3	Реалізація алгоритмів	22.05.2018 - 02.06.2018	
4	Розробка розділу «Охорона праці»	02.06 .2018- 11.06.2018	
5	Оформлення пояснювальної записки та електронних плакатів	11.06.2018 - 16.06.2018	

Здобувач вищої освіти

_____ (підпис)

Грошкова О.О.

_____ (прізвище та ініціали)

Керівник

_____ (підпис)

Барбарук Л.В.

_____ (прізвище та ініціали)

РЕФЕРАТ

Пояснювальна записка дипломної роботи бакалавра: 88 с., 26 рис., 5 табл., 2 додатки, 30 джерел посилань.

Робота присвячена розробці алгоритмів і програми хмарного сервісу для розпізнавання тексту в потоці відеоданих. В основі алгоритму лежить розпізнавання зображень за допомогою нейронних мереж з вчителем. Був проведений аналіз фільтрів зображення, в результаті чого, для виділення текстової інформації на зображенні, в розробленому алгоритмі використовується фільтр Гауса. Реалізований алгоритм виявлення символів на зображенні.

Ключові слова: детектор Канні, оператор Собеля, фільтр Гауса, перцептор, відеопотік, розпізнавання образів, нейронна мережа.

Умови одержання дипломного проекту: СНУ ім. В. Даля, пр. Центральний 59-А, м. Сєвєродонецьк, 93400.

ЗМІСТ

ВСТУП.....	6
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ.....	8
1.1 Системи аналізу відеопотоку	8
1.2 Методи розпізнавання тексту	9
1.3 Постановка задачі	13
2 ВИЯВЛЕННЯ ТЕКСТОВИХ ОБЛАСТЕЙ І ВИДІЛЕННЯ СИМВОЛІВ НА ЗОБРАЖЕННЯХ.....	15
2.1 Перетворення зображення в градацію сірого.....	16
2.2 Згладжування фільтром Гауса	17
2.3 Виявлення контурів Канні.....	18
2.4 Пошук градієнтів	20
2.5 SWT-обробка зображення	22
2.6 Фільтрація отриманого SWT-зображення.....	23
2.8 Групування літер на слова та рядки	24
3 РОЗПІЗНАВАННЯ СИМВОЛІВ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ.....	26
3.1 Розпізнавання образів інтелектуальними системами	26
3.1.1 Огляд задачі розпізнавання образів інтелектуальними системами	26
3.1.2 Задача моделювання нейронних мереж для класифікації зображень символів.....	27
3.1.3 Навчання перцептрона.....	29
3.1.4 Дельта-правило	31
3.2 Термін образу. Проблема навчання розпізнавання образів	33
3.3 Нейронні мережі	35
3.3.1 Біологічний нейрон.....	35
3.3.2 Штучний нейрон	36
3.3.3 Формальний нейрон.....	38

3.3.4 Одношарові штучні нейронні мережі	43
3.3.5 Багатошарові штучні нейронні мережі	46
4 ПРОГРАМНА РЕАЛІЗАЦІЯ Й ДОСЛІДЖЕННЯ РОЗРОБЛЕНИХ АЛГОРИТМІВ.....	48
4.1 Архітектура хмарного сервісу та клієнтської програми для розпізнавання тексту в відео потоці.....	48
4.2 Реалізація алгоритму виявлення та розпізнавання тексту.....	49
4.3 Реалізація алгоритму розпізнавання тексту	53
4.4 Аналіз роботи алгоритму виявлення тексту.....	54
5 ОХОРОНА ПРАЦІ	57
5.1 Аналіз потенційних небезпечних і шкідливих виробничих чинників проєктованого об'єкту, що мають вплив на персонал	57
5.2 Заходи щодо техніки безпеки	59
5.3 Заходи, що забезпечують виробничу санітарію і гігієну праці....	61
5.4 Рекомендації по пожежній безпеці.....	66
ВИСНОВКИ	70
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ.....	72
ДОДАТОК А. Вихідний код класу, що описує нейрон.....	75
ДОДАТОК Б. Електронні плакати	82

ВСТУП

На сьогоднішній день існує багато алгоритмів щодо вирішення проблеми розпізнавання символів, але більшість із них або вузько спрямовані на певну сферу розпізнавання, або якість розпізнавання дуже низька, й алгоритм працює повільно. Ефективний розв'язок завдань розпізнавання зображень відкриває широкі можливості в розв'язку завдань робототехніки, побудові систем керування для мобільних роботів і створенні штучного зору. Серед методів, що застосовуються в розв'язку завдань розпізнавання зображень є багато різних методів, які використовують різні математичні підходи. Тому для вирішення задач розпізнавання символів часто використовуються високоінтелектуальні системи на основі штучних нейронних мереж [1].

Проблема розпізнавання тексту на сьогоднішній день вважається надзвичайно серйозною і відіграє надзвичайно велику роль в аналізі графічної інформації. Це зумовлено як складністю опису об'єкту розпізнавання, його виділення з фону на дискретному полі уваги, так і потребою застосування (аерокосмічна галузь, системи безпеки, автоматизовані системи розпізнавання для прийняття рішень).

Завдяки сучасному розвитку комп'ютерних технологій та інформаційних систем, задача обробки відеозображення, що надходить в режимі реального часу постає досить часто, причому така необхідність виникає не тільки в деяких галузях науки й техніки, а й на конвеєрному виробництві, різноманітних підприємствах, банках, супермаркетах, охоронних й оборонних організацій, адже саме фіксування неправомірних дій людини та миттєва реакція на них є головною метою відеозображення.

Найбільш очевидним прикладом проблеми, що не може бути вирішена людиною через велику кількість даних, є проблема аналізу відеоданих. З моменту появи відеозображення його застосування полягало в нагромадженні

відеоданих для наступної обробки людиною у випадку потреби. Згодом системи відеозображення стали здатними виконувати аналіз зображень самостійно та розпізнавати на них об'єкти. Але залишається низка проблем, які перешкоджають можливості комп'ютера обробляти зображення. До цих проблем відносяться: освітленість, погодні умови, електромагнітний вплив на апаратуру, що реєструє дані та на передавальну апаратуру, і т.п. фактори, що сильно ускладнюють завдання обробки.

В ході виконання дипломної роботи був розроблений алгоритм та його програмна реалізація для виявлення та розпізнавання тексту у потоці відеоданих з використанням нейронних мереж. Розроблений алгоритм також може бути використаний в будь-якій системі, що потребує пошуку текстових даних, наприклад: системи виміру швидкості руху автомобілів, для автоматичного пошуку номерних знаків; системи моніторингу реклам та телепередач.

Більшість додатків, що розпізнають текст, розроблені для встановлення на локальному комп'ютері користувача, і, як наслідок, їхня продуктивність залежить від характеристик комп'ютера. Рішенням даної проблеми є хмарний сервіс. Користувач у такому випадку не прив'язаний до певного робочого місця й може підключитися до сервісу з будь-якої точки світу, а сервіс у свою чергу не залежить від програмного забезпечення й операційної системи на стороні користувача. У рамках даної роботи був реалізований хмарний сервіс, що буде надавати функціональність по розпізнаванню тексту на зображеннях.

У розділі «Охорона праці», вирішуються питання, пов'язані із забезпеченням здорових і безпечних умов праці на робочих місцях, захисту людини і навколишнього середовища, від дії шкідливих факторів, що перевищують нормативно допустимі рівні підтримки оптимальних умов діяльності і відпочинку, що створить передумови для вищої працездатності.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Системи аналізу відеопотоку

Обробка відеопослідовностей є дуже важливою на сьогоднішній день. Її застосовують починаючи від обробки відео на камері мобільного пристрою, передачі відео по телебаченню і закінчуючи різними складними системами відслідковування об'єктів в відеопотоці [2].

Однією із складнощів обробки відео є обробка відео в реальному часі, що зумовлює проводити обробку 25 і більше кадрів в секунду. Ці труднощі в першу чергу пов'язані з тим, що дуже швидко потрібно проводити багато математичних операцій над великими обсягами даних (для прикладу на відеопотоці з розміром кадру 640x480, одним каналом та 8-мибітним представленням, щоразу потрібно обробляти 307200 байт даних). Тому завжди проводять оптимізацію обробки даних як на програмному рівні (удосконалюються алгоритми) так і на апаратному (впроваджуються нові архітектури процесорів, що оптимізовані для обробки векторів та матриць).

Введемо ряд визначень, які будуть необхідні впродовж всієї роботи.

Відеокадр – растрове зображення у форматі RGB.

Потік відеокадрів – послідовність відеокадрів, отримана від цифрового пристрою реєстрації відео зображень (цифрової відеокамери) або синтезована за допомогою спеціального ПЗ (аніматора).

Під відеоданими звичайно розуміється потік даних, що мають структуру послідовних кадрів відеозображення. Ці кадри можуть бути отримані від пристрою реєстрації (цифрової / аналогової камери) або синтезовані за допомогою спеціалізованого програмного забезпечення, а потім неодноразово перетворені для зручності їх зберігання та передачі. Всі кадри подібного потоку мають один і той же розмір зображення і зазвичай надходять з певним постійним часовим інтервалом. Величина, зворотна до часового інтервалу між сусідніми кадрами – це кількість кадрів, що

необхідно опрацювати за одну секунду . Крім того, всі кадри потоку мають однакову глибину кольору, що вимірюється в кількості використовуваних бітів пам'яті для зберігання одного пікселя кадру.

Зауважимо, що будемо мати справу виключно з цифровими відеоданими, тобто, з тими, які зберігаються в пам'яті ЕОМ і обробляються на ній програмно або апаратно.

Реєстрація (відео захват) зазвичай відбувається апаратно за допомогою відеокамери, решту етапів частіше реалізуються програмно. Кодування / декодування виконується різними кодеками, що реалізують алгоритми стиснення растрових даних без втрат або стиснення з втратами. Решта перетворення можуть включати такі операції, як обробка фільтрами згладжування, виділення країв, зміни контрастності, яскравості, колірного простору.

Перетворювачі, що виконують подібні дії будемо називати фільтрами, оскільки ділянки їх програмного коду, реалізують не що інше, як алгоритми фільтрації растрів.

Робота із зображеннями у відтінках сірого (8 bpp, grayscale) є окремим випадком роботи з форматом RGB24, оскільки повністю аналогічна обробці будь-який однією з компонентів 24-бітного кольору.

1.2 Методи розпізнавання тексту

Розпізнавання тексту складається з наступних процедур і методів [3]:

- попередня обробка зображення;
- фільтрація зображення;
- розпізнавання тексту.

Процедура попередньої обробки використовується практично завжди після отримання інформації, і являє собою застосування операцій

усереднення і вирівнювання гістограм, різного типу фільтрів для виключення перешкод, а також придушення зовнішніх шумів. Одні з основних дій, які проводяться на етапі попередньої обробки – це зміна контрастності, яскравості та згладжування зображення.

Найбільш важливим етапом являється етап фільтрації зображення, важливість даного етапу полягає у знаходженні об'єктів, що схожі на літери.

Кінцевий етап обробки – розпізнавання текстової інформації. Для цього етапу вхідними даними є зображення, отримані в результаті SWT-перетворення зображення.

Про те що існують спеціальні системи, які «автоматично вводять текст в комп'ютер», знають навіть ті користувачі, які тільки почали знайомитись з персональним комп'ютером. На зображенні, яке було відскановано система знаходить ті фрагменти, в яких «впізнає» літери, а потім замінює ці зображення справжніми літерами, або їх ще називають машинними кодами. Так здійснюється перехід від зображення тексту до «справжнього» тексту, з яким можна працювати в текстовому редакторі.

На даний момент відомо три основні підходи розпізнавання символів – шаблонний, структурний та ознаковий [4]. Кожен з цих методів орієнтований на свої умови застосування, для яких вони є ефективними. Разом з тим всім цим методам властиві недоліки. Найбільш істотні з них – низька стійкість до змін масштабу, зсувів, поворотам, зміні ракурсу та іншим спотворень. Ці недоліки особливо яскраво проявлялися при масштабній експлуатації програмно-технологічних систем, що використовують в своїй основі ці методи.

Шаблонний опис простіше в реалізації, однак, на відміну від структурного, він не дозволяє описувати складні об'єкти з великою різноманітністю форм. Саме тому шаблонний опис застосовується для розпізнавання лише друкованих символів, у той час як структурний – для рукописних, що мають, звичайно, набагато більше варіантів накреслення. Принципу цілісності відповідає лише перші два.

А тепер розглянемо більш детально кожен метод окремо.

Шаблонні методи розпізнавання символів. Першим етапом роботи шаблонного методу є перетворення відсканованого зображення в растрове. Далі проводиться його порівняння з усіма наявними в базі системи шаблонами. Найбільш підходящим шаблоном вважається той, у якого буде найменша кількість точок, відмінних від досліджуваного зображення. Шаблон для кожного класу зазвичай отримують, усереднюючи зображення символів навчальної вибірки. У цих методів досить висока точність розпізнавання дефектних символів (склеєних або розірваних). Недолік даного методу - неможливість розпізнати шрифт, який хоч трохи відрізняється від закладеного в систему (розміром, нахилом або накресленням), шаблонні методи можуть робити помилки навіть при обробці дуже якісних зображень.

Алгоритм, заснований на шаблонному методі, повинен заздалегідь знати шрифт, який йому представляють для розпізнавання. При існуючому багатстві друкованої продукції в процесі навчання неможливо охопити всі шрифти та їх модифікації. Іншими словами, цей фактор обмежує універсальність таких методів.

Мірою подібності в даному методі обраний коефіцієнт подібності зображення символу з узагальненим образом S -го класу, виражається наступною формулою:

$$R_S = \sum_{j=1}^n \left(\ln \frac{P_{js}}{1-P_{js}} \right) x_j + \sum_{j=1}^n (1-P_{js}),$$

де R_S – коефіцієнт подібності розпізнавального символу до еталонного зображення S -го класу символів; R_{js} – ймовірність появи чорного кольору в j -му елементі еталонного зображення S -го класу; x_j – значення інтенсивності, відповідному j -му елементу розпізнавального символу.

Зображення символу ототожнюється із еталонним класом, який дав максимальний коефіцієнт подібності R серед всіх R_S .

Структурні або топологічні методи розпізнавання зберігають інформацію не про растрове написання символу, а про його топологію. Іншими словами, еталон містить інформацію про взаємне розташування окремих складових частин символу. Перевага методу стійкість до зсуву і повороту символу на невеликий кут, до різних стильових варіацій шрифтів. Однак, при повороті на кут, більший десяти градусів, даний метод не може бути використаний для розпізнавання символів. Ясно, що при цьому стає неважливим розмір літери, що розпізнається і навіть шрифт, яким вона надрукована. Але основною проблемою структурних методів розпізнавання залишається ідентифікація знаків, що мають певні дефекти (наприклад розрив ліній або злиття сусідніх ліній). Крім того, для цих методів, на відміну від шаблонних та ознакових, досі не створені ефективні автоматизовані процедури навчання. Тому структурний опис найчастіше доводиться створювати вручну.

Ознакові методи базуються на тому, що зображенню ставиться у відповідність N -мірний вектор ознак. Розпізнавання полягає в порівнянні його з набором еталонних векторів тієї ж розмірності. Завдання розпізнавання, прийняття рішення про приналежність образу того чи іншого класу, на підставі аналізу обчислених ознак, має цілий ряд математичних рішень в рамках ймовірнісного підходу. У системах розпізнавання символів найчастіше використовується класифікація, заснована на підрахунку евклідової відстані між вектором ознак розпізнавального символу і векторами ознак еталонного опису. Тип і кількість ознак в чималому ступені визначають якість розпізнавання. Формування вектора проводиться під час аналізу попередньо підготовленого зображення. Даний процес називають витяганням ознак. Еталон для кожного класу отримують шляхом аналогічної обробки символів навчальної вибірки.

Основні переваги ознакових методів – простота в реалізації, хороша стійкість до зміни форми символів, хороша узагальнююча здатність, висока швидкість розпізнавання. Найбільш серйозний недолік цих методів нестійкість до різних дефектів зображення. Крім того, ознакові методи володіють іншим серйозним недоліком на етапі вилучення ознак відбувається незворотня втрата значної частини інформації про символ. Вилучення ознак ведеться незалежно, тому інформація про взаємне розташування елементів символу втрачається. Ознакові системи не відповідають принципу цілісності.

Для вирішення поставленої задачі ефективно використовувати шаблонну систему розпізнавання символів за допомогою нейронних мереж, у зв'язку з тим, що вони є слабо чутливими до спотворень вхідного сигналу, так само забезпечують можливість отримання класифікатора, добре моделює складну функцію розподілу зображень символів, містять у собі велику стійкість до зміни ознакового простору в межах знімка, меншу вимогливість до стандартизації зображення (куту сонця, атмосферному перекручуванню), легку відтворюваність, можливість навчання у випадку чи зміни ускладнення задач, простоту при подальшому використанні як стандарти виділення тих чи інших об'єктів, тим самим збільшуючи точність рішення в порівнянні з іншими методами.

1.3 Постановка задачі

Метою дипломної роботи є розробка алгоритму розпізнавання тексту в потоці відеоданих.

Для працездатності алгоритму використовують наступні вимоги:

- можливість роботи в реальному часі;
- можливість виявлення та розпізнавання текстової інформації в

умовах дії завад;

– алгоритм повинен заздалегідь знати шрифт, що йому представляють для розпізнавання друкованого тексту, тобто він повинен зберігати в базі різні характеристики шрифту та стилю написання.

Для досягнення цієї мети необхідно:

– підготувати набір тестових зображень та реалізувати програму для генерації еталонних зображень для різних варіацій шрифтів;

– реалізувати алгоритм виявлення тексту на зображенні;

– реалізувати алгоритм розпізнавання тексту шляхом використання нейронної мережі;

– програмно реалізувати хмарний сервіс.

2 ВИЯВЛЕННЯ ТЕКСТОВИХ ОБЛАСТЕЙ І ВИДІЛЕННЯ СИМВОЛІВ НА ЗОБРАЖЕННЯХ

З кожним днем зростає кількість текстової інформації, яку сприймає і аналізує людина. Оголошення, вивіски, написи, візитні картки, номерні знаки транспортних коштів лише короткий перелік форм подання такої інформації. З ростом кількості інформації назріває питання – як автоматизувати і поліпшити процес сприйняття інформації людиною за допомогою інформаційних технологій? В таких умовах досить актуальною є розробка алгоритмів обробки візуальної інформації з подальшим використанням в інформаційних системах різного призначення. Одним із видів таких систем є системи автоматичного розпізнавання тексту.

Базовою процедурою, що застосовується на початковому етапі обробки в таких системах, є виявлення текстових областей [5]. Існує безліч алгоритмів і методів для її виконання, проте до сьогоднішнього часу немає остаточного вирішення цього завдання.

Існуючі методи виявлення текстових областей на зображеннях можна умовно розділити на три категорії: текстурні, методи на основі областей і гібридні. Текстурні розглядають текст як особливий тип текстури і використовують для виявлення такі його характеристики як локальна інтенсивність або вейвлет-коефіцієнти. Ці методи в основному визначають текстові області, рядки і стовпці яких паралельні відповідним координатним осям. Методи на основі областей спочатку знаходять кандидатів у текстові області, використовуючи виділення контурів або кластеризацію, а потім проводять фільтрацію областей, використовуючи евристичні правила для відсіювання зайвих. Третя категорія, гібридні методи, являють собою суміш методів перших двох категорій.

Перераховані методи дають хороші результати лише при експериментальній оцінці багатьох параметрів (наприклад, співвідношення

інтенсивності тексту і фону, розмір шрифту). Це негативно позначається на швидкодії самих методів або звужує область їх застосування.

У даній роботі при аналізі текстових областей пропонується використовувати таку характеристику як товщина штриха символу тексту. Вона ґрунтується на апріорних знаннях про те, що в межах однієї текстової області (написи, рядки тексту) товщина штриха символу тексту і співвідношення висоти/ширини символу залишаються приблизно незмінними.

2.1 Перетворення зображення в градацію сірого

Градація сірого – режим кодування кольорів у зображенні, що містить тільки білий, чорний і відтінки сірого кольору.

Перетворення кольорового зображення у відтінки сірого здійснюється наступним чином. Вся кольорова палітра представляється у вигляді куба, вершини якого відповідають різним кольорам. Сіра шкала розташовується на діагоналі куба, що з'єднує чорну і білу вершини [6].

Для перетворення зображення у відтінки сірого для кожної точки зображення виділяються інтенсивності червоною, зеленою і синьою складовою кольору, а потім здійснюється перетворення кольору за наступною формулою:

$$Y = 0.255R + 0.255G + 0.255B,$$

де Y – нове значення кольору, R – інтенсивність червоної складової кольору, G – інтенсивність зеленої складової кольору, B – інтенсивність синьої складової кольору.

2.2 Згладжування фільтром Гауса

Існує досить багато методів для згладжування зображень, які відрізняються своїми параметрами і вихідними результатами. це такі фільтри: лінійні, Гауса, нелінійні, медіанний, адаптовані, комбіновані, гібридні [7]. В роботі для зменшення чутливості алгоритму до шуму виконується згладжування зображення за допомогою фільтру Гауса. Завдяки даному фільтру можна зменшити відстань впливу пікселей один на одного.

При згладжуванні зображень використовують апертуру (радіус) – це прямокутна або квадратна ділянка зображення, на якому визначається функція. Задана функція в апертурі називається ваговою, або функцією вікна, а сама апертура разом з цією функцією називається маскою зображення.

В обробці сигналів, фільтром Гауса називають фільтр, чия імпульсна характеристика є функцією Гауса. Імпульсна характеристика – вихідний сигнал динамічної системи як реакція на вхідний сигнал, тобто наші дані (імпульс). Гаусів фільтр спроектований таким чином, щоб звести до мінімуму відхилення від вхідних даних перехідної функції, реакцію системи на вхідний поетапне вплив при нульових початкових умовах під час наростання і спаду.

Така поведінка пов'язана з тим, що фільтр Гауса має мінімальну групову затримку, міру проходження даних через ядро фільтра. Фільтр Гауса зазвичай використовується в цифровому вигляді для обробки двовимірних сигналів з метою зниження рівня шуму. Варто відзначити досить обмежену швидкість фільтра Гауса при реалізації за допомогою явного методу, особливо помітну на великих обсягах даних.

Імпульсна характеристика одновимірного фільтра Гауса може бути представлена у вигляді:

$$G(x) = \sqrt{\frac{a}{n}} e^{-ax^2}.$$

А з середньоквадратичним відхиленням:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}.$$

Для двовимірного випадку ми представляємо фільтр як добуток двох одновимірних випадків, тому отримаємо:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}},$$

де x, y – координати пікселя зображення, σ – середньоквадратичне відхилення розподілу Гауса.

Після застосування фільтра, зображення стає злегка розмитим, як при спостереженні через каламутне скло.

2.3 Виявлення контурів Канні

Актуальною проблемою при обробці зображень залишається вибір адекватного поставленій задачі набору характерних ознак, які використовуватимуться для аналізу і розпізнавання зображень. В багатьох технічних задачах необхідно розпізнавати або об'єкт, або текстуру зображення, або об'єкт на фоні текстури.

На сьогодні широке застосування для виділення контуру зображення має оптимальний детектор Канні.

Детектор Канні орієнтується на три основні критерії:

- підвищення рівня сигналу / шуму;

- хороша локалізація (правильне визначення положення межі);
- єдиний відгук на одну межу.

Після обчислення градієнта згладженого зображення в контурі межі залишаються тільки точки максимуму градієнта зображення. Використовується інформація про напрямлення межі для того, щоб видаляти точки саме поряд з границею і не розривати саму межу поблизу локальних максимумів градієнта [8].

Потім за допомогою двох порогів видаляються слабкі межі. Фрагмент межі при цьому обробляється як ціле. Якщо значення градієнта де-небудь на простежуваному фрагменті перевищить верхній поріг, то цей фрагмент залишається також «допустимою» межею і в тих місцях, де значення градієнта падає нижче цього порога, до тих пір поки вона не стане нижче нижнього порога. Якщо ж на всьому фрагменті немає жодної точки зі значенням вище верхнього порогу, то він видаляється. Такий гістерезис дозволяє знизити число розривів у вихідних межах.

Включення в алгоритм Канні шумозаглушення з одного боку підвищує стійкість результатів, а з іншого – збільшує обчислювальні витрати та призводить до спотворення і навіть втрати подробиць меж. Так, наприклад, таким алгоритмом округляються кути об'єктів і руйнуються межі в точках з'єднань.

Перевагами детектора Канні є висока стійкість до шуму завдяки використанню методу Гауса, що дає кращу якість сегментації зображення і орієнтації меж областей, те, що він чітко виділяє контур і дозволяє виявляти внутрішні контури об'єкта. Крім того він виключає помилкове виявлення контуру там, де об'єктів немає.

Основними недоліками детекторів такого типу є:

- складність реалізації і великий час оброблення;
- можливе деяке округлення кутів об'єкта, що призводить до зміни параметрів контура.

2.4 Пошук градієнтів

Оператор Собеля використовується в обробці зображень для виділення границь. Це дискретний диференціальний оператор, що обчислює наближене значення градієнта яскравості зображення [8]. Результатом застосування оператора Собеля в кожній точці зображення є або вектор градієнта яскравості в цій точці, або його норма. Оператор Собеля базується на згортці зображення невеликими цілочисленими фільтрами в вертикальному та горизонтальному напрямках, тому його відносно легко обчислювати [2]. З іншого боку, використовувана їм апроксимація градієнта досить груба, особливо на високочастотних коливаннях зображення.

Оператор обчислює градієнт яскравості зображення в кожній точці. Так знаходиться напрямок найбільшого збільшення яскравості і величина її зміни в цьому напрямку. Результат показує, наскільки «різко» або «плавно» змінюється яскравість зображення в кожній точці, а отже, ймовірність знаходження точки на межі, а також орієнтацію границі. На практиці, обчислення величини зміни яскравості (ймовірності приналежності до грані) надійніше і простіше в інтерпретації, ніж розрахунок напрямки.

Математично градієнт функції двох змінних для кожної точки зображення (якої i є функція яскравості) – двовимірний вектор, компонентами якого є похідні яскравості зображення по горизонталі і вертикалі. У кожній точці зображення градієнтний вектор орієнтований в напрямку найбільшого збільшення яскравості, а його довжина відповідає величині зміни яскравості. Це означає, що результатом оператора Собеля в точці області постійної яскравості буде нульовий вектор, а в точці, що лежить на границі областей різної яскравості - вектор, що перетинає границю у напрямку збільшення яскравості.

Оператор використовує ядра 3×3 , з якими згортають вихідне зображення для обчислення наближених значень похідних по горизонталі і

по вертикалі. Нехай A вихідне зображення, а G_x і G_y – два зображення, де кожна точка містить наближені похідні по x і по y . Вони обчислюються наступним чином:

$$G_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & 0 & +1 \end{pmatrix} \times A,$$

$$G_x = \begin{pmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{pmatrix} \times A,$$

де \times позначає двовимірну операцію згортки.

Координата x тут зростає «направо», а y – «вниз». У кожній точці зображення наближене значення величини градієнта можна обчислити, використовуючи отримані наближені значення похідних:

$$G = \sqrt{G_x^2 + G_y^2}.$$

Використовуючи цю інформацію, ми також можемо обчислити напрямок градієнта:

$$\theta = \arctan\left(\frac{G_y}{G_x}\right),$$

де, наприклад, кут θ дорівнює нулю для вертикальної межі, у якої темна сторона зліва.

2.5 SWT-обробка зображення

Вихідне зображення, як правило, може мати ряд спотворень через наявність шуму, низького дозволу, відблисків, недостатньою або зайвою експозиції і т.д. Все це знижує ймовірність виявлення текстової області і наступного виділення символів на зображеннях.

Для виявлення текстових областей і виділення символів на зображеннях скористаємося алгоритмом SWT (Stroke Width Transform – перетворення по товщині штриха) [9]. Даний алгоритм використовує товщину штриха символу тексту, яка практично не змінюється в межах текстової області зображення і, отже, є підходящою характеристикою для виявлення текстових областей.

SWT-перетворення зображення реалізується за наступним алгоритмом:

Крок 1. Вибирається піксель контурного зображення. Його координати зберігаються в пам'яті як P_1 . У зарезервованому зображенні з розмитим контуром знаходиться піксель, відповідний цим координатам, його околиці перевіряються на найбільшу (якщо текст світліше фону) або меншу (в іншому випадку) інтенсивність. Координати цього пікселя також зберігаються в пам'ять (як P_2).

Крок 2. Проводиться промінь з пікселя з координатами P_1 в напрямку пікселя з координатами P_2 . Якщо побудований промінь перетинає будь-який необроблений піксель контуру, або досягає межі зображення – він заповнюється пікселями певної інтенсивності, що розраховується виходячи з довжини отриманого відрізка.

Крок 3. Повторюються кроки 1-2 до тих пір, поки кожен піксель нерозбавленого контурного зображення не буде належати якому-небудь побудованому променю.

В результаті даного етапу обробки отримуємо так зване SWT-

зображення.

Особливістю методу виявлення текстових областей на етапі побудови SWT-зображення є облік залежності співвідношення середньої інтенсивності кольору шрифту (I_c) і фону (I_ϕ). При значенні $(I_c)/(I_\phi) > 1$ колір шрифту вважається світліше фону, отже, промені слід будувати від контурів символу до фону. В іншому випадку напрямок градієнта змінюється у зворотний бік і промені будуються у напрямку від фону до контуру символу.

Використання дворазової SWT-обробки зображення для різного співвідношення I_c/I_ϕ з наступним об'єднанням отриманих результатів дозволяє поліпшити точність виявлення текстових областей на неоднорідному фоні.

2.6 Фільтрація отриманого SWT-зображення

Для перевірки належності контуру до символів розраховується середнє значення товщини штриха в цьому контурі [5]. Для кожного символу тексту товщина штриха повинна відрізнятися від середнього значення не більше ніж на 10%. При перевищенні даної межі символ не буде вважатися кандидатом на символ тексту, відповідна область зображення, обмежена таким контуром, виключається з подальшої обробки.

Для виділених символів перевіряється співвідношення пропорцій за формулою:

$$c = h/w,$$

де h – висота символу, w – його ширина.

При значеннях $c > 1,5$ символ виключається з подальшого процесу обробки.

2.8 Групування літер на слова та рядки

Якщо координати центрів яких-небудь кандидатів на символ знаходяться на одній прямій або прямокутнику, в які вписані виділені символи, мають спільні сторони або відрізки, то аналізовані символи є частиною послідовності символів і мають бути об'єднані.

У результаті отримуємо об'єднані в компоненти символи, готові до подальшого процесу класифікації або розпізнавання. При цьому фон видаляється з зображення, залишається тільки масив текстових областей. Зниження результатів правильного виявлення і виділення символів пов'язано з залежністю методу в цілому від параметрів детектора меж Канні і складністю отримання якісного початкового зображення. Поліпшити результат виявлення та фільтрації можна застосуванням попередньої обробки зображення, але обчислювальна складність при цьому збільшується [5].

Кількість виявлених та виділених символів залежить від особливостей алгоритму і шрифтів. Зокрема, символи шрифтів із зарубками і змінною товщиною штриха, виявлялися набагато гірше. Цей недолік можуть частково усунути морфологічні операції. Застосування операції дилатації для світлих, або ерозії – для темних шрифтів збільшувало кількість виявлених символів на 10 ... 15%.

Запропонований метод виявлення текстових областей і виділення символів з використання SWT-перетворення дозволяє виявити і виділити в середньому 85% символів на зображеннях незалежно від розміру шрифту, що доводить його працездатність в системах автоматичного розпізнавання тексту. Використання дворазової SWT-обробки для різного співвідношення

середньої інтенсивності кольору шрифту і фону дозволяє поліпшити точність виявлення текстових областей на неоднорідному фоні.

3 РОЗПІЗНАВАННЯ СИМВОЛІВ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

3.1 Розпізнавання образів інтелектуальними системами

3.1.1 Огляд задачі розпізнавання образів інтелектуальними системами

Розпізнавання образів є однією з найфундаментальніших проблем теорії інтелектуальних систем, що вивчає методи класифікації об'єктів. За традицією об'єкт, що піддається класифікації, називається образом. Образом може бути цифрова фотографія (розпізнавання зображень), буква або цифра (розпізнавання символів), запис мови (розпізнавання мови) тощо.

В межах теорії штучного інтелекту розпізнавання образів включається в більш широку наукову дисципліну – теорію машинного навчання, метою якої є розробка методів побудови алгоритмів, що здатні навчатися.

В цілому проблема розпізнавання образів складається з двох частин: навчання та розпізнавання [10]. Існує два підходи до навчання: індуктивне і дедуктивне. Індуктивне навчання, або навчання за прецедентами, засноване на виявленні загальних властивостей об'єктів на підставі неповної інформації, отриманих емпіричним шляхом. Дедуктивне навчання передбачає формалізацію знань експертів у вигляді баз знань (експертних систем тощо).

Слід зауважити, що, як кожна математична дисципліна, розпізнавання образів має власний математичний апарат, який включає математичну статистику, методи оптимізації, дискретну математику, алгебру і геометрію.

Розпізнавання образів має широке застосування і використовується при створенні усіх комп'ютерних систем, на які покладаються інтелектуальні функції, тобто функції, пов'язані із прийняттям рішень замість людини: медична діагностика, криміналістична експертиза, пошук інформації та інтелектуальний аналіз даних тощо.

Прецедент – це об’єкт, приналежність якого до заданого класу визначена заздалегідь. Прецедентом може бути, наприклад, набір ознак пацієнта із відомим діагнозом, з яким слід порівнювати набір ознак людини, діагноз якої ще невідомий.

Кожний образ являє собою набір чисел, що описують його властивості і називаються ознаками. Упорядкований набір ознак об’єкта називається вектором ознак. Вектор ознак – це точка в просторі ознак.

Задача розпізнавання тексту, написаного рукою людини унікальна в сенсі індивідуальних особливостей написання як самих символів, так і структурних складових текстових даних.

У ході виконання дипломного проекту необхідно формалізувати задачу розпізнавання печатних символів і представити її в контексті нейромережних технологій. Необхідно розробити архітектуру штучної мережі розпізнавання символів або комплексу нейронних мереж, що дозволяє з максимальною точністю виконувати процес розпізнавання вхідних даних.

3.1.2 Задача моделювання нейронних мереж для класифікації зображень символів

Розроблено багато систем, що демонструють можливості штучних нейронних мереж: мережі здатні подавати текст фонетично, розпізнавати рукописні букви, стискувати зображення. Більшість потужних мереж, що здатні розпізнавати символи та звуки, використовують як основу принцип зворотного поширення [11], що є систематичним методом для навчання багат шарових мереж, тим самим, переборює обмеження. Нейронні мережі, які працюють за вищевказаним принципом прийшли на зміну системам, що складались з одного шару штучних нейронів і використовувались для широкого класу завдань, у тому числі для організації штучного зору. Однак

такі системи мають недоліки. Насамперед немає гарантії, що мережа може бути навчена за скінченний час.

Кожен із розроблених алгоритмів навчання мереж має свої специфічні переваги, але спільним недоліком є обмеження у своїх можливостях навчатися й згадувати.

При навчанні мережі кожна вхідна (або вихідна) множина сигналів розглядається як вектор. Навчання здійснюється шляхом послідовного пред'явлення вхідних векторів з одночасним налаштуванням ваг відповідно до певної процедури. У процесі навчання ваги мережі поступово стають такими, щоб кожний вхідний вектор відповідно виробляв вихідний. Навчальні алгоритми можуть бути класифіковані як алгоритми навчання без учителя й з учителем.

У першому випадку, при пред'явленні вхідних символів мережа самоорганізується за допомогою настроювання своїх ваг у відповідності з визначеним алгоритмом. Унаслідок відсутності вказівки наперед визначеного виходу, у процесі навчання результати непередбачені з погляду визначення збудливих символів для конкретних нейронів. При цьому, мережа організується у формі, що відбиває істотні характеристики навчального набору. Наприклад, вхідні символи можуть бути класифіковані відповідно до ступеня їхньої подібності так, що символи одного класу активізують той самий вихідний нейрон.

Зазначенні методи навчання мережі передбачають, що відомі лише вхідні вектори, а на їхній основі мережа вчиться давати найкращі значення вихідної функції.

У другому випадку існує вчитель, що надає вхідні символи мережі, порівнює результуючі виходи з необхідними, а потім налаштовує ваги мережі таким чином, щоб зменшити розходження.

Методи навчання мережі з учителем передбачають, що існують пари «вхід-вихід», тобто відомі значення вхідних векторів, і значення вихідних векторів, що їм відповідають.

Таким чином, нейронна мережа, яка навчається з учителем, є більш надійною, оскільки при певному вхідному сигналі на виході формується відповідне вихідне значення.

Навчаючи мережу розпізнавати нові символи, дуже часто знищуються або змінюються результати попереднього навчання. Якщо існує тільки фіксований набір навчальних векторів, вони можуть пред'являтися при навчанні циклічно. У мережах зі зворотним поширенням, навчальні вектори подаються на вхід мережі послідовно, поки мережа не «навчиться» всьому вхідному набору. Проте цілком навчена мережа повинна запам'ятати новий навчальний вектор, і він може змінити ваги настільки, що буде потрібно повне перенавчання мережі.

3.1.3 Навчання перцептрона

В якості головного компонента системи виступає видозмінений (багатошаровий) перцептрон [12, 14]. Кількість нейронів перцептрона визначається кількістю вивчених ним образів (m), кількість синапсів (n) визначається розширенням рисунків, які розпізнаються.

Розглянемо алгоритм зворотного поширення для перцептрона. Символ подається рисунком, розміри якого $79,38 \times 52,92$ мм, тип — бітовий, розширення *.bmp.

Мережею рисунок розбивається на 600 частин (по горизонталі на 20, по вертикалі на 30). Отже, кількість синапсів для одного нейрона буде незмінним числом і визначатися як: $n = 20 \times 30 = 600$.

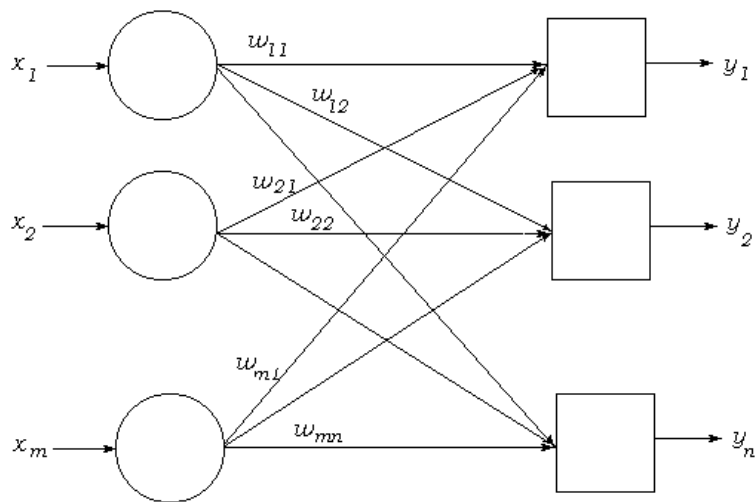


Рисунок 3.1 – Структура видозміненого перцептрона

Навчання перцептрона з урахуванням запропонованого підходу можна подати наступним чином [18]:

- ініціалізуються всі ваги мережі в малі ненульові величини.
- на вхід мережі подається вхідний навчальний вектор X вхідних символів і обчислюється сигнал NET вихідної функції від кожного нейрона, використовуючи формулу:

$$NET_j = \sum_i x_i w_{ij},$$

де x_i – i -й елемент вхідного вектора X ; w_{ij} – вага входу i нейрона j ;

NET_j – рівень збудження післясинаптичного нейрона.

- а) обчислюється значення граничної функції активації для сигналу NET від кожного нейрона в такий спосіб:

$$OUT_j = 1, \text{ якщо } NET_j \text{ більше ніж порогови } j,$$

$$OUT_j = 0 \text{ – у протилежному випадку.}$$

б) обчислюється помилка для кожного нейрона:

$$error_j = target_j - OUT_j$$

де $error_j$ – помилка для j -го нейрона; $target_j$ – необхідний вихід j -го нейрона; OUT_j – отриманий вихід j -го нейрона.

в) кожна вага модифікується в такий спосіб:

$$w_{ij}(t+1) = w_{ij}(t) + ax_j error_j,$$

де $w_{ij}(t+1)$ – вага входу i нейрона j в момент часу $(t+1)$; $w_{ij}(t)$ – вага входу i нейрона j в момент часу t ; a – нормуючий коефіцієнт навчання, початкове значення якого 0,1, і який зменшується в процесі навчання.

Повторюються кроки з другого до п'ятого, доки помилка не стане досить малою.

Робота інтелектуального модуля з розпізнавання символів полягає в підрахуванні виходів перцептрона й, якщо два або більше нейронів на виході отримали максимальне значення, то переглядається який нейрон має більшу суму на виході його суматора, і вносяться зміни до результату [13].

Активаційна функція нейронів має вигляд лінійного порога :

$$Y = S \text{ при } S < T,$$

$$Y = S \text{ при } S \geq T,$$

де Y – значення активаційної функції; S – аргумент активаційної функції; T – величина порога.

3.1.4 Дельта-правило

Подальше навчання перцептрона здійснюється під час роботи інтелектуального модуля штучних нейронних мереж із розпізнавання символів [14]. Ваги синапсів корегуються відповідно Дельта-правилу:

$$\delta = (D - Y),$$

де δ – рівна різниці між необхідним або цільовим виходом D і реальним виходом Y .

Алгоритм навчання зберігається, якщо δ множиться на величину кожного входу x_i і цей добуток додається до відповідної ваги. З метою організації управління середньою величиною зміни ваг уведемо коефіцієнт «швидкості навчання» η , який множиться на δx_i , що дозволяє управляти середньою величиною зміни ваги. У алгебраїчній формі запису:

$$\Delta_i = \eta \delta x_i,$$

$$w_i(n+1) = w_i(n) + \Delta_i,$$

де Δ_i – корекція, пов'язана з i -м входом x_i ; η – коефіцієнт швидкості навчання; $w_i(n+1)$ – i -е значення ваги після корекції; $w_i(n)$ – i -е значення ваги до корекції.

Дельта-правило модифікує ваги відповідно до i -го необхідного і дійсного значень виходу кожної полярності як для безперервних, так і для бінарних входів і виходів. Таким чином, використання зворотного поширення для перцептрона, дозволяє досягти ефективного розпізнавання навіть зашумлених символів. Результати практичної реалізації розроблених

алгоритмів необхідно проаналізувати з точки зору ефективності та точності роботи.

Незважаючи на свої обмеження перцептрони широко вивчалися (хоч не дуже широко використовувалися). Теорія перцептронів є основою для багатьох інших типів штучних нейронних мереж, і перцептрони ілюструють важливі принципи. Внаслідок цих причин вони є логічною початковою точкою для вивчення штучних нейронних мереж.

3.2 Термін образу. Проблема навчання розпізнавання образів

Розпізнавання образів (а часто говорять - об'єктів, сигналів, ситуацій, явищ або процесів) – найпоширеніша завдання, яку людині доводиться вирішувати практично щомиті від першого до останнього дня свого існування. Для цього вона використовує величезні ресурси свого мозку.

Методи навчання розпізнавання образів – перцептрони, нейронні мережі, метод потенційних функцій, метод групового обліку аргументів, метод граничних спрощень, колективи вирішальних правил [15].

Предмет розпізнавання образів об'єднує ряд наукових дисциплін. Їх пов'язує пошук рішення загальної задачі – виділити елементи, що належать конкретному класу, серед множини розмитих елементів, що відносяться до кількох класів.

Клас образів – деяка категорія, що визначається рядом властивостей загальних для всіх її елементів.

Образ – класифікаційне угруповання в системі класифікації, яке об'єднує (виділяє) певну групу об'єктів за певною ознакою.

Образне сприйняття світу – один з загадкових властивостей живого мозку, що дозволяє розібратися в нескінченному потоці інформації і зберігати орієнтацію в океані розрізнених даних про зовнішньому світі.

Образи володіють характерною властивістю, що виявляється в тому, що ознайомлення з кінцевим числом явищ з одного і того ж дає можливість дізнаватися як завгодно велике число його представників. Образи володіють характерними об'єктивними властивостями в тому сенсі, що різні люди, що навчаються на різному матеріалі спостережень, здебільшого однаково і незалежно один від одного класифікують одні й ті самі об'єкти. Саме ця об'єктивність образів дозволяє людям усього світу розуміти один одного.

У класичній постановці задачі розпізнавання універсальне розбивається на частини-образи. Кожне відображення будь-якого об'єкта на сприймаючі органи розпізнавальної системи, незалежно від його положення щодо цих органів, прийнято називати зображенням об'єкта, а безлічі таких зображень, об'єднані будь-якими загальними властивостями, являють собою образи.

Методика віднесення елемента до якого-небудь образом називається вирішальним правилом. Ще одне важливе поняття – метрика, спосіб визначення відстані між елементами універсальної множини. Чим менше ця відстань, тим більше схожими є об'єкти (символи, звуки та інші) – те, що ми розпізнаємо. Звичайно елементи задаються у вигляді набору чисел, а метрика – у вигляді функції. Від вибору представлення образів і реалізації метрики залежить ефективність програми, один алгоритм розпізнавання з різними метриками буде помилятися з різною частотою.

Сприймаючи зовнішній світ, ми завжди робимо класифікацію сприйманих відчуттів, тобто розбиваємо їх на групи подібних, але не тотожних явищ. Наприклад, незважаючи на істотну відмінність, до однієї групи належать усі літери А, написані різними почерками, або всі звуки, що відповідають одній і тій же ноті, взятої у будь-якій октаві і на будь-якому інструменті. Дитині можна показати лише один раз будь-яку літеру, щоб вона змогла знайти цю літеру в тексті, написаному різними шрифтами, або знайти її, навіть якщо вона написана в навмисно спотвореному вигляді. Ця властивість мозку дозволяє сформулювати таке поняття, як образ.

Здатність сприйняття зовнішнього світу у формі образів дозволяє з певною вірогідністю впізнавати нескінченну кількість об'єктів.

3.3 Нейронні мережі

3.3.1 Біологічний нейрон

Нервова система людини, побудована з елементів, які називаються нейронами і має приголомшливу складність. Близько 1011 нейронів беруть участь в приблизно 1015 передавальних зв'язках, що мають довжину метр і більше. Кожен Нейрон володіє багатьма якостями, спільними з іншими елементами тіла, але його унікальною здатністю є прийом, обробка і передача електрохімічних сигналів по нервових шляхах, які утворюють комунікаційну систему мозку [16].

Нейрон (нервова клітка) складається з тіла клітини, і двох типів зовнішніх деревоподібних відгалужень: аксона і дендритів. Тіло клітини вміщує ядро, що містить інформацію про властивості нейрона, і плазму, яка продукує необхідні для нейрона матеріали. Нейрон отримує сигнали (імпульси) від інших нейронів через дендрити (приймачі) і передає сигнали, що згенеровані тілом клітки, вздовж аксона (передавач), який наприкінці розгалужується на волокна. На закінченнях волокон знаходяться синапси.

Синапс є функціональним вузлом між двома нейронами (волокно аксона одного нейрона і дендрит іншого). Коли імпульс досягає синаптичного закінчення, там продукуються хімічні речовини – нейротрансмітери. Нейротрансмітери проходять через синаптичну щілину і в залежності від типу синапсу, збуджують або гальмують здатність нейрона-приймача генерувати електричні імпульси. Результативність синапсу налаштовується сигналами, які проходять скрізь нього, тому синапси навчаються у відповідності до активності процесів, у яких вони приймають

участь. Нейрони взаємодіють за допомогою короткої серії імпульсів, повідомлення передається за допомогою частотно-імпульсної модуляції.

Біологічні нейрони є структурно складнішими, ніж існуючі штучні нейрони. Оскільки нейрофізіологія надає науковцям розширене розуміння дії нейронів, а технологія обчислень постійно вдосконалюється, розробники мереж мають необмежений простір для вдосконалення моделей біологічного мозку.

3.3.2 Штучний нейрон

Штучний нейрон є базовим модулем нейронних мереж. Він моделює основні функції природного нейрона (рис. 3.2).

При функціонуванні нейрон одночасно отримує багато вхідних сигналів [17]. Кожен вхід має свою власну синаптичну вагу, яка надає входу вплив, необхідний для функції суматора елемента обробки. Ваги є мірою сили вхідних зв'язків і моделюють різноманітні синаптичні сили біологічних нейронів. Ваги суттєвого входу підсилюються і, навпаки, вага несуттєвого входу примусово зменшується, що визначає інтенсивність вхідного сигналу. Ваги можуть змінюватись відповідно до навчальних прикладів, топології мережі та навчальних правил.

Вхідні сигнали x_n зважені ваговими коефіцієнтами з'єднання w_n додаються, проходять через передатну функцію, генерують результат і виводяться.

В програмних реалізаціях штучні нейрони називають «елементами обробки» або «процесорами» і вкладають в них більше можливостей, ніж в базовому штучному нейроні, що описаний вище.

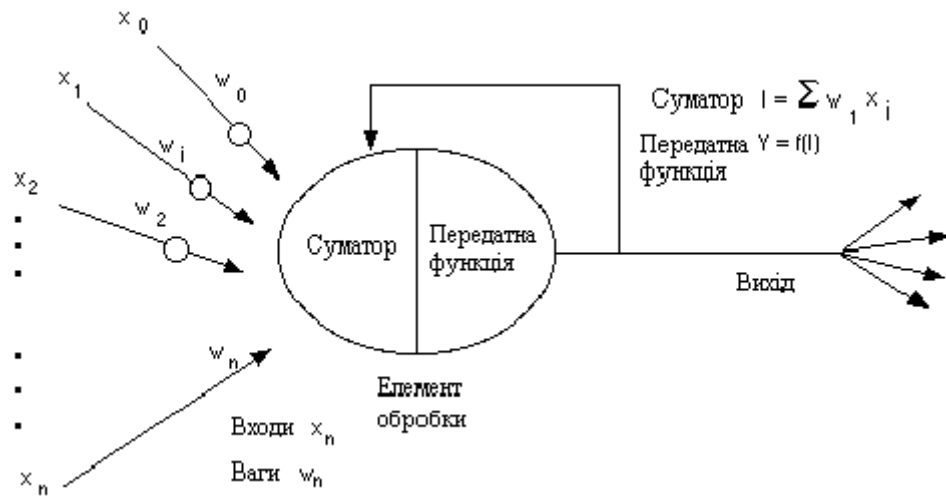


Рисунок 3.2 – Базовий штучний нейрон

На рисунку 3.3 зображена детальна схема штучного нейрону.

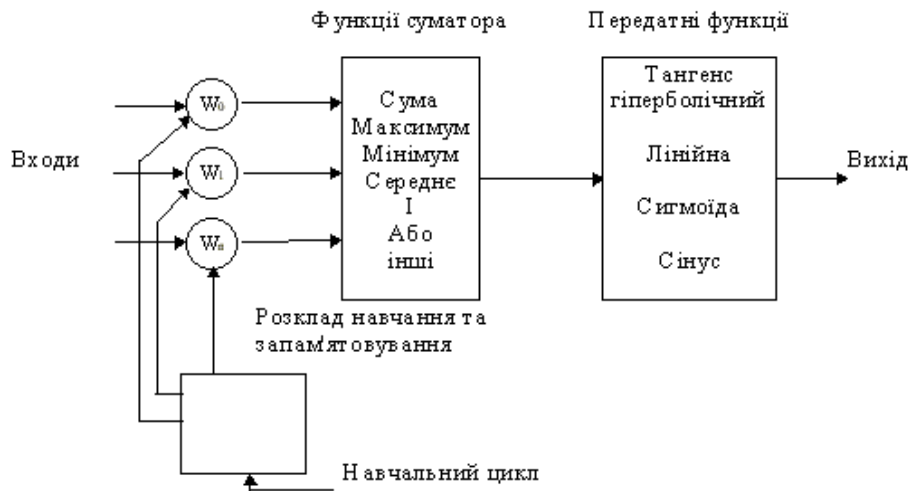


Рисунок 3.3 – Модель «елементу обробки»

Функція суматора може бути складнішою, наприклад, вибір мінімуму, максимуму, середнього арифметичного, добутку або обчислюватися за іншим алгоритмом.

Перед надходженням до передатної функції вхідні сигнали та вагові коефіцієнти можуть комбінуватись багатьма способами. Алгоритми для комбінування входів нейронів визначають відповідно до мережної архітектури та парадигми.

3.3.3 Формальний нейрон

У роботі Уоррена С. Мак-Каллока та Вальтера Пітса було започатковано теорію, в основі якої лежав той факт, що всі аспекти нервової діяльності можна моделювати за допомогою мережі елементів, які мають два стійкі стани [16, 17]. Модель такого елемента, яка показана на рис.3.4, отримала назву формального нейрона.

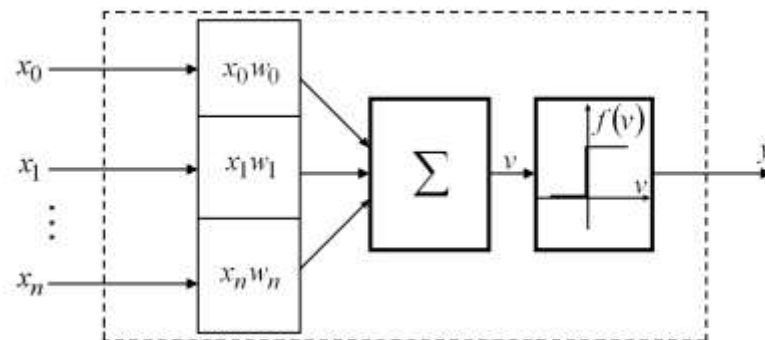


Рисунок 3.4 – Формальний нейрон

Моделювання функцій синапсів відбувається шляхом масштабування вхідних сигналів (x_0, x_1, \dots, x_n) за допомогою вагових коефіцієнтів (w_0, w_1, \dots, w_n) . Одержані сигнали подаються на вхід суматора, який виконує їх обробку за формулою:

$$v = \sum_{i=0}^n w_i x_i.$$

Змінна v є аргументом порогової активаційної функції:

$$y = f(v).$$

В деяких нейронних мережах суматор виконує додаткову обробку, так звану функцію активації, яка зміщує вихід функції суматора в часі. Цю функцію найкраще використовувати як компоненту мережі в цілому, ніж як компоненту окремого нейрона. Часто, ця функція є відсутньою. Активаційна функція, що вперше була запропонована, має вигляд:

$$y = \begin{cases} 1, v \leq 0, \\ 0, v \geq 0. \end{cases}$$

У сучасній літературі зустрічається велика кількість парадигм штучних нейронних мереж, елементи яких реалізують різні активаційні функції. Найбільш поширеною є сигмоїдальна функція, що може бути представлена в дискретному та аналоговому варіанті.

Прикладом найпростішої активаційної функції є дискретний сигмоїд із параметром, що має назву функції Хевісайда або тета функції:

$$y = \begin{cases} 1, v \leq a, \\ 0, v > a. \end{cases}$$

Якщо величина збудження v нейрона з активаційною функцією Хевісайда не перевищує значення параметра a , то нейрон залишається пасивним, а при перевищенні порога видає фіксоване значення функції, що приймається за логічну одиницю.

Найпоширенішою нелінійною аналоговою активаційною функцією є сигмоїдальна логічна функція з параметрами, b c і d , яку задають виразом:

$$y = \frac{b}{c + e^{dv}}.$$

При одиничних значеннях параметрів $b=1$, $c=1$, $d=-1$ одержуємо:

$$y = \frac{1}{1 + e^{-v}}.$$

Амплітуда вихідного сигналу нейрона з даною активаційною функцією залежить від амплітуди вхідних сигналів. З рис.3.5 видно, що крутизна S-подібної функції залишається високою тільки у деякому визначеному діапазоні.

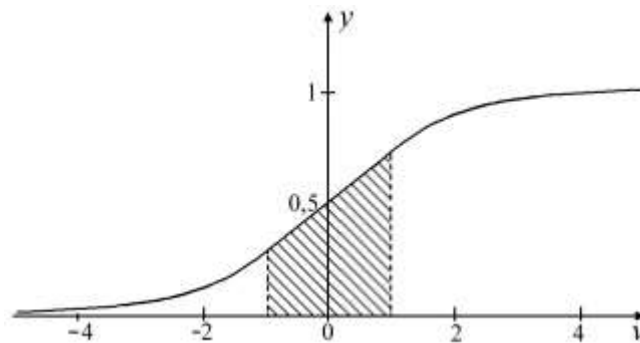


Рисунок 3.5 – Сигмоїдальна логічна функція

Ця властивість була вперше використана Гросбергом для вирішення поставленої ним дилеми шумового насичення. Завдяки нелінійному коефіцієнту підсилення нейрон може надійно працювати в широкому діапазоні рівнів вхідних сигналів. Ще однією корисною властивістю аналогової сигмоїдальної активаційної функції є диференційованість на всій осі абсцис та простий вираз для похідної

$$y' = y(1 - y),$$

де $y = \frac{1}{1 + e^{-v}}.$

Подібною до функції Хевісайда є функція гіперболічного тангенса (рис.3.6).

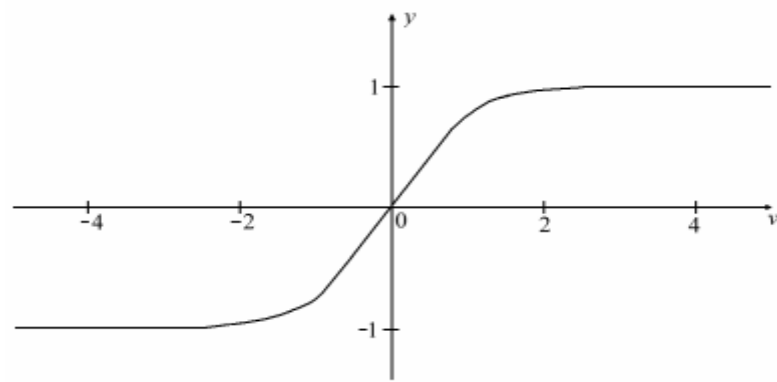


Рисунок 3.6 – Функція гіперболічного тангенса

Основна відмінність полягає у симетрії відносно початку координат, що дозволяє працювати у від'ємному діапазоні вхідних сигналів. Гіперболічну функцію задано виразом, які призначені для регулювання параметрів вихідного сигналу нейрона:

$$y = k \tanh(mv).$$

При $k = 1$, $m = 1$ одержимо варіант активаційної функції

$$y = \tanh(v),$$

що має найбільше поширення. Подібно до сигмоїдальної функції, в дискретному та аналоговому варіанті існує дзвонovidна активаційна функція. Її графічне зображення показано на рис.3.7.

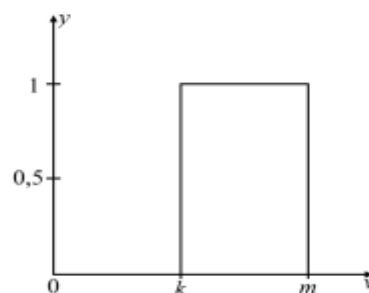


Рисунок 3.7 – Дискретна дзвонovidна функція

$$y = \begin{cases} 1, v > k, v \leq m \\ 0, v \leq k, v > m. \end{cases}$$

Аналогова дзвонovidна функція використовується у випадку аналогового моделювання нейронів із властивостями смугового фільтра. Розрізняють дзвонovidні експоненціальні функції першого та другого порядку.

Аналогову дзвонovidну функцію першого порядку у загальному випадку задають виразом, що містить параметри k, m :

$$y = \frac{1}{ke^{mv} + ne^{pv} + r}.$$

Задавши $k=1, m=1, n=1, p=-1, r=-1$, одержимо базовий варіант дзвонovidної функції першого порядку, графік якої представлений на рис.3.8:

$$y = \frac{1}{e^v + e^{-v} - 1}.$$

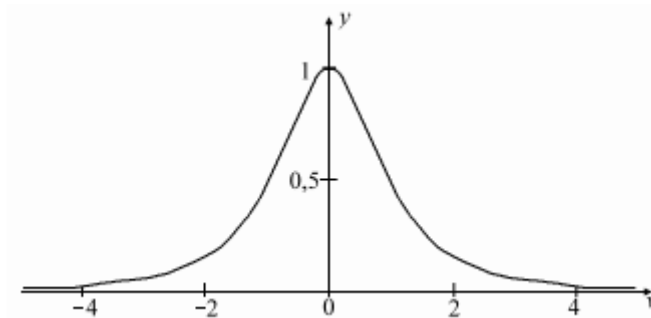


Рисунок 3.8 – Аналогова дзвонovidна функція першого порядку

Дзвонovidну функцію другого порядку задають виразом $Y = e^{-v^2}$. Як видно з рисунку 3.9, її графік нагадує дзвонovidну функцію першого

порядку, але має більшу крутизну, тому може використовуватися в смугових фільтрах із жорсткішими характеристиками.

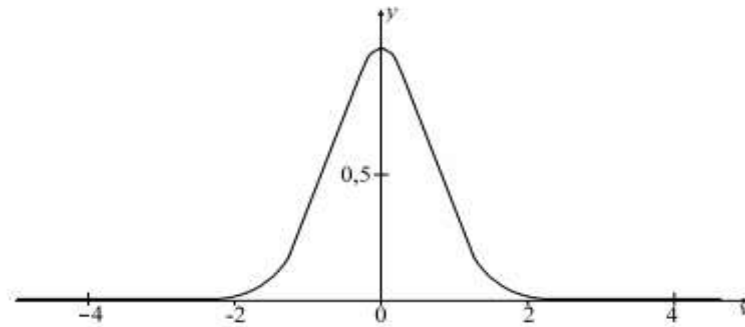


Рисунок 3.9 – Аналогова дзвоновидна функція другого порядку

3.3.4 Одношарові штучні нейронні мережі

Штучна нейронна мережа – паралельно розподілений процесор, який володіє здатністю до навчання, збереження і представлення знань, набутих на основі досвіду.

Штучні нейронні мережі, зокрема багатошаровий перцептрон, вирішують задачі регресії і класифікації. Проте, на відміну від дерев рішень, нейронні мережі не здатні пояснити отримане рішення, тому їх функціонування нагадує «чорний ящик» з входами і виходами.

Нейронні мережі є обчислювальними структурами, що моделюють прості біологічні процеси, подібні до тих, що відбуваються в людському мозку. Вони здатні до адаптивного навчання шляхом реакції на позитивні і негативні дії.

Подібність штучної нейронної мережі з мозком полягає в двох аспектах:

- а) знання набуваються мережею під час навчання;
- б) для збереження знань використовуються міжнейронні з'єднання.

Як відомо, один перцептрон може виконувати найпростіші процедури класифікації та розпізнавання. При цьому потужність нейронних обчислень збільшується від з'єднання перцептронів в мережах. Найпростіша НМ - одношаровий перцептрон, який складається з групи перцептронів, що утворюють шар [17, 19]. Каскадуванням таких шарів отримують багатошарові мережі, в яких вихід одного шару є входом для наступного. Саме таким нейронна мережа властиві великі обчислювальні можливості.

Проста модель перцептрона має трьохелементну структуру (помножувач, суматор, нелінійний перетворювач), в якій реалізується перемноження вектора вагових коефіцієнтів на вектор входних операндів з наступним підсумуванням зважених входів та порівнянням отриманої суми із зовнішнім порогом обробки. Математично ці операції можна записати так:

$$S = \sum_{i=1}^n w_i * x_i,$$

$$Y = f(S) = \text{sign}(S - \theta),$$

де x_i – i -й входний сигнал; w_i – i -й ваговий коефіцієнт; S – суматорна функція; $f(S)$ – активаційна функція(нелінійне перетворення), Y – вихідний сигнал перцептрона, що реалізує найпростішу знакову функцію – функцію одиничного стрибка, θ – зовнішній поріг обробки.

Така модель перцептрона реалізує простіший класифікатор, тобто є лінійним дискримінатором, оскільки вихідний сигнал Y формує значення так званої дискримінантної функції, яка залежить лише від лінійної комбінації входів. Так, якщо входні вектори можуть належати одному з двох класів, перцептрон здатен розрізнити тип входу таким чином: $Y \geq 0$ – входний вектор належить першому класу, в протилежному випадку – другому.

Простота такої моделі перцептрона не зменшує її цінності, оскільки на її основі ефективно розв'язуються задачі обробки двійкових даних, класифікації, розпізнаванням, ідентифікації тощо.

Всі вагові коефіцієнти одношарової нейронної мережі утворюють матрицю вагових коефіцієнтів:

$$W = \begin{pmatrix} w_{11} & w_{1j} & w_{1m} \\ w_{i1} & w_{ij} & w_{im} \\ w_{n1} & w_{nj} & w_{nm} \end{pmatrix}.$$

Тоді вектор аргументів визначається як добуток $V = XW$, а вектор вихідних сигналів є вектором значень активаційних функцій:

$$Y = F(V) = \begin{cases} f_1(v_1), \\ f_j(v_j), \\ f_m(v_m). \end{cases}$$

Згадані функції можуть бути однаковими для всіх нейронів, що складають мережу. У цьому випадку її називають гомогенною. Мережу, в якій вигляд активаційної функції залежить від номера нейрона, називають гетерогенною.

На рис.3.10 показаний загальний випадок одношарової нейронної мережі, у якій кожен нейрон приймає участь в обробці всіх елементів вхідного вектора даних. Такий підхід не завжди є економічно та технічно виправданим. Тому для вирішення конкретних задач можуть використовуватись архітектури зі структурою зв'язків, яка є підмножиною повної зв'язності.

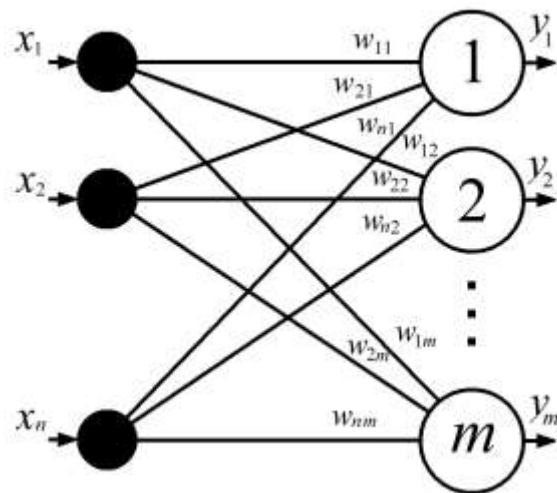


Рисунок 3.10 – Одношарова нейронна мережа

3.3.5 Багатошарові штучні нейронні мережі

Спроби застосувати одношарові нейронні мережі для розв'язування широкого кола задач наštовхнулися на ряд труднощів, пов'язаних з проблемою лінійної роздільності. Природним вирішенням цієї проблеми стало застосування багатошарових штучних нейронних мереж, що нагадують багатошарові структури мозку [20].

Багатошарові мережі володіють значно більшими можливостями, ніж одношарові. Проте багатошарові мережі можуть привести до збільшення обчислювальної потужності в порівнянні з одношаровими лише в тому випадку, якщо активаційна функція між шарами буде нелінійною.

Розглянемо ієрархічну структуру, в якій нейрони структуровані за шарами. Вона складається з нейронів першого (прихованого) шару, які одночасно здатні прийняти вхідний вектор сигналів $X = (x_1, \dots, x_i, \dots, x_n)$ та k нейронів другого (вихідного) шару. Така структура зображена на рис.3.11 і є узагальненою структурою багатошарової нейронної мережі прямого поширення.

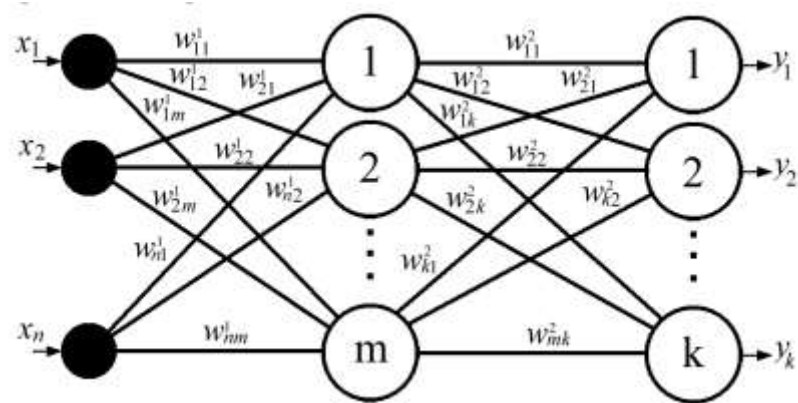


Рисунок 3.11 – Багатошарова нейронна мережа прямого поширення

У нейронних мережах прямого поширення синаптичні зв'язки організовані таким чином, що кожний нейрон даного рівня ієрархії сприймає інформацію тільки від деякої непустої множини нейронів, які розташовані на більш низькому рівні. Назва мереж вказує на те, що у них існує виділений напрям поширення сигналів, які рухаються, починаючи з входу, через один або декілька прихованих шарів до вихідного шару.

Легко помітити, що багатошарова нейронна мережа може бути одержана шляхом каскадного об'єднання одношарових мереж з матрицями вагових коефіцієнтів w^1, w^2, \dots, w^p , де p – кількість шарів нейронної мережі.

У випадку лінійності активаційних функцій багатошарова нейронна мережа може бути зведена до еквівалентної одношарової з матрицею вагових коефіцієнтів $w^1 \times w^2 \times \dots \times w^p$, тому формування подібних структур має сенс тільки у випадку застосування у нейронах нелінійних активаційних функцій.

4 ПРОГРАМНА РЕАЛІЗАЦІЯ Й ДОСЛІДЖЕННЯ РОЗРОБЛЕНИХ АЛГОРИТМІВ

4.1 Архітектура хмарного сервісу та клієнтської програми для розпізнавання тексту в відео потоці

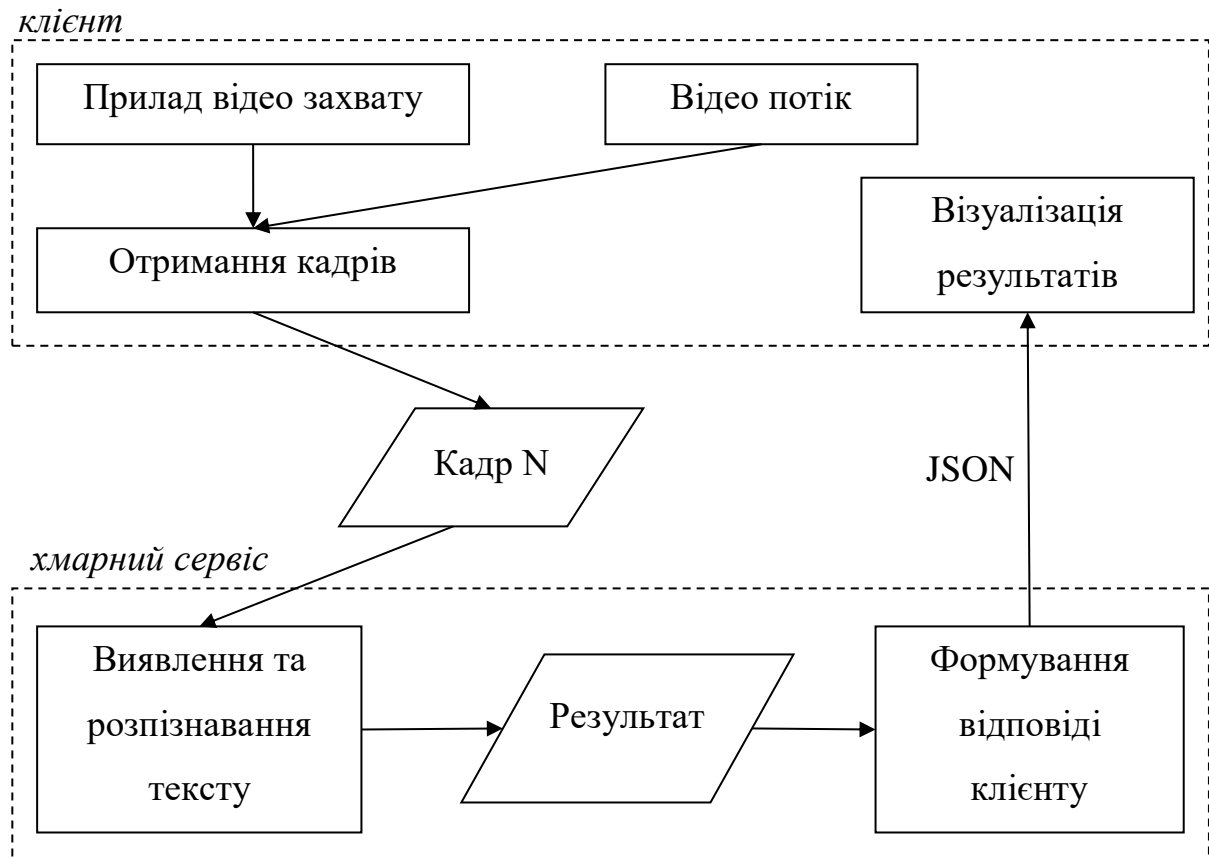


Рисунок 4.1 – Архітектура хмарного сервісу

Розглянемо детально компоненти системи.

Клієнтська частина реалізована на базі JavaFX платформи. Це програма, яка відповідає за підготовку відео кадрів, що передаються на обробку до хмарного сервісу, а також за візуалізацію результатів, які повертає у відповідь хмарній сервіс.

Обмін даними між хмарним сервісом та клієнтською програмою реалізований з використанням протоколу HTTP.

Хмарний сервіс – це спеціалізована служба, яка фізично розташована на віддаленому сервері та має змогу обробляти чи зберігати данні, які

надходять до служби від віддалених клієнтів. У нашому випадку, хмарний сервіс обробляє відео кадри, що надходять з клієнта, та повертає клієнту результати обробки у форматі JSON.

4.2 Реалізація алгоритму виявлення та розпізнавання тексту

Розглянемо детально алгоритм виявлення та розпізнавання тексту, який зображений на рисунку 4.2.

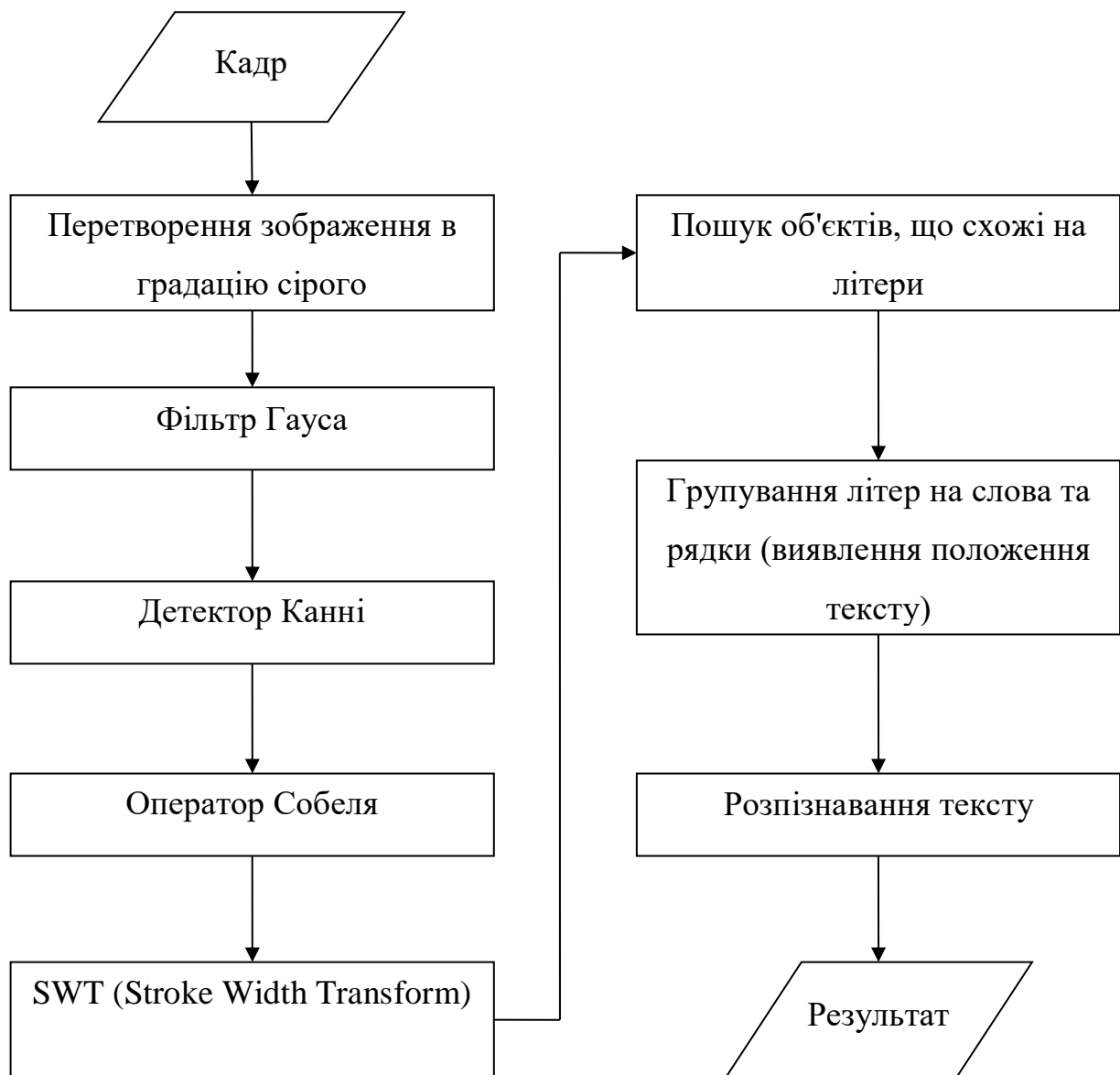


Рисунок 4.2 – Алгоритм виявлення та розпізнавання тексту

На вхід системі подається зображення у форматі RGB.



Рисунок 4.3 – Вхідне зображення

Далі зображення перетворюється із кольорового у відтінки сірого та згладжується з використанням фільтра Гауса радіусом 3*3 або 5*5.



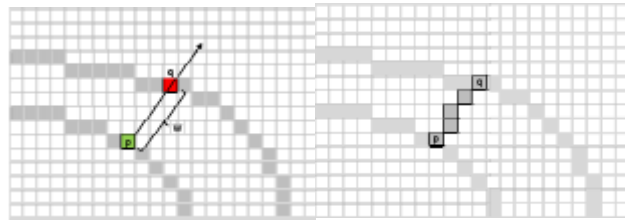
Рисунок 4.4 – Результат застосування фільтру Гауса

SWT перетворення основане на аналізі границь, то ж наступним кроком застосовуємо детектор Канні.



Рисунок 4.5 – Результат виділення контурів з використанням детектора Канні

Спочатку кожному пікселю SWT зображення присвоюємо -1. Далі починаємо аналізувати кожний піксель p зображення, отриманого детектором Канні, враховуючи напрямок градієнту D_p . Напрямок градієнту обчислюємо за допомогою оператора Собеля. Якщо p знаходиться на межі границі штриха, проводимо промінь перпендикулярно напрямку градієнту. Продовжуємо промінь $r = p + nd_p$, $n > 0$ до тих пір, поки не зустрінемо іншу межу штриха q . Пікселі SWT зображення s , що відповідають променю $[p, q]$, оновлюється за правилом $s = \frac{\sum_{i=p}^q \text{img}[i]}{\|p - q\|}$. Алгоритм побудови променя графічно зображений на рисунку 4.6.



а)

б)

Рисунок 4.6 – Результат побудови променя

В результаті SWT перетворення отримуємо зображення на рисунку 4.7.



Рисунок 4.7 – SWT зображення

Видно, що основна частина шумів видалена. Лишилися лише області, які фактично представляють собою контури.

На підготовленому зображенні виконуємо фільтрацію, яка відповідає за знаходження об'єктів, що схожі на літери. Отриманий результат зображений на рисунку 4.8.



Рисунок 4.8 – Результат знаходження літер

4.3 Реалізація алгоритму розпізнавання тексту

Отримані контури розпізнаємо за допомогою одношарової нейронної мережі. Спочатку потрібно виконати навчання мережі. За допомогою розробленого генератора, згенеруємо набір еталонних зображень для декількох варіацій шрифтів російської мови.

Графічний інтерфейс користувача програми-генератора зображений на рисунку 4.9.

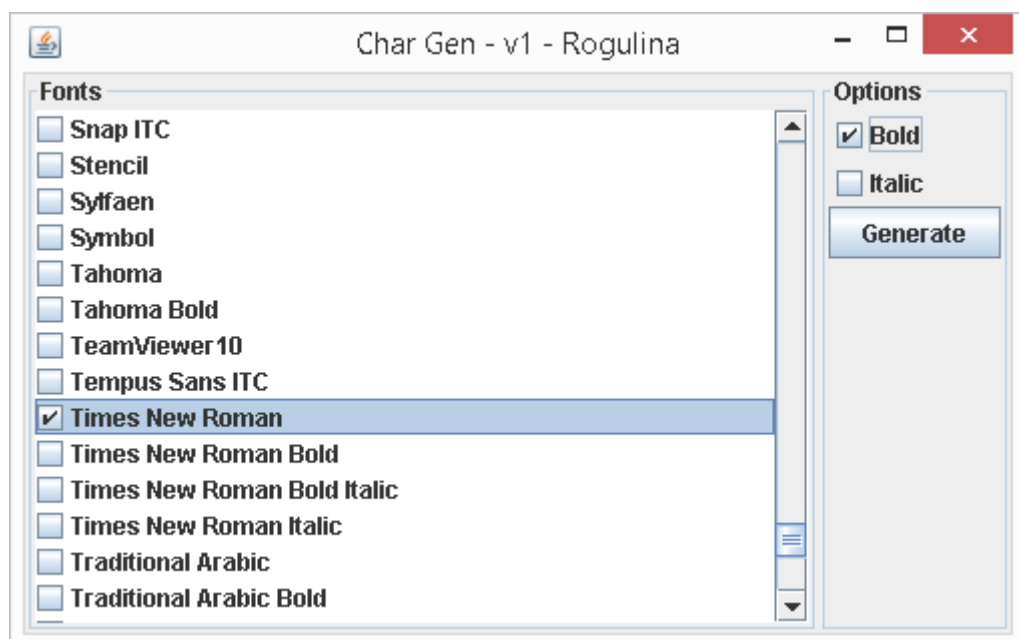


Рисунок 4.9 – Інтерфейс користувача програми-генератора

Навчання нейронної мережі виконаємо за допомогою алгоритма з вчителем.

Навчання нейронної мережі з учителем припускає, що для кожного вхідного вектора з навчальної множини існує необхідне значення вихідного вектора, званого цільовим. Ці вектора утворюють навчальну пару. Ваги мережі змінюють доти, поки для кожного вхідного вектора не буде отриманий прийнятний рівень відхилення вихідного вектора від цільового.

Таким чином, після навчання мережі, виконуємо розпізнавання контурів. Результат роботи програми зображений на рисунку 4.10.

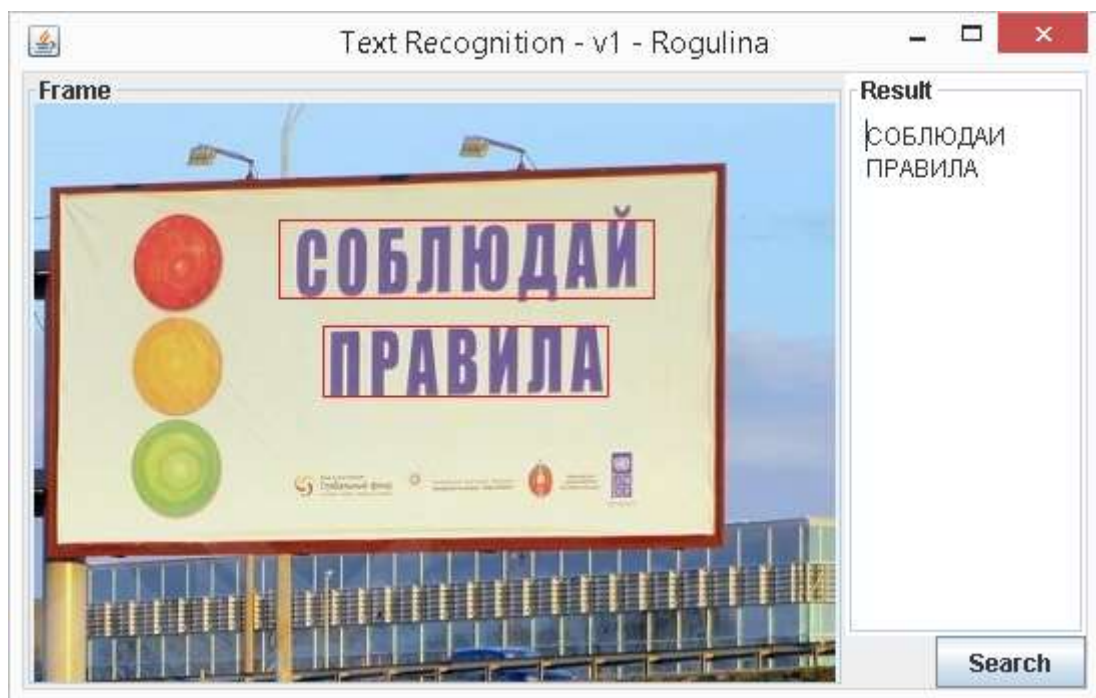


Рисунок 4.10 – Інтерфейс розпізнавання контурів

4.4 Аналіз роботи алгоритму виявлення тексту

Шум є однією з основних завд при роботі з алгоритмами розпізнавання та виявлення об'єктів на зображенні.

Проаналізуємо якість розробленого алгоритму виявлення тексту шляхом штучного впливу шуму Гауса на еталонне зображення.

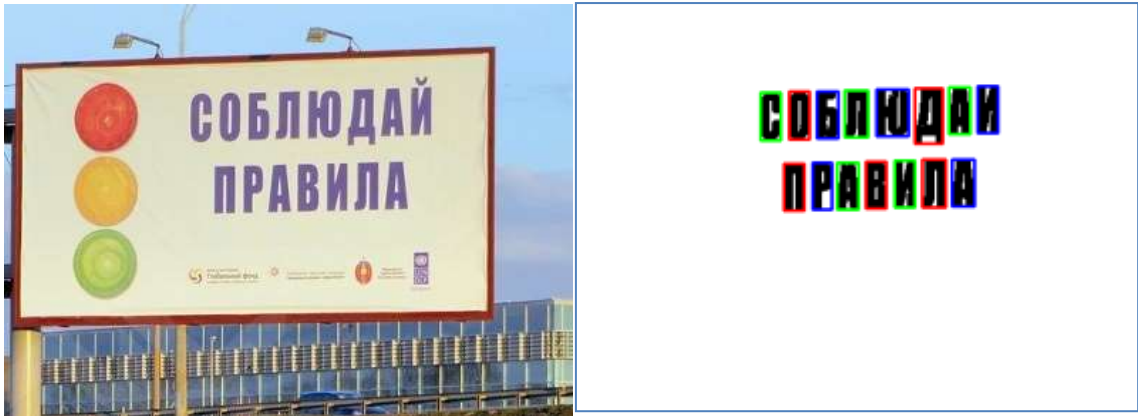


Рисунок 4.11 – Результат роботи алгоритму в ідеальних умовах

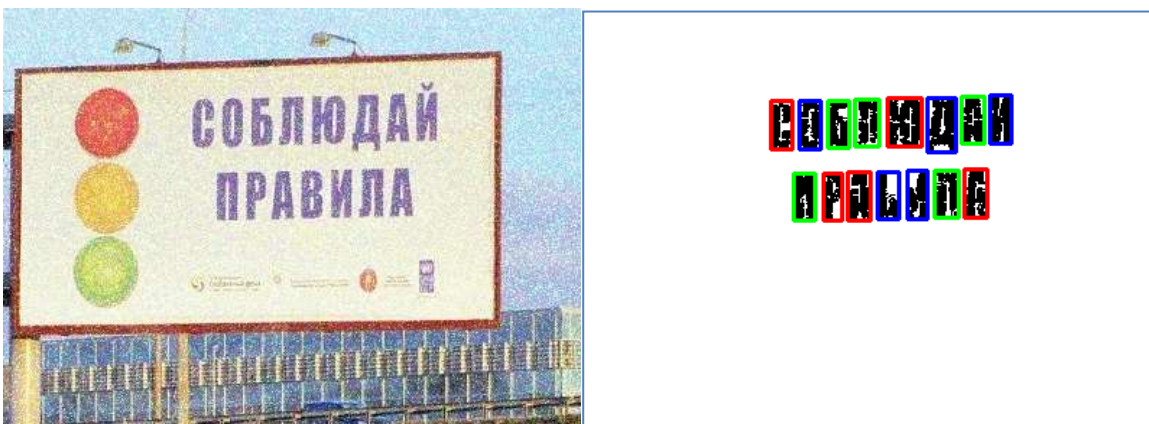


Рисунок 4.12 – Результат роботи алгоритму при зашумленості 20%



Рисунок 4.13 – Результат роботи алгоритму при зашумленості 30%

Залежність успішного виявлення тексту від зашумленості зображень наведена на рисунку 4.14

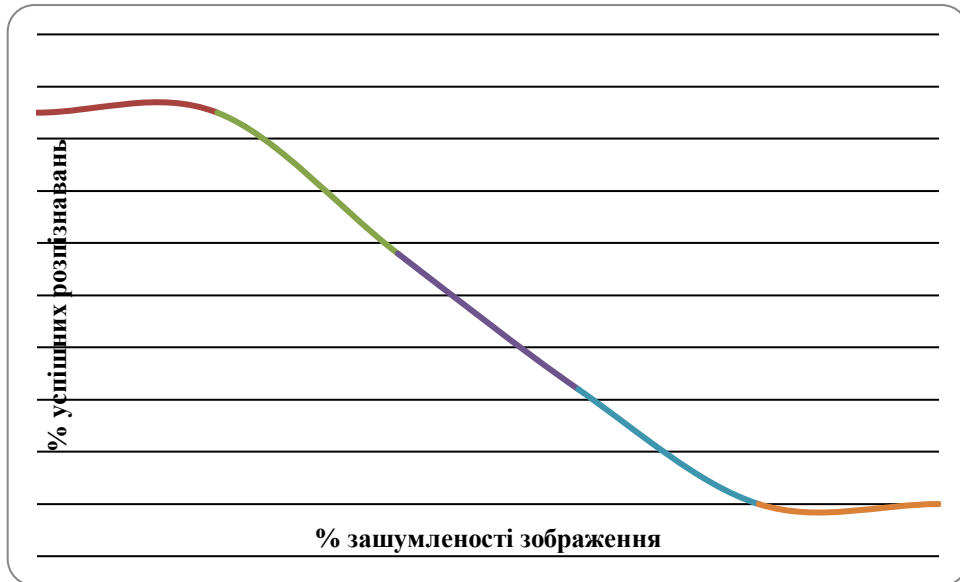


Рисунок 4.14 – Вплив шуму на успішність виявлення

5 ОХОРОНА ПРАЦІ

5.1 Аналіз потенційних небезпечних і шкідливих виробничих чинників проєктованого об'єкту, що мають вплив на персонал

У даному дипломному проєкті розробляється програмне забезпечення.

Розроблене програмне забезпечення орієнтоване на роботу з персональним комп'ютером. Експлуатовані для вирішення внутрішньовиробничих завдань ПЕОМ типу IBM PC мають наступні характеристики:

споживана потужність	220 Вт;
робоча напруга	220 В;
напруга джерел живлення	+12 В; - 12 В; +5 В;
робоча частота	50 Гц.

Виходячи з приведених характеристик, вочевидь, що для людини існує небезпека поразки електричним струмом, унаслідок недбалого поводження з комп'ютером і порушення правил експлуатації, залишення частин ПЕОМ, що знаходяться під напругою, відкритими або знятих для ремонту вузлів.

Відповідно до [20] до легкої фізичної роботи відносяться всі види діяльності, виконувані сидячи і ті, що не потребують фізичної напруги. Робота користувача ПК відноситься до категорії 1а.

При роботі на ПЕОМ користувач піддається ряду потенційних небезпек. Унаслідок недотримання правил техніки безпеки при роботі з машиною (невиконання огляду відкритих частин ПЕОМ, що знаходяться під напругою або знятих для ремонту вузлів) для користувача існує небезпека поразки електричним струмом.

Джерелами підвищеної небезпеки можуть служити наступні елементи:

- розподільний щит;

- джерела живлення;
- блоки ПЕОМ і друку, що знаходяться в ремонті.

Ще одна проблема полягає у тому, що спектр випромінювання комп'ютерного монітора включає рентгенівську, ультрафіолетову і інфрачервону області, а також широкий діапазон хвиль інших частот. Небезпека рентгенівського проміння мала, оскільки цей вид випромінювання поглинається речовиною екрану. Проте велику увагу слід приділяти біологічним ефектам низькочастотних електромагнітних полів (аж до порушення ДНК).

Відповідно до [21], при обслуговуванні ПЕОМ мають місце фізичні і психофізичні небезпечні, а також шкідливі виробничі чинники:

- підвищене значення напруги в електричному ланцюзі, замикання якої може відбутися через тіло людини;
- підвищений рівень статичної електрики;
- підвищений рівень електромагнітних випромінювань;
- підвищена або знижена температура повітря робочої зони;
- підвищений або знижений рух повітря;
- підвищена або знижена вологість повітря;
- відсутність або недостатність природного світла;
- підвищена пульсація світлового потоку;
- недостатня освітленість робочого місця;
- підвищений рівень шуму на робочому місці;
- розумове перенапруження;
- емоційні навантаження;
- монотонність праці.

5.2 Заходи щодо техніки безпеки

Основним небезпечним чинником при роботі з ЕОМ є небезпека поразки людини електричним струмом, яка посилюється тим, що органи чуття людини не можуть на відстані знайти наявності електричної напруги на устаткуванні.

Проходячи через тіло людини, електричний струм чинить на нього складну дію, що є сукупністю термічної (нагрів тканин і біологічних середовищ), електролітичної (розкладання крові і плазми) і біологічної (роздратування і збудження нервових волокон і інших органів тканин організму) дій.

Тяжкість поразки людини електричним струмом залежить від цілого ряду чинників:

- значення сили струму;
- електричного опору тіла людини і тривалості протікання через нього струму;
- роду і частоти струму;
- індивідуальних властивостей людини і навколишнього середовища.

Розроблений дипломний проект передбачає наступні технічні способи і засоби, що застерігають людину від ураження електричним струмом:

- заземлення електроустановок;
- занулення;
- захисне відключення;
- електричне розділення мережі;
- використання малої напруги;
- ізоляція частин, що проводять струм;
- огорожа електроустановок.

Занулення зменшує напругу дотику і обмежує години, протягом яких

людина, ткнувшись до корпусу, може потрапити під дію напруги.

Струм однофазного короткого замикання визначається по наближеній формулі:

$$I_k = \frac{U_{\phi}}{Z_{\Pi} + \frac{Z_T}{3}}, \quad (5.1)$$

де U_{ϕ} - номінальна фазна напруга мережі, В;

Z_{Π} - повний опір петлі, створене фазними і нульовими дротами, Ом;

Z_T - повний опір струму короткого замикання на корпус, Ом.

Згідно таблиці 4 [22]: $Z_T / 3 = 0,1$ Ом.

Для провідників і жил кабелю для розрахунку повного опору петлі використовуємо формулу(5.2.) :

$$Z_{\Pi} = \sqrt{R_{\Pi}^2 + X_{\Pi}^2}, \quad (5.2)$$

де $R_{\Pi} = R_{\phi} + R_0$ - сумарний активний опір фазного R_{ϕ} і нульового R_0 дротів, Ом;

X_{Π} - індуктивний опір паяння дротів, Ом.

Перетин 1 км мідного дроту $S = 2.5$ мм, тоді згідно таблицям 5 і 6 [22], має такий опір:

$$X_{\Pi} = 0,11 \text{ Ом};$$

$$R_{\phi} = 7,55 \text{ Ом};$$

$$R_0 = 7,55 \text{ Ом}.$$

$$\text{Отже, } R_{\Pi} = 7,55 + 7,55 = 15,1 \text{ Ом}.$$

Тоді по формулі (4.2) знаходимо повний опір петлі :

$$Z_{\Pi} = \sqrt{15,1^2 + 0,11^2} \approx 15,1 \text{ (Ом)}.$$

Струм однофазного короткого замикання рівний:

$$I_k = \frac{220}{15,1 + 0,1} = 14,47 \quad (\text{А}).$$

Дія плавкої вставки на ПЕОМ забезпечується, якщо виконується співвідношення:

$$I_k \geq k * I_n, \quad (5.3)$$

де I_n - номінальний струм спрацьовування плавкої вставки, А;

k - коефіцієнт кратності нелінійного струму I_n , А.

Коефіцієнт кратності нелінійного струму I_n розраховується по формулі (5.4.) :

$$I_n = P / U, \quad (5.4)$$

де $P = 220$ Вт - споживана потужність;

$U = 220$ В - робоча напруга;

$k = 3$ А - для плавких вставок.

Отже, $I_n = 220 / 220 = 1$ А.

Підставивши значення у вираз (5.3), одержимо:

$$14,47 > 3 * 1.$$

Таким чином, доведено, що апарат забезпечить спрацьовування(і захист) при підвищенні номінального струму.

3.3 Заходи, що забезпечують виробничу санітарію і гігієну праці

Вимоги до виробничих приміщень встановлюються [20], ДБН, відповідними ГОСТами і ОСТами з урахуванням небезпечних і шкідливих чинників, що утворюються в процесі експлуатації електроустаткування.

Підвищення працездатності людини і збереження її здоров'я забезпечується стабільними метеорологічними умовами.

Мікроклімат виробничих приміщень визначається діючими на організм людини поєднаннями температури, вологості і швидкості руху повітря, а також температури навколишніх поверхонь. Значне коливання параметрів мікроклімату приводить до порушення систем кровообігу, нервової і потовидільної, що може викликати підвищення або пониження температури тіла, слабкість, запаморочення і навіть непритомність.

Відповідно до [20] встановлюють оптимальну і допустиму температуру, відносну вологість і швидкість руху повітря в робочій зоні. За відсутності надмірного тепла, вологи, шкідливих речовин в приміщенні досить природної вентиляції.

У приміщенні для виконання робіт операторського типу(категорія 1а), пов'язаних з нервово-емоційною напругою, проектом передбачається дотримання наступних нормованих величин параметрів мікроклімату (табл. 5.1).

Таблиця 5.1 - Санітарні норми мікроклімату робочої зони приміщень для робіт категорії 1а.

Пора року	Температура, С	Відносна вологість, %	Швидкість руху повітря, м/с
Холодна	22...24	40...60	0,1
Тепло	23...25	40...60	0,1

У приміщенні, де знаходиться ПЕОМ, повітрообмін реалізується за допомогою природної організованої вентиляції (з пристроєм вентиляційних

каналів в перекриттях будівлі і вертикальних шахт) й установленого промислового кондиціонера фірми Mitsubishi, який дозволяє вирішити переважну більшість завдань по створінню та підтримці необхідних параметрів повітряного середовища. Цей метод забезпечує приток потрібної кількості свіжого повітря, визначеного в ДБН (30 м³ в годину на одного працівника).

Шум на виробництві має шкідливу дію на організм людини. Стомлення операторів через шум збільшує число помилок при роботі, призводить до виникнення травм. Для оператора ПЕОМ джерелом шуму є робота принтера. Щоб усунути це джерело шуму, використовують наступні методи. При покупці принтера слід вибрати найбільш шумозахисні матричні принтери або з великою швидкістю роботи (струменеві, лазерні). Рекомендується принтер поміщати в найбільш віддалене місце від персоналу, або застосувати звукоізоляцію та звукопоглинання (під принтер підкладають демпфуючі підкладки з пористих звукопоглинальних матеріалів з листів тонкої повсті, поролону, пеноплону).

При роботі на ПЕОМ, проектом передбачені наступні методи захисту від електромагнітного випромінювання: обмеження часом, відстанню, властивостями екрану.

Обмеження годині роботи на ПЕОМ складає 3,5-4,5 години. Захист відстанню передбачає розміщення монітора на відстані 0,4-0,5 м від оператора. Передбачений монітор 20" TFT, Samsung 2043BW відповідає вимогам стандарту ТСО'03.

ТСО'03 пред'являє жорсткі вимоги в таких областях: ергономіка (фізична, візуальна і зручність користування), енергія, випромінювання (електричних і магнітних полів), навколишнє середовище і екологія, а також пожежна та електрична безпека, які відповідають всім вимогам [13].

Для зниження стомлюваності та підвищення продуктивності праці обслуговуючого персоналу в колірній композиції інтер'єру приміщень для ПЕОМ дипломним проектом пропонується використовувати спокійні колірні

поєднання і покриття, що не дають відблисків.

У проекті передбачається використання сумісного освітлення. У світлий час доби приміщення освітлюватиметься через віконні отвори, в решту часу використовуватиметься штучне освітлення.

Як штучне освітлення необхідно використовувати штучне робоче загальне освітлення. Для загального освітлення необхідно використовувати люмінесцентні лампи. Вони володіють наступними перевагами: високою світловою віддачею, тривалим терміном служби, хоча мають і недоліки: високу пульсацію світлового потоку.

При експлуатації ПЕОМ виробляється зорова робота. Відповідно до [27] ця робота відноситься до розряду 5а. При цьому нормоване освітлення на робочому місці(E_n) при загальному освітленні рівна 200 лк.

Приміщення завдовжки 12 м, шириною 10 м, заввишки 4 м обладнується світильниками типу ЛП02П, оснащеними лампами типу ЛБ зі світловим потоком 3120 лм кожна.

Виконаємо розрахунок кількості світильників в робочому приміщенні завдовжки $a=12$ м, шириною $b=10$ м, заввишки $z=4$ м, використовуючи формулу (5.5) розрахунку штучного освітлення при горизонтальній робочій поверхні методом світлового потоку:

$$n = (E \cdot S \cdot Z \cdot k) / (F \cdot U \cdot M), \quad (5.5)$$

де F - світловий потік = 3120 лм;

E - максимально допустима освітленість робочих поверхонь = 200 лк;

S - площа підлоги = 120 м²;

Z - поправочний коефіцієнт світильника = 1,2;

k - коефіцієнт запасу, що враховує зниження освітленості в процесі експлуатації світильників = 1,5;

n - кількість світильників;

U - коефіцієнт використання освітлювальної установки = 0,6;

M - кількість ламп у світильнику =2.

З формули (5.5) виразимо n (5.6) і визначимо кількість світильників для даного приміщення:

$$n = (E \cdot S \cdot Z \cdot k) / (F \cdot U \cdot M), \quad (5.6)$$

Отже, $n = (200 \cdot 120 \cdot 1,2 \cdot 1,5) / (3120 \cdot 0,6 \cdot 2) = 12$.

Виходячи з цього, рекомендується використовувати 12 світильників. Світильники слід розміщувати рядами, бажано паралельно стіні з вікнами. Схема розташування світильників зображена на рис. 5.1.

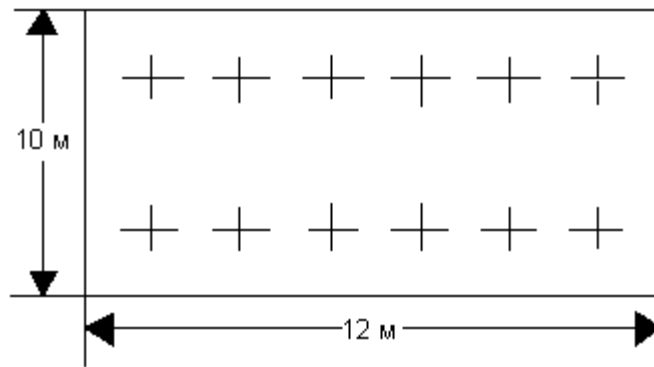


Рисунок 5.1 - Схема розташування світильників

5.4 Рекомендації по пожежній безпеці

Пожежі в приміщеннях, де встановлена обчислювальна техніка, представляють небезпеку для життя людини. Пожежі також пов'язані як з матеріальними втратами, так і з відмовою засобів обчислювальної техніки, що у свою чергу спричиняє за собою порушення ходу технологічного процесу.

Пожежа може виникнути при наявності горючої речовини та внесення джерела запалювання в горюче середовище. Пальними матеріалами в приміщеннях, де розташовані ПЕОМ, є:

- поліамід - матеріал корпусу мікросхеми, горюча речовина, температура самозаймання аерогелю 420 °С ;
- полівінілхлорид - ізоляційний матеріал, горюча речовина, температура запалювання 335 °С, температура самозаймання 530 °С, кількість енергії, що виділяється при згоранні - 18000 - 20700 кДж/кг;
- стеклотекстоліт ДЦ - матеріал друкарських плат, важкозаймистий матеріал, показник горючості 1.74, не схильний до температурного самозаймання;
- пластика кабельний №489 - матеріал ізоляції кабелю, горючий матеріал, показник горючості більш 2.1;
- деревина - будівельний і обробний матеріал, матеріал з якого виготовлені меблі, горючий матеріал, показник горючості більше 2.1, теплота згорання 18731 - 20853 кДж/кг, температура запалювання 399 °С, схильна до самозаймання.

Згідно [29] приміщення відносяться до категорії В (пожежовибухонебезпечним) і згідно правилам побудови електроустановок простір усередині приміщення відноситься до вогнебезпечної зони класу П - Па (зони, розташовані в приміщеннях, в яких зберігаються тверді горючі речовини).

Потенційними джерелами запалення при роботі ПЕОМ є:

- іскри при замиканні і розмиканні ланцюгів;
- іскри і дуги коротких замикань;
- перегріву від тривалого перевантаження і наявності перехідного опору.

Продуктами згорання, що виділяються при пожежі, є : оксид вуглецю, сірчистий газ, оксид азоту, синильна кислота, акролеїн, фосген, хлор та ін. При горінні пластмас, окрім звичайних продуктів згорання, виділяються різні продукти термічного розкладання: хлорангідридні кислоти, формальдегіди, хлористий водень, фосген, синильна кислота, аміак, фенол, ацетон, стирол та ін., що шкідливо впливають на організм людини.

Для захисту персоналу від дії небезпечних і шкідливих чинників пожежі проектом передбачається застосування промислового протигаза з коробкою марки В(жовта).

Пожежна безпека об'єктів народного господарства регламентується [24] і забезпечується системами запобігання пожежам і протипожежному захисту. Для успішного гасіння пожеж вирішальне значення має швидке виявлення пожежі і своєчасний виклик пожежних підрозділів до місця пожежі.

Зменшити горюче навантаження не представляється можливим, тому проектом передбачається застосувати наступні способи і їх комбінації для запобігання утворенню(внесення) джерел запалення :

- застосування устаткування, що задовольняє вимогам електростатичної безпеки;
- застосування в конструкції швидкодіючих засобів захисного відключення можливих джерел запалення;
- виключення можливості появи іскрового заряду статичної електрики в горючому середовищі з енергією, рівної і вище мінімальної енергії запалення;

- підтримка температури нагріву поверхні машин, механізмів, устаткування, пристроїв, речовин і матеріалів, які можуть увійти до контакту з палим середовищем, нижче гранично допустимої, становить 80% якнайменшої температури самозаймання пального.

- заміна небезпечних технологічних операцій більш безпечними;
- ізольоване розташування небезпечних технологічних установок і устаткування;

- зменшення кількості палих і вибухонебезпечних речовин, що знаходяться у виробничих приміщеннях;

- запобігання можливості утворення палих сумішей на лінії, вентиляційних системах і ін.;

- механізація, автоматизація та справність(потокова) виробництва;

- суворе дотримання стандартів і точне виконання встановленого технологічного режиму;

- запобігання можливості появи в небезпечних місцях джерел запалення;

- запобігання розповсюдженню пожеж і вибухів;

- використання устаткування і пристроїв, при роботі яких не виникає джерел запалення;

- виконання вимог сумісного зберігання речовин і матеріалів;

- наявність громовідводу;

- ліквідація можливості самозаймання речовин і матеріалів .

Для запобігання пожежі в обчислювальних центрах проектом пропонується виконання наступних вимог :

- електроживлення ЕОМ повинно мати автоматичне блокування відключення електроенергії на випадок зупинки системи охолодження і кондиціонування;

– система вентиляції обчислювальних центрів повинна бути обладнана блокуючими пристроями, що забезпечують її відключення на випадок пожежі;

– робочі місця повинні бути оснащені пожежними щитами, сигналізацією, засобами для сповіщення про пожежну небезпеку (телефонами), медичними аптечками для надання першої медичної допомоги, розробленим планом евакуації.

Для зниження пожежної небезпеки в приміщеннях використовуються первинні засоби гасіння пожеж, а також система автоматичної пожежної сигналізації, яка дозволяє знайти початкову стадію загоряння, швидко і точно оповістити службу пожежної охорони про час і місце виникнення пожежі.

Відповідно до правил пожежної безпеки для промислових підприємств приміщення категорії В підлягають устаткуванню системами автоматичної пожежної сигналізації. Проектом передбачається застосування датчика типу ІДФ - 1(димовий фотоелектричний датчик), оскільки специфікою пожеж обчислювальної техніки і радіоапаратури є, в першу чергу, виділення диму, а потім - підвищення температури.

При виникненні пожежі в робочому приміщенні обслуговуючий персонал зобов'язаний негайно вжити заходи по ліквідації пожежі. Для ліквідації пожежі використовують вогнегасники (хімічно-пінні, пінні для повітря ОП-5, ОП-6, ОП-9, вуглекислотні ОУ-5), пісок, пожежний інвентар (сокири, ломи, багри, шерстяну або азбестову ковдри). Як засіб індивідуального захисту проектом передбачається використання промислового протигаза з маскою, фільтруючої коробки В.

В якості організаційно-технічних заходів рекомендується проводити навчання робочого персоналу правилам пожежної безпеки.

ВИСНОВКИ

Дана дипломна робота присвячена розробці алгоритмів і програми хмарного сервісу для розпізнавання тексту в потоці відеоданих. У результаті виконання випускної роботи були розроблені й програмно реалізовані алгоритми фільтрації тексту, нейронних мереж й проведені їхні практичні дослідження. Програма реалізована мовою програмування C++ та Java.

Для досягнення поставленої мети, були вирішені наступні завдання:

а) розроблений алгоритм виявлення та розпізнавання текстів у потоці відеоданих на основі SWT перетворення;

б) був проведений аналіз фільтрів зображення, в результаті чого, для виділення текстової інформації на зображенні, в розробленому алгоритмі використовується фільтр Гауса;

в) було досліджено детектор Канні та оператор Собеля;

г) реалізований хмарний сервіс;

д) задача розпізнавання символів була розв'язана шляхом використання одношарової нейронної мережі з вчителем, так як переваги використання алгоритмів, що базуються на роботі нейронних мереж перед іншими алгоритмами обробки растра містять у собі велику стійкість до зміни ознакового простору в межах кадру, меншу вимогливість до стандартизації зображення, легку відтворюваність та можливість навчання у випадку чи зміни ускладнення задач.

Практичні дослідження випускної роботи показали що алгоритм пошуку тексту на основі SWT перетворення може бути використаний у реальних умовах так як має високий коефіцієнт розпізнавання тексту навіть в умовах 50% зашумленості зображення.

У розділі «Охорона праці» виконано аналіз потенційних небезпек при роботі із засобами обчислювальної техніки і механізмами, розроблені заходи щодо техніки безпеки, заходи, які забезпечують виробничу санітарію і гігієну

праці, розраховане штучне освітлення, виконані рекомендації по пожежній безпеці.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

- 1) Шапиро, Л. Компьютерное зрение [Текст] / Л. Шапиро, Дж. Стокман; Пер. с англ. – М.: Бином. Лаборатория знаний, 2006. – 752 с.
- 2) Обробка відео потоку в реальному часі: накладання фільтрів [Електронний ресурс] – Режим доступу: <http://fpga.in.ua/fpga/examples/obrobka-video-potoku-v-realnomu-chasi-nakladannya-filtriv.html>.
- 3) Путятін Є.П. Методи та алгоритми комп'ютерного зору: навчальний посібник [Текст] / Є.П. Путятін, В.О. Гороховатський, О.О. Матаг. – Х.: ТОВ «Компанія СМІТ», 2006. – 236 с.
- 4) Афонасенко, А.В. Обзор методов распознавания структурированных символов [Текст] / А.В. Афонасенко, А.И. Елизаров / Доклады Томского государственного университета систем управления и радиоэлектроники. 2008. Вып. 2(18). Ч.1. С.83-88.
- 5) Николенко А.А. Обнаружение текстовых областей и выделение символов на изображениях с неоднородным фоном [Електронний ресурс] / А.А. Николенко, А.В. Котляр, Т. Нгуен / Праці Одеського політехнічного університету, 2013. – № 1. – С. 55–60. – Режим доступу: http://nbuv.gov.ua/j-pdf/Poru_2013_1_10.pdf.
- 6) Аунг Ч.Х. Разработка алгоритмов обработки изображений интеллектуальными мобильными роботами на основе нечеткой логики и нейронных сетей. [Електронний ресурс] / Ч.Х. Аунг, З.П. Тант, А.Р. Федоров, П.А. Федоров / Современные проблемы науки и образования, 2014. – № 6. Режим доступу: www.science-education.ru/120-15579 – 10.06.2015.
- 7) Гонсалес, Р. Цифровая обработка изображений [Текст] / Р.Гонсалес, Р. Вудс; пер. с англ. – М.: Техносфера, 2005. – 1072 с.
- 8) Детектор границ Канни [Електронний ресурс] – Режим доступу: <http://habrahabr.ru/post/114589/>.

9) Detecting Text in Natural Scenes with Stroke Width Transform [Електронний ресурс] / Microsoft Corporation – Режим доступу: <http://research.microsoft.com/pubs/149305/1509.pdf>.

10) Форсайт, Д., Компьютерное зрение. Современный подход [Текст] / Д.Форсайт, Ж.Понс; пер. с англ. – М.: Издательский дом «Вильямс», 2004.– 928 с.

11) Фисенко В.Т., Фисенко Т.Ю., Компьютерная обработка и распознавание изображений: учеб. пособие. [Текст] / В.Т. Фисенко, Т.Ю. Фисенко – СПб: СПбГУ ИТМО, 2008. – 192 с.

12) Журавлев Ю.И. Распознавание. Математические методы. Программная система. Практические применения. [Текст] / Ю.И. Журавлев, В.В. Рязанов, О.В. Сенько – М.: Фазис, 2005.

13) Хайкин С. Нейронные сети: полный курс – 2-е изд. [Текст] / С. Хайкин; пер. с .англ. – М.: «Вильямс», 2006. – 1104 с.

14) Перцептрони [Електронний ресурс] – Режим доступу: http://victoria.lviv.ua/html/wosserman/rozdil2.htm#r2_3.

15) Системы распознавания образов (идентификации) [Електронний ресурс] – Режим доступу: http://www.codenet.ru/progr/alg/ai/htm/gl3_1.php.

16) Нейронні мережі: Означення та основні властивості [Електронний ресурс] – Режим доступу: <http://novotarskiy.narod.ru/cgi-bin/mono2/2.pdf>.

17) Основы штучних нейронних мереж [Електронний ресурс] – Режим доступу: <http://victoria.lviv.ua/html/wosserman/rozdil1.htm>.

18) Використання нейронних мереж для розпізнавання символів [Електронний ресурс] – Режим доступу: <http://dspace.nbuiv.gov.ua/bitstream/handle/123456789/50791/09-Savchuk.pdf?sequence=1>.

19) Розенблатт, Ф. Принципы нейродинамики (персептрон и теория механизмов мозга) [Текст]/ Ф. Розенблатт. –М.: «Мир», 1965. – 480с.

20) ГОСТ 12.1.005-88. Міждержавний стандарт. Система стандартів безпеки праці. Загальні санітарно-гігієнічні вимоги до повітря робочої зони

- 21) ГОСТ 12.0.003-74 Небезпечні і шкідливі виробничі фактори.
Класифікація
- 22) ДСТУ 7237:2011 Національний стандарт України. Система стандартів безпеки праці. Електробезпека. Загальні вимоги та номенклатура видів захисту
- 23) ДСанПіН 3.3.2.007-98. Державні санітарні правила і норми. Гігієнічні вимоги до організації роботи з візуальними дисплейними терміналами електронно-обчислювальних машин.
- 24) ГОСТ 12.1.004-91. Пожежна безпека. Загальні вимоги .
- 25) ДБН В.2.5-67. Опалення вентиляція та кондиціонування.
- 26) ГОСТ 12.1.006-84. Електромагнітні поля радіочастот. Допустимі рівні на робочих місцях і вимоги до проведення контролю
- 27) ДБН В.2.5-28-2006. Природне і штучне освітлення.
- 28) ГОСТ 12.4.009-83. Пожежна техніка для захисту об'єктів. Основні види. Розміщення і обслуговування.
- 29) ДСТУ Б В.1.1-36-2016. Визначення категорії приміщень, будинків та зовнішніх установок за вибухопожежною та пожежною безпекою.
- 30) ДСП 173-96. Державні санітарні правила планування та забудови населених пунктів

ДОДАТОК А.

Вихідний код класу, що описує нейрон

```

#pragma once
#include "networklib.h"
#include <vector>
#include <boost/function.hpp>

class Neuron {
public:
    Neuron(size_t numInputs);
    void Reset(boost::function<float ()> randFunc);
    std::vector<float>& weights() { return m_weights; }
    float GetDelta() const { return m_delta; }
    void SetDelta(float delta) { m_delta = 0; }
    float Compute(std::vector<float>& inputs);
    void IncWeights(std::vector<float>& inputs);
    void DecWeights(std::vector<float>& inputs);

private:
    std::vector<float> m_weights;
    float m_delta;
};

#include "Neuron.h"

Neuron::Neuron(size_t numInputs)
: m_delta(-1)
{
    this->m_weights.resize(numInputs);
    for (size_t i=0; i<numInputs; i++) {
        this->m_weights[i] = 0;
    }
}

void Neuron::Reset(boost::function<float ()> randFunc)
{
    for (size_t i=0; i < this->m_weights.size(); i++) {
        this->m_weights[i] = randFunc();
    }
}

float Neuron::Compute(std::vector<float>& inputs)
{
    if (this->m_delta < 0) {
        return -1;
    }

    float res = 0;

```

```

        for (size_t i=0; i < inputs.size(); i++) {
            res += inputs[i]*this->m_weights[i];
        }

        return this->m_delta <= res ? res : -1;
    }

void Neuron::IncWeights(std::vector<float>& inputs)
{
    float tmpDelta = 0;
    for (size_t i=0; i < inputs.size(); i++) {
        this->m_weights[i] += inputs[i];
        tmpDelta += inputs[i];
    }

    tmpDelta = tmpDelta * 0.7F; // 80% from the max signal

    if (!tmpDelta) {
        return;
    } else if (this->m_delta < 0) {
        this->m_delta = tmpDelta;
    } else {
        this->m_delta = min(this->m_delta, tmpDelta);
    }
}

void Neuron::DecWeights(std::vector<float>& inputs)
{
    for (size_t i=0; i < inputs.size(); i++) {
        this->m_weights[i] -= inputs[i];
    }
}

```

Вихідний код класу, що описує нейронну мережу

```

#pragma once
#include "networklib.h"
#include <vector>
#include <string>
#include <boost/shared_ptr.hpp>

class NetAlgorithm;

class NeuronNet {
public:
    NeuronNet();

    void Init(size_t numNeurons, size_t numInputs);
    /**
     * Resets the weights of the neurons to the random values
     */
    void Reset();
    void Study(TTrainSignals& signals);
}

```

```

    TResult Test(TSignal& inputSignal);
    void SaveState(const std::string& fileName);
    bool LoadState(const std::string& fileName);

    size_t GetNeuronsCount() const { return this-
>m_neurons.size(); }

private:
    TNeurons m_neurons;
    size_t m_inputsCount;
    TAlgorithm m_algorithm;
};

#include "NeuroNet.h"
#include "Neuron.h"
#include <math.h>
#include <iostream>
#include <sstream>
#include <boost/foreach.hpp>
#include <boost/archive/text_oarchive.hpp>
#include <boost/archive/text_iarchive.hpp>
#include "TeacherBasedAlgorithm.h"
#include "easylogging++.h"

NeuronNet::NeuronNet()
: m_inputsCount(0)
, m_algorithm(new TeacherBasedAlgorithm())
{
}

void NeuronNet::Init(size_t numNeurons, size_t numInputs)
{
    this->m_neurons.clear();
    for(size_t index=0; index<numNeurons; index++) {
        this->m_neurons.push_back(TNeuron(new
Neuron(numInputs)));
    }
    this->m_inputsCount = numInputs;
}

void NeuronNet::Reset()
{
    this->m_algorithm->Reset(this->m_neurons);
}

void NeuronNet::Study(TTrainSignals& signals)
{
    this->m_algorithm->Study(signals, this->m_neurons);
}

TResult NeuronNet::Test(TSignal& inputSignal)
{

```

```

    return this->m_algorithm->Test(inputSignal, this-
>m_neurons);
}

void NeuronNet::SaveState(const std::string& fileName)
{
    try {
        std::ofstream ofs(fileName);

        // save data to archive
        {
            boost::archive::text_oarchive oa(ofs);

            size_t numNeurons = this->m_neurons.size();
            oa << numNeurons;
            oa << this->m_inputsCount;

            std::for_each(this->m_neurons.begin(), this-
>m_neurons.end(), [&](TNeuron neuron) {
                float delta = neuron->GetDelta();
                oa << delta;

                std::for_each(neuron->weights().begin(), neuron-
>weights().end(), [&](float wi) {
                    oa << wi;
                });
            });
        }
    } catch(std::exception& ex) {
        LOG(ERROR) << "Can't save state! Error occurred: " <<
ex.what();
    }
}

bool NeuronNet::LoadState(const std::string& fileName)
{
    try {
        std::ifstream ifs(fileName);

        // save data to archive
        {
            boost::archive::text_iarchive ia(ifs);
            size_t numNeurons;
            ia >> numNeurons;
            size_t numInputs;
            ia >> numInputs;

            this->Init(numNeurons, numInputs);
            float wi;

            for (size_t numN=0; numN<numNeurons; numN++) {
                float delta;
                ia >> delta;
            }
        }
    }
}

```

```

        this->m_neurons[numN]->SetDelta(delta);
        for (size_t indexW=0; indexW<numInputs;
indexW++) {
            ia >> wi;
            this->m_neurons[numN]->weights()[indexW] =
wi;
        }
    }
} catch (std::exception& ex) {
    LOG(ERROR) << "Can't load state! Error occurred: " <<
ex.what();
}
return true;
}

```

Вихідний клас, що описує алгоритм навчання мережі за допомогою вчителя

```

#pragma once
#include "NetAlgorithm.h"

class TeacherBasedAlgorithm : public virtual NetAlgorithm {
public:
    virtual void Study(TTrainSignals& signals, TNeurons&
neurons);
    virtual TResult Test(TSignal& inputSignal, TNeurons&
neurons);
    virtual void Reset(TNeurons& neurons);
    TResult FindWinner(TNeurons& neurons, TSignal& inputSignal);
};

#include "TeacherBasedAlgorithm.h"
#include "Neuron.h"
#include <easylogging++.h>

void TeacherBasedAlgorithm::Reset(TNeurons& neurons)
{
    std::for_each(neurons.begin(), neurons.end(), [](TNeuron&
neuron) {
        neuron->Reset([&] { return (float)0;});
    });
}

TReturn TeacherBasedAlgorithm::FindWinner(TNeurons& neurons,
TSignal& inputSignal)
{
    float power = -1;
    float tmpPower = -1;
    TResult index = -1;

    for (size_t nNum = 0; nNum < neurons.size(); nNum++) {

```



```

        tmpPower = neurons[nNum]->Compute(inputSignal);
        if (tmpPower < 0) {
            continue;
        } else if (tmpPower > power) {
            power = tmpPower;
            index = nNum;
        }
    }

    return index;
}

void PrintVectors(std::vector<float>& v1, std::vector<float>&
v2) {
    auto print = [] (std::vector<float>& v) {
        std::stringstream str;
        for_each(v.begin(), v.end(), [&](float elem) {
            str << elem << " ";
        });
        LOG(INFO) << str.str().c_str();
    };

    print(v1);
    print(v2);
}

void TeacherBasedAlgorithm::Study(TTrainSignals& signals,
TNeurons& neurons)
{
    bool done = true;
    TResult winnerIndex = 0;

    int k=0;

    do {
        done = true;

        LOG(INFO) << "Studying step: " << k;

        for (size_t i=0; i < signals.size(); i++) {
            TTrainSignal signal = signals[i];

            if (signal->correctAnswer < 0 || signal-
>correctAnswer > (TResult)neurons.size()) {
                LOG(ERROR) << "wrong correctAnswer provided";
                continue;
            }

            winnerIndex = this->FindWinner(neurons, signal-
>signal);

            if (winnerIndex > UNKNOWN_RESULT) {

```

```

        if (winnerIndex != signal->correctAnswer) { //
this is wrong result
            LOG(INFO) << "    wrong result for signal: "
<< signal->correctAnswer << " neuron: " << winnerIndex;
            neurons[winnerIndex]->DecWeights(signal-
>signal);
            neurons[signal->correctAnswer]-
>IncWeights(signal->signal);
            done = false;
        }
        } else {
            LOG(INFO) << "    winner not wound, image: " <<
signal->correctAnswer;
            neurons[signal->correctAnswer]-
>IncWeights(signal->signal);
            done = false;
        }
    }

    LOG(INFO) << "Studying step: " << k << " finished,
result: " << done;
    k++;
} while (!done);
}

TResult TeacherBasedAlgorithm::Test(TSignal& inputSignal,
TNeurons& neurons)
{
    return this->FindWinner(neurons, inputSignal);
}

```

ДОДАТОК Б. Електронні плакати

Міністерство освіти і науки України
Східноукраїнський національний університет ім. В. Даля
Кафедра комп'ютерних наук та інженерії

ДИПЛОМНА РОБОТА БАКАЛАВРА

Інформаційна система виявлення та розпізнавання тексту з відеопотоку

Науковий керівник:
ст. викладач Барбарук Л.В.

Виконала:
студент групи KI-14з Грошкова О.О.

Сєвєродонецьк 2018

Актуальність проблеми



Робототехніка



Реклама



Штрих-коди продуктів



Автомобільні номери



Дорожні знаки

(2)

Постановка задачі

Мета роботи – розробити алгоритм та програму хмарного сервісу для розпізнавання текстів в відеопотоці.

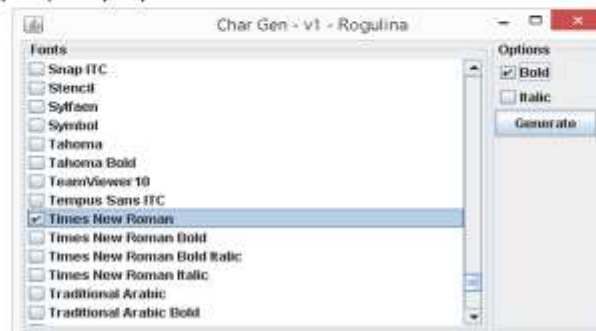
Для досягнення поставленої мети, необхідно вирішити наступні завдання:

- підготувати набір тестових зображень та реалізувати програму для генерації еталонних зображень для різних варіацій шрифтів;
- реалізувати алгоритм виявлення тексту на зображенні;
- реалізувати алгоритм розпізнавання тексту шляхом використання нейронної мережі;
- програмно реалізувати хмарний сервіс.

3

Підготовка тестового матеріалу

- пошук та збір зображень, що мають текстову інформацію (зображення плакатів, дорожніх знаків, автомобільних номерів та рекламної інформації);
- розробка програми генерації зображень літер для різних варіацій шрифтів



4

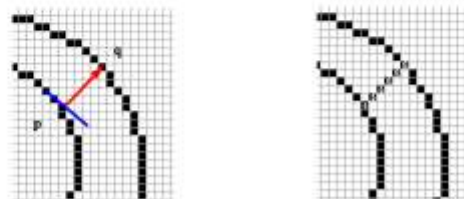
Алгоритм виявлення та розпізнавання тексту



5

Алгоритм Stroke width Transform перетворення

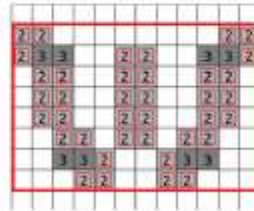
- для кожного пікселя p , що належить лінії контуру, будуємо промінь перпендикулярно напрямку градієнту, доки не зустрінемо інший піксель q на лінії контуру
- обчислюємо можливу ширину штриха w , шляхом знаходження відстані між p та q , та позначаємо пікселі SWT зображення відповідним значенням



6

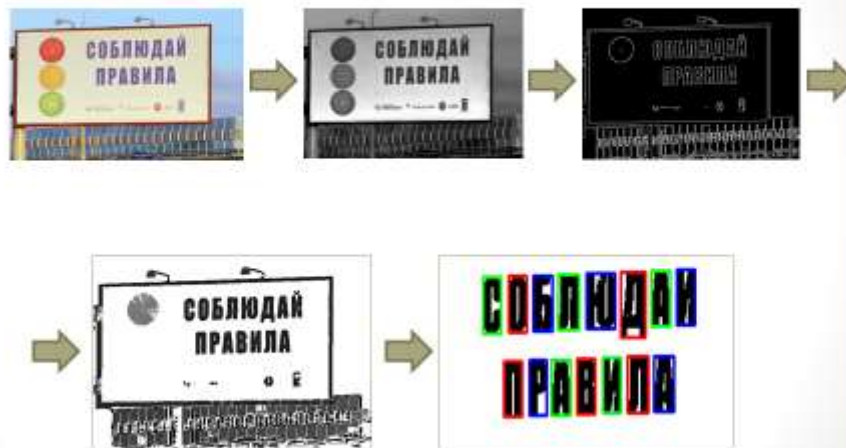
Алгоритм фільтрації літер

- * групуємо пікселі SWT зображення, що мають однакові значення
- * обчислюємо розміри отриманого об'єкту
- * порівнюємо розміри з пороговими значеннями `minSize` та `maxSize`



(7)

Аналіз результатів



(8)

Результат роботи програми



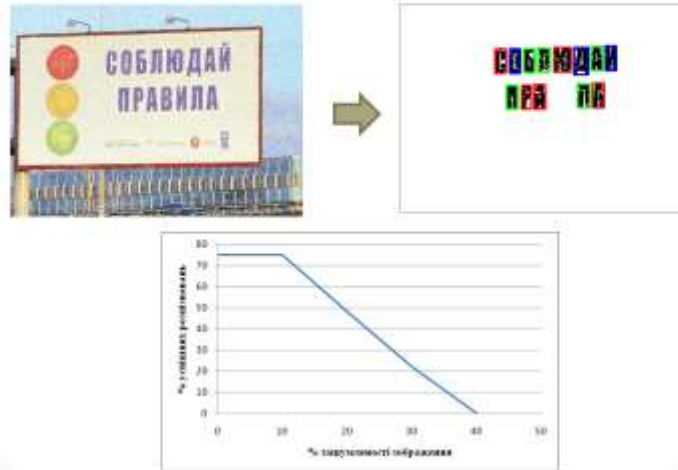
9

Архітектура хмарного сервісу



10

Дослідження впливу шуму Гауса на якість роботи алгоритму



[11]

Висновки

- Реалізовано та досліджено алгоритм виявлення тексту на основі SWT перетворення;
- реалізовано одношарову нейронну мережу з вчителем;
- розроблено хмарний сервіс.

[12]

Доповідь закінчена, дякую за увагу