

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Методи аналізу та класифікації зображень з
використанням гібридного класифікатора

(тема)

Виконав:

студент II курсу, групи СПМ-22-5
Кайда В.В.
(прізвище, ініціали)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва спеціальності)

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

Освітня програма Системне програмування
(повна назва освітньої програми)

Керівник: проф. Міхаль О.П.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет _____ комп'ютерної інженерії та управління _____

Кафедра _____ електронних обчислювальних машин _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 123 «Комп'ютерна інженерія» _____
(код і повна назва)

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

Освітня програма _____ Системне програмування _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту _____ Кайді Валерії Віталіївни _____
(прізвище, ім'я, по батькові)

1. Тема роботи Методи аналізу та класифікації зображень з використанням гібридного класифікатору

затверджена наказом по університету від “ 01 ” квітня 2024 р. № 257 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 15 червня 2024 р.

3. Вхідні дані до роботи _____
гібридний класифікатор

_____ обробка даних

_____ класифікація даних

4. Перелік питань, що потрібно опрацювати у роботі _____

_____ Аналіз предметної області та постановка завдання

_____ Методи класифікації даних з використанням машинного навчання

_____ Програмна реалізація

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 18 слайдів

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

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

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання та аналіз літератури	01.04.2024 – 06.04.2024	
2	Огляд існуючих рішень та методів	07.04.2024 – 12.04.2024	
3	Розробка моделі	13.04.2024 – 18.04.2024	
4	Вибір програмних засобів	19.04.2024 – 25.04.2024	
5	Програмна реалізація	26.04.2024 – 02.05.2024	
6	Аналіз отриманих результатів	03.05.2024 – 16.05.2024	
7	Оформлення записки	17.05.2024 – 14.06.2024	
8	Представлення роботи в ЕК	15.06.2024 – 15.06.2024	

Дата видачі завдання 01 квітня 2024 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис)

проф. Міхаль О.П.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 69 с., 32 рис., 3 табл., 2 дод., 10 джерел.

МАШИННЕ НАВЧАННЯ, КОМП'ЮТЕРНИЙ ЗІР, КЛАСИФІКАЦІЯ, НЕОДНОРІДНІСТЬ, НАВЧАННЯ, КЛАСИФІКАТОР.

Метою кваліфікаційної роботи є дослідження методів аналізу та класифікації зображень з використанням машинного навчання.

У ході виконання кваліфікаційної роботи досліджено методи аналізу та класифікації зображень з використанням комп'ютерного зору. Розроблено гібридний класифікатор, заснований на двох методах аналізу зображень для аналізу і класифікації зображень, окремих видів і підвидів з урахуванням невисловленої структури бажаного класу і неоднорідності його характеристик.

ABSTRACT

Master's thesis: 69 pages, 32 figures, 3 tables, 2 appendices, 10 sources.

MACHINE LEARNING, COMPUTER VISION, CLASSIFICATION,
HETEROGENEITY, LEARNING, CLASSIFIER.

The major goal of this thesis is to study the methods of analysis and classification based on machine knowledge.

In order to the qualification work methods of analysis and classification were studied using computer data. A hybrid classifier has been developed, based on two methods of analysis, to analyze and classify images of different species and subspecies based on the unknown structure of the fruit class and the heterogeneity of its characteristics.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	7
ВСТУП	8
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАВДАННЯ.....	11
1.1 Сучасні профорієнтаційні методики як система підтримки та прийняття рішень	11
1.2 Супровід процесу прийняття рішення	14
1.3 Застосування системного підходу до профорієнтаційної СППР	19
2 МЕТОДИ КЛАСИФІКАЦІЇ ДАНИХ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ	22
2.1 Комп'ютерний зір	22
2.2 Ознаки класів об'єктів	24
2.3 Аналіз та класифікація зображень методами машинного навчання.....	30
3 ПРОГРАМНА РЕАЛІЗАЦІЯ.....	44
3.1 Проведення експериментів	44
3.2 Гібридний класифікатор.....	48
3.3 Аналіз отриманих результатів	51
ВИСНОВКИ.....	55
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	56
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	58
ДОДАТОК Б Апробація	68

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

ГНМ – глибока нейронна мережа

КЗ – комп'ютерний зір

МН – машинне навчання

ПЗ – програмні засоби

ШНМ – штучна нейронна мережа

ВСТУП

Важко переоцінити роль процесу автоматизації в будь-якій сфері діяльності сучасної людини. Покликана виключити, або істотно знизити необхідність участі людини в ряді трудомістких операцій, автоматизація стає невід'ємною частиною різних процесів.

Так, комп'ютери підвищують ефективність виробничої сфери, в якій відбувається повсюдна автоматизація рутинних операцій та заміщення ручної праці промисловими роботами. Бортові комп'ютери керують різними транспортними засобами, швидкими темпами розвивається мобільний зв'язок, створюючи нові сектори у споживчій сфері та забезпечуючи істотне підвищення якості життя. Завдяки сучасному ринку мобільних додатків та сервісів, через смартфон чи ПК можливо здійснювати досить складні та багатоступінчасті процеси, такі як дистанційне навчання на сертифікованих онлайн-курсах та віддалена робота.

Автоматизація та комп'ютеризація охоплюють сьогодні і таку важливу сферу діяльності людини як освіту. Для швидкої адаптації до шостому технологічному укладу, що розвивається, людині доводиться сприймати і переробляти величезну кількість інформації. Разом з підвищенням якості життя через повсюдне впровадження комп'ютерів і мобільних пристроїв настав момент, коли інформаційне навантаження почало негативно позначатися на здатності людини приймати рішення і орієнтуватися у величезному просторі можливостей. Перевтома та перенасиченість інформацією ведуть до втрати мотивації до навчання, створюючи загрозу рівню якості кадрів, що готуються вузами.

Удосконалення технологій відбувається на стику наук, тому до питання професійного відбору до адрів необхідно підходити уважніше, глибоко і всебічно.

Найбільш важлива тенденція сучасної науки – співвідношення диференціації та інтеграції наук. Сучасна наука характеризується процесами інтеграції з такими передумовами: поява суміжних дисциплін; поява міждисциплінарних досліджень; поява проблем, які є орієнтирами дослідження; поява об'єктів, які мають міждисциплінарний характер.

Так, поширення міждисциплінарних досліджень, інтегративних сфер діяльності породжує нові професії на стику дисциплін і, відповідно, нових спеціалістів, що зумовлює необхідність актуалізації класичних методик профорієнтаційної діагностики та розробки нових. Крім того, більшість методик передбачає вибір найбільш відповідної відповіді з обмеженої кількості вже закладених розробником варіантів, що, в першу чергу, покаже швидше суб'єктивні переваги випробуваного, ніж відобразить рівень реальних навичок. Також, за деякої кількості вже готових варіантів для вибору, існує можливість недостовірності результатів тестування через навмисний вибір піддослідним більш «престижних» варіантів, а не тих, які справді відповідали б його поточним навичкам і перевагам [3,4]. У короткостроковій перспективі недостатньо комплексний та актуалізований - підхід до профорієнтації призводить до переривання навчання майбутнім фахівцем, недостатнього рівня підготовки кадрів, низької мотивації та включеності до майбутньої професійної діяльності, помилок, недбалості, техногенних катастроф, обумовлених людським фактором.

Вимога підвищення точності оцінок робить актуальною задачу – системного аналізу та моделювання процесів класифікації зображень, розробки та автоматизації нових алгоритмів обробки зображень з метою уточнення та доповнення класифікації типів індивідів

Метою кваліфікаційної роботи є дослідження методів аналізу та класифікації зображень з використанням машинного навчання.

Об'єктом дослідження є функціонування методи обробки зображень.

Предметом дослідження є методи аналізу зображень, створені задля підвищення ефективності процедури класифікації.

Завдання:

- огляд та аналіз сучасного стану проблем класифікації образів, професійних орієнтаційних програм та програмних комплексів як систем підтримки прийняття рішень;

- розробка гібридного класифікатора на основі двох методів аналізу зображень для поділу типів та підтипів з урахуванням неявних структур та неоднорідності ознак шуканих класів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАВДАННЯ

У вступі були розглянуті основні недоліки класичних – профорієнтаційних опитувальників, анкет і матриць, проте процес супроводу прийняття рішень у профорієнтації не обмежується проблемами структури та можливостей самих методик.

Не менш важливими є такі аспекти як постановка задачі для випробуваного, наявність чітко позначених критеріїв оцінки та класифікації, об'єктивність експерта, ясність викладу рекомендацій, доступність методики. Паперові опитувальники можуть бути доступні і прості для застосування в якості інструментів оцінки тих чи інших особливостей випробуваного, в той час як робота з матрицями та зображеннями вручну може викликати – труднощі, особливо якщо йдеться про кольорові зображення.

Методика тестування «Колірний простір», що розглядається в рамках даного дослідження, передбачає роботу з зображеннями та кольором, припускаючи створення кожним випробуваним свого унікального цифрового «портрета».

1.1 Сучасні профорієнтаційні методики як система підтримки та прийняття рішень

Промислові та побутові роботи замінили людину в ряді випадків, проте більшість операцій, як і раніше, контролюється, або здійснюється оператором або експертом. Особливо важливою є присутність експерта всередині людино-машинних систем.

У зв'язку з обмеженнями людської здатності переробки інформації, зростанням обсягу інформації та ускладнення завдань стає актуальним – завдання з'єднання можливостей людини та ЕОМ для мінімізації помилок і суб'єктивності суджень, властивих людині, при збереженні властивих тій же

людині гнучкості, креативності та емпатичності. Останнє вкрай актуально в професіях, що допомагають, і сферах соціальної спрямованості.

Так, перші реалізації систем підтримки прийняття рішень (СППР) знайшли своє застосування у сфері бізнесу, потім почали поширюватися й інші галузі діяльності, зокрема й соціальні.

Термін «система підтримки прийняття рішень (СППР)» має кілька визначень. Для завдань профорієнтації – розпізнавання елементів та класифікації зображень – найбільш близьким можна вважати визначення СППР як комп'ютерної інформаційної системи, що використовується для - підтримки різних видів діяльності та при прийнятті рішень у ситуаціях, де неможливо або небажано цілком і повністю надавати управління автоматизованій системі [15].

Роль СППР не зводиться до того, щоб повністю замінити людину в будь-якому процесі, завдання СППР – підвищити ефективність людини, свести помилки, що провокуються людським фактором, до мінімуму. Використання СППР не спрямоване на повну автоматизацію процесу прийняття рішення, а здійснює кооперації, забезпечує взаємодію між системою та людиною в процесі прийняття рішень [16].

Чи комп'ютерні методики профорієнтаційного тестування можуть називатися СППР? Щоб відповісти на це питання, необхідно звернутися до ознак СППР та найбільш поширених категорій завдань для СППР.

До ознак СППР належать:

- концептуальні моделі;
- користувачі системи;
- розв'язувані завдання;
- які забезпечують засоби;
- області застосування.

Основні завдання СППР [16]:

- прийняття рішень (у т. ч. планування та прогнозування);
- вилучення та надання знань;

- побудова людино-машинних (діалогових) систем.

Основні завдання СППР [18]:

- оптимізація (вибір оптимального рішення з кількох можливих);
- ранжування (упорядкування можливих рішень щодо переваги/ймовірності).

Розглянуті у введенні класичні методики профорієнтаційного – моніторингу та супроводу доступні в мережі Інтернет: як на спеціалізованих сайтах, так і на тематичних ресурсах або форумах, що передбачає їхню автоматизацію як комп'ютерні програми.

Класична концептуальна модель СППР виглядає наступним чином – (рисунок 1.1):

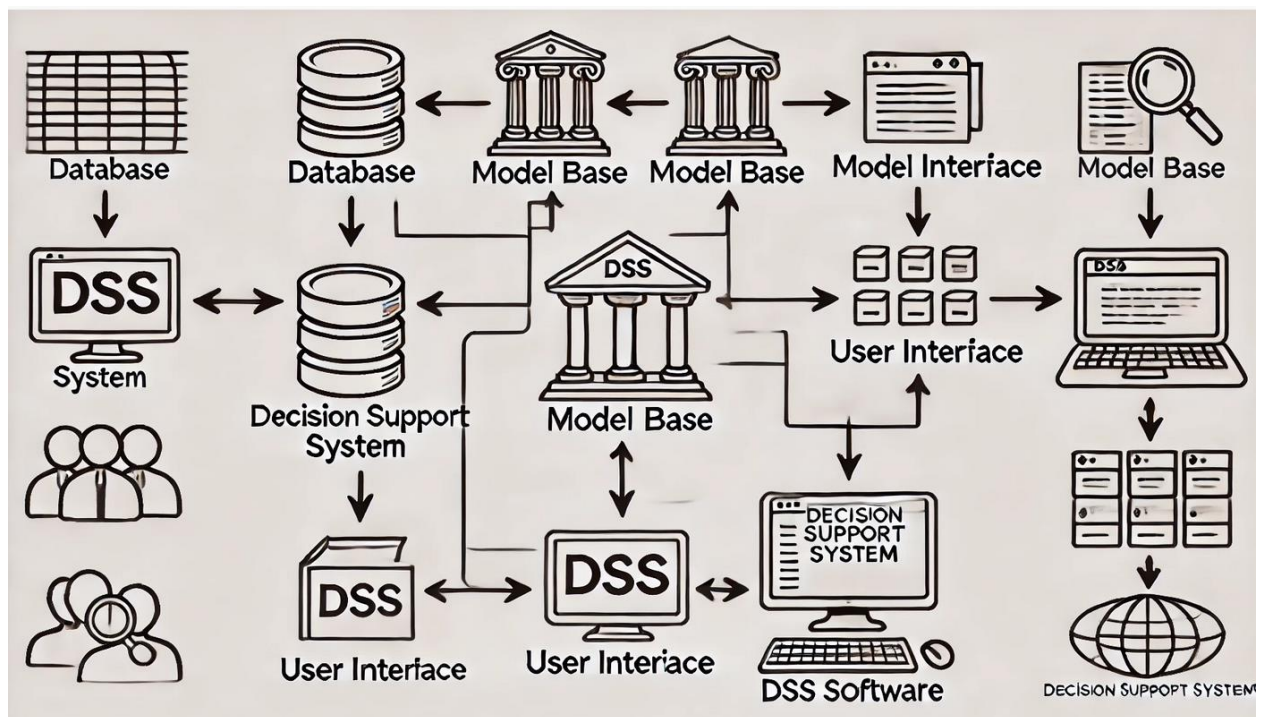


Рисунок 1.1 – Класична СППР

У процесі профорієнтаційного тестування завжди присутній користувач, який у даному сценарії є особою, яка приймає рішення (ЛІР), оскільки саме користувач, за результатами взаємодії з профорієнтаційною методикою і іноді ще й експертом, приймає рішення про вибір подальшого

напряму для самої освіти, або вступі до навчального закладу.

Оскільки в цьому розділі розглядаються автоматизовані профорієнтаційні методики, тобто методики, переведені у формат комп'ютерної програми або комплексу, то, відповідно, дані програми розташовують інтерфейсом для здійснення взаємодії системи і користувача.

Блоки аналізу проблем зазвичай містять ряд тестів, опитувальників і тренажерів, заснованих на різних моделях збору і візуалізації даних, а також алгоритми оцінки тих чи критеріїв. Блок прийняття рішень спрямований на обробку результатів діагностики. У програмному комплексі «Effecton» враховуються такі статистичні показники: обсяг даних, середня арифметична, середнє квадратичне відхилення, середня помилка, оцінка достовірності різниці між середніми арифметичними, коефіцієнт кореляції показників.

Повертаючись до ознак і завдань СПР, сучасні профорієнтаційні програми та програмні комплекси можуть бути представлені у вигляді концептуальних моделей, мають користувачами системи, мають завдання, що потребують вирішення, мають у своєму розпорядженні засоби їх вирішення (моделі, методики, алгоритми), забезпечують умови вирішення проблем (аналіз, результат, інтерпретація), застосовуються в різних сферах діяльності людини. Також програмні комплекси вирішують наступні завдання: прийняття рішень (рекомендації для ЛПР), вилучення та надання знань (інтерпретація результатів), побудова людино-машинних діалогових систем (інтерактивний інтерфейс), оптимізації (вибір найбільш відповідного рішення з усіх можливих), ранжування (оцінка можливих варіантів рішення за перевагою, ймовірності, доцільності та ін.).

1.2 Супровід процесу прийняття рішення

Як правило, при роботі з профорієнтаційними програмними комплексами до складу експертної групи, яка здійснює контроль та

підтримку прийняття рішення, входять педагоги, психологи та керівники піддослідних [19].

Крім зазначеної раніше проблеми індивідуального сприйняття кольору і форми, експертам, як і всім людям, властиві певна перевищення і рівень очікувань від результатів тестування або конкретного випробуваного, що також є однією з вагомих причин для автоматизації методики, що розглядається в рамках даного дослідження.

З метою виявлення особливостей зорового сприйняття людини, а також оцінки точності та єдності думок у класифікації тестових форм відповідно до зразків, було поставлено «Експеримент на зіставлення».

Вихідні дані експерименту:

- дві групи піддослідних – експертна (50 чол.) і експертна (50 чол.);
- виведено критерії відбору експертів та не експертів;
- підготовлені зразки (чорно-білі зображення яскраво виражених типів тестових форм, всього 4 типи);
- підібрані об'єкти дослідження для зіставлення з стандартами (кольорові зображення тестових форм, 25 шт.);
- сформульовано завдання для випробуваних;
- розроблено опитувальники для піддослідних (2 опитувальники для експертних та не експертних груп відповідно);
- сформульована гіпотеза, яка потребує підтвердження, доповнення або спростування в результаті експерименту.

Групи випробуваних . Для експерименту, відповідно до виведених критеріїв, було набрано дві групи піддослідних - експертів і експертів. Кожна група налічувала рівну кількість піддослідних – 50 чол., Розподіл за статевою ознакою становило 56% респондентів-жінок і 44% респондентів чоловіків у групі експертів; 56% респондентів-жінок та 44% респондентів чоловіків в обох групах (рисунок 1.2).

Критерії відбору групи піддослідних. Як експерти для участі в «Експерименті на зіставлення» запрошувалися випробувані, які мають –

кваліфікацію бакалавра, магістра, спеціаліста, або мають науковий ступінь у галузі мистецтва, психології, педагогіки, естетики, дизайну, архітектури, культурології, психофізіології, інформатики, а також фахівці, що є на момент участі в експерименті працевлаштованими в одній з перерахованих областей.

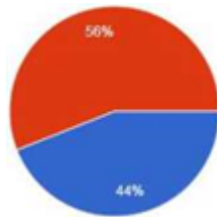


Рисунок 1.2 – Діаграма розподілу респондентів за статевою ознакою в результаті

Групам піддослідних пропонувалося 4 зразки шуканих структур (рисунок 1.3):

- тип А "лінії";
- тип В "промені";
- тип С «градієнт та лінії»;
- тип D «шахи».

Важливо, що у обох групах випробувані не навчалися оцінці тестових форм за методикою тестування, оскільки у разі навчання було складніше оцінювати особливості зорового сприйняття як експертів, і не-експертів.

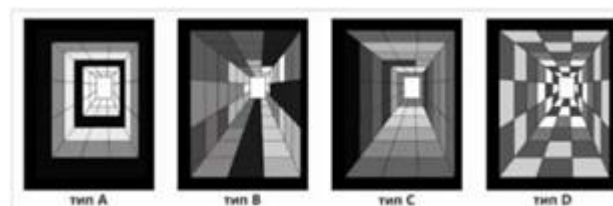


Рисунок 1.3 – Еталони структур

Еталони являли собою чорно-білі зображення, оскільки завданням піддослідних було зіставлення структур, що формуються за різними принципами групування темних і світлих секторів, які можуть мати різний колір.

Об'єктами дослідження були 25 зразків – кольорових зображень тестових форм, виконаних студентами різних років навчання. Спочатку 75% добірки (19 зображень) були змішані або неявні для класифікації варіанти зображень, тоді як 25% добірки (6 зображень) оцінювалися яскраво виражені (рисунок 1.4).



Рисунок 1.4 – Результати опитування

Завдання для досліджуваних полягало у співвіднесенні еталонів із зразками з метою класифікації кольорових зображень. У опитувальниках для груп експертів і не-експертів, складених за допомогою сервісу Google Form, завдання звучало наступним чином: «Будь ласка, уважно дивіться на еталони (для зручності роботи з тестом можна виконати print screen). На Вашу думку, структура даного кольорового зображення найбільш схожа... (далі респонденту пропонувалося кольорове зображення з виведеними варіантами

відповідей)». Вибрати можна було лише один із чотирьох представлених варіантів (рисунок 1.5).



Рисунок 1.5 – Вибір варіанта

24 питання, що залишилися, мали аналогічне уявлення. На стадії підготовки експерименту автором методики планувалося представити груп пам'яті 50 зразків для зіставлення з еталонами, проте число зразків було скорочено у зв'язку з імовірно високим ступенем втоми більшості респондентів, так як багато з них вперше брали участь у подібному експерименті.

Опитувальник розроблявся на платформі сервізу Google Forms, містив питання, що стосуються кваліфікації респондента, його статі та згоди на участь у дослідженні, а також 25 питань на зіставлення з еталонами шуканих структур (слайд з еталонами пропонувався вивести в окремому вікні за допомогою команди print screen або роздрукувати на принтері).

Опитувальники для груп експертів і неекспертів містили однакові зразки для зіставлення, відмінними були лише питання щодо кваліфікації шуканих піддослідних.

Припущення, що вимагає підтвердження, доповнення або спростування в результаті проведення експерименту, сформульовано таким чином: «В обох групах піддослідних передбачаються розбіжності в класифікації зображень загалом. У групі експертів розбіжностей передбачається істотно менше і очікується більш виражену єдність у відсотковому співвідношенні (від 60%

для одного з варіантів). У групі не-експертів очікується менш виражена єдність і більша кількість варіантів присвоєння типу для кожного зразка.

Висновок: рівень кваліфікації експерта підвищує шанси на отримання точного результату за запропонованою методикою профорієнтаційного тестування, проте, індивідуальні особливості сприйняття структури та сприйняття кольору будуть вести до помилок і різночитання, що не є прийнятним для потенційного реципієнта».

1.3 Застосування системного підходу до профорієнтаційної СППР

Основою системного підходу є цілісне бачення досліджуваних об'єктів, явищ чи процесів, являючи собою найбільш універсальний метод аналізу чи дослідження різних систем, зокрема технічних і соціальних. Іримування системного підходу в управлінні сприяє розкриттю сутності та змісту механізмів управління, а також побудові нових концепцій управління [21].

Система є сукупністю елементів, які перебувають у зв'язку між собою та об'єднані єдністю мети, а також і функціональною цілісністю, але властивість самої системи не зводиться до суми властивостей її елементів. Елементи об'єднуються в систему в результаті узгодженого взаємодії (складання зусиль), спрямованого на утворення нового, яке мало б таку якість або якості, які не були властиві елементам раніше, до об'єднання. Система повинна мати цілісність, тобто виступати як ціле щодо навколишнього середовища, ціле, здатне протистояти зовнішнім впливам за рахунок внутрішніх зв'язків між елементами. Чим стійкіше зв'язку всередині системи, тим система загалом стійкіша до впливу різних зовнішніх чинників.

Системний аналіз – наука, що займається проблемою прийняття рішень за умов аналізу великої кількості інформації різної природи. Застосування методів системного аналізу до профорієнтаційної СППР дозволяє виділити цю систему із середовища, зрозуміти, що входить у систему, як система

здатна протистояти зовнішнім фактором, що надходить у систему на вході і що вона здатна дати на виході.

Декомпозиція. Це завдання полягає у поданні системи як сукупності її підсистем, які, у свою чергу, складаються з дрібніших елементів [26]. У процесі декомпозиції насамперед визначається головна та кінцева мета функціонування системи.

Метою аналізованої системи є підвищення точності процесу профорієнтації з урахуванням ІСД, метою ж дисертаційного дослідження є - підвищення ефективності процесу класифікації зображень за наявності неявних структур та елементів, що належать до різних класів, що шукаються. Тобто мета дисертаційної роботи спрямована на підвищення точності процесу профорієнтації за рахунок застосування методів машинного аналізу та класифікації зображень до об'єктів, які отримують у процесі взаємодії людини та програмного комплексу.

Також вся аналізована система, що включає випробуваного, програмний комплекс та експерта, виробляється з навколишнього середовища, визначаються запити навколишнього середовища до системи і ті знання або результати, які система віддає навколишньому середовищу у відповідь на запит. Навколишнє середовище включає безліч процесів і явищ у всьому їхньому різноманітті. Розглянута ж система повинна відрізнятися від навколишнього середовища тим, що входять до неї елементи і підсистеми різних рівнів зв'язані і взаємодіють між собою сильніше, об'єднані тіснішими зв'язками і відносинами, ніж система з навколишнім середовищем.

Впливаючими чинниками довкілля може бути чинники раз особистої природи. Як біологічний фактор може виступати пандемія COVID-19, що негативно впливає на всі сфери діяльності людини. Крім впливу на смертність, пандемія провокує логістичні збої, ускладнює соціальні контакти, що не дає можливості випробуваним довірливо поспілкуватися з експертами і отримати допомогу в інтерпретації результатів різних профорієнтаційних тестів.

Природними факторами можуть бути несприятливі погодні умови, як політичні чинники виступають загострення міжнародних відносин, що також впливають на логістику і мобільність як експертів, так і піддослідних.

Соціальними факторами є насамперед очікування суспільства та роботодавців від кадрів, що надходять. У різні часи були потрібні різні професії, що не може не впливати як на випробуваних, так і на експертів, і на систему підтримки прийняття рішень у вигляді профорієнтаційної програми або програмного комплексу.

Серед інформаційних факторів можуть бути нові класифікації та методики, що з'явилися в загальному доступі, актуалізовані і відмінні від попередніх. У таких умовах потрібна максимально валідна класифікація усередині профорієнтаційного програмного комплексу, з можливістю її доповнення, уточнення, а головне – чітке розуміння механізмів утворення підтипів.

Аналіз. Завдання аналізу полягає у визначенні властивостей системи або середовища, в якому система функціонує.

Основне середовище, в якому функціонують профорієнтаційні програмні комплекси – соціальне, оскільки функції профорієнтації зводяться до тих, хто допомагає (супровід прийняття рішення, консультування). Але оскільки користувачі вибирають різні професії, то так чи інакше профорієнтація стосується практично кожної сфери діяльності людини.

Так, якщо з середовища надходить запит на найрізноманітніші кадри, то система повинна виділяти найбільш кращі типи для кожної професії. Це дозволяють аналоги – профорієнтаційні програми та комплекси, проте в ситуаціях, коли користувач хоче реалізуватися в одному середовищі, але має навички, необхідні для іншого, то потрібна більш «тонка» діагностика.

Так як сьогодні в науці і праці простежується тенденція до міждисциплінарності, то і в профорієнтації має бути можливість для рекомендацій людині технічного складу, що вибрала мистецтво, і людині, що має художні навички, але прагне до сфери ІТ-технологій.

2 МЕТОДИ КЛАСИФІКАЦІЇ ДАНИХ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

2.1 Комп'ютерний зір

Завдання CV (Computer Vision, комп'ютерний зір) полягають, головним чином, отримання корисної інформації (знань) у вигляді різних зображень – як статичних, і динамічних.

Серед найактуальніших завдань CV:

- калібрування камер та оптичних систем;
- визначення руху за зображеннями;
- визначення перешкод у процесі руху;
- розпізнавання об'єктів на сцені;
- просторова реконструкція сцени;
- локалізація зображення в задалегідь відомій сцені;
- аналіз відмінностей у наборі зображень.

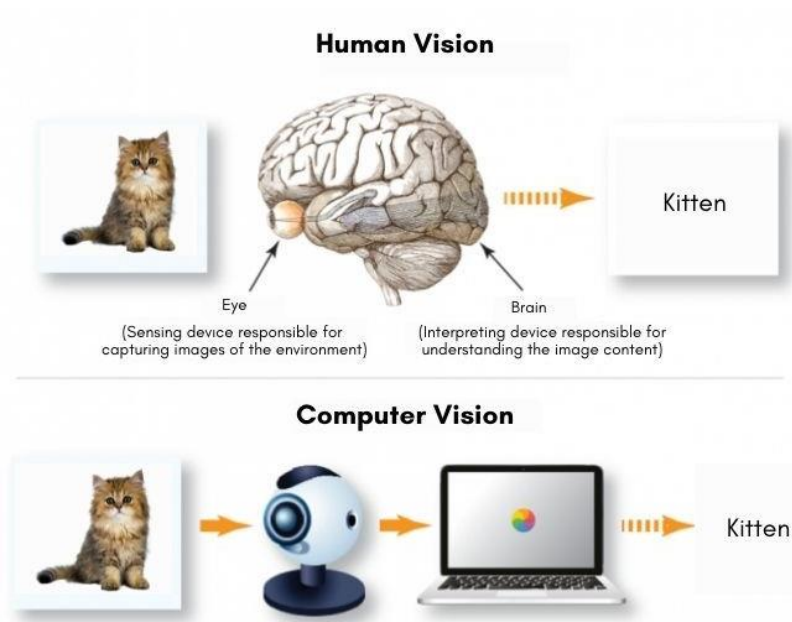


Рисунок 2.1 – Схема порівняння людського та комп'ютерного зору

В рамках цього дослідження найбільший інтерес становлять методи, придатні для вирішення задач п. 4, а саме – розпізнавання об'єктів на сцені або площині (планігоні). Схема розв'язання задач комп'ютерного зору представлена на рисунку 2.1.

У разі поняття «розпізнавання об'єктів» пов'язані з поняттям «розпізнавання образів». Методи розпізнавання образів можна поєднати у три великі групи:

- методи перебору;
- методи з використанням характеристик образів;
- методи штучних нейронних мереж.

Перш ніж конкретизувати методи всередині зазначених груп, потрібно вирішити додаткове завдання формалізації ознак, що шукаються. Завдання частково вирішується при генерації зображення користувачем, так як квазіпростір вже розділено на комірки, які мають чіткі краї, кольори оптимізованої палітри підібрані таким чином щоб відтінки, розташовані в спектрі близько відносного один одного, легко розрізнялися візуально. Так, можна виділити деякі апріорні або детерміновані знання про шукані ознаки, а саме: розміри областей визначення, діапазон розподілу деяких характеристик, еталонні приклади.

Сегментація. Якщо розглянути шуканий об'єкт як певну область інтенсивності або їх сукупність, то для їх виділення можна використовувати методи розбиття або сегментації областей. Алгоритми сегментації також є частиною формалізації ознак шуканих класів об'єктів.

Сегментація визначається як процес підрозділу текстури або зображення на складові або об'єкти. Процес сегментації входить в елементи роботи системи машинного зору, оскільки на даному етапі обробки зображення відбувається відділення шуканих об'єктів від фону з метою їх подальшого аналізу та дослідження.

У процесі сегментації поділ на елементи чи області відбувається за подібністю властивостей їх точок. Можна сказати, що при генерації

квазіпростору користувачем виникає якась «самосегментація», тому що випробовуваний довільно і мимоволі виділяє деякі області в характерні структури, в той час як «об'єкти» або «пошукані області» рідко виявляються візуально простими і зрозумілими в багатьох інших областях, де використовується комп'ютерний зір.

2.2 Ознаки класів об'єктів

Усі типи, крім F, передбачають проходження класифікації та мають виділені апріорні ознаки.

Тип А характеризується переважною (від 50% і більше, відповідно 100% – чистий тип) присутністю елементів «горизонтальні лінії» (рисунок 2.2). Лінії (від 3 осередки в довжину) можуть становити як кольори строго однакових координат, так і кольори різних координат, але відносяться до одного тону, наприклад «зеленого». Колірний тон – це особливість кольору, що відрізняє його від інших кольорів спектру, або це діапазон кольору, що позначається в колірному колі як "жовтий", "червоний", "помаранчевий" і т.д. Також саме тон визначає назву кольору.

У поточній вибірці елементи типу А переважно представлені кольорами однакових координат, передбачається, що смуг із кольорів різних координат, але одного тону буде менше.

Ще однією умовою присутності типу А є мала кількість відтінків, а саме – не більше одного тону з кількістю відтінків 5 і більше.

Ознаки домінуючого типу А:

- 50% розгортки і більше заповнені шуканими елементами «горизонтальна лінія» (від трьох осередків одного кольору або тону);
- мала кількість відтінків, а саме не більше одного тону з кількістю відтінків 5 і більше.

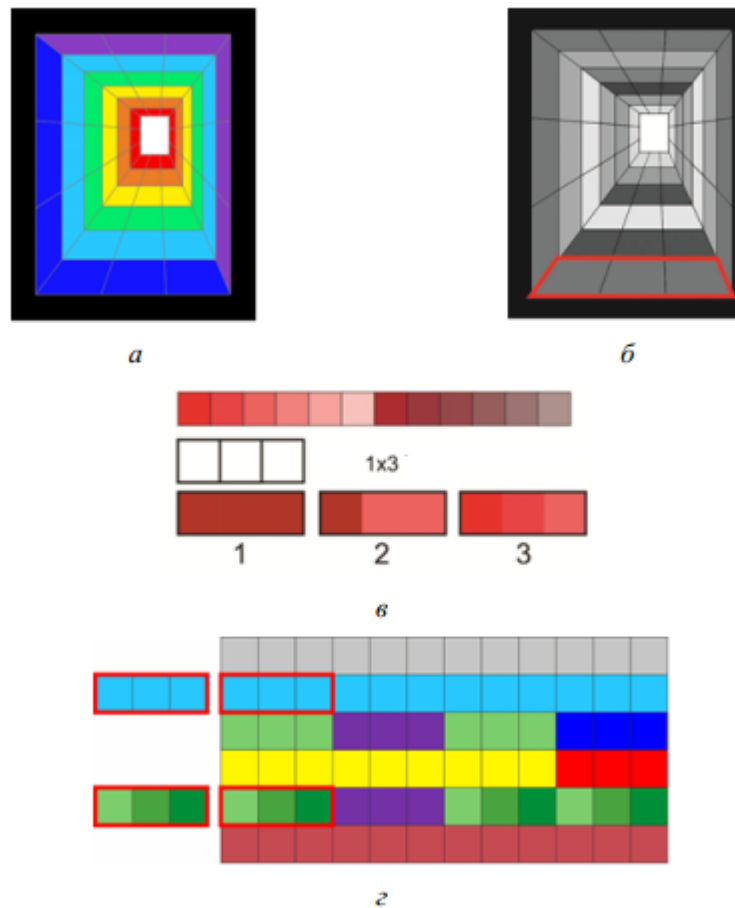


Рисунок 2.2 – Тип А: а – готова тестова форма; б – шуканий елемент (горизонтальні лінії); в – варіанти шуканих елементів в області визначення 3x3; г – розгортка квазіпростору з шуканими елементами

Тип В характеризується наявністю (3 елементи на розгортку і більше) елементів «вертикальні лінії» (від трьох осередків у висоту строго кольору однакових координат, відповідно, область визначення може досягати 6x1. Загальна кількість відтінків може дуже різнитися, головне – 4-х кратне і більше наявність елемента від 3-х осередків у висоту строго однакового кольору.

Ознакою домінуючого типу В є наявність 3-х і більше елементів «вертикальні лінії».

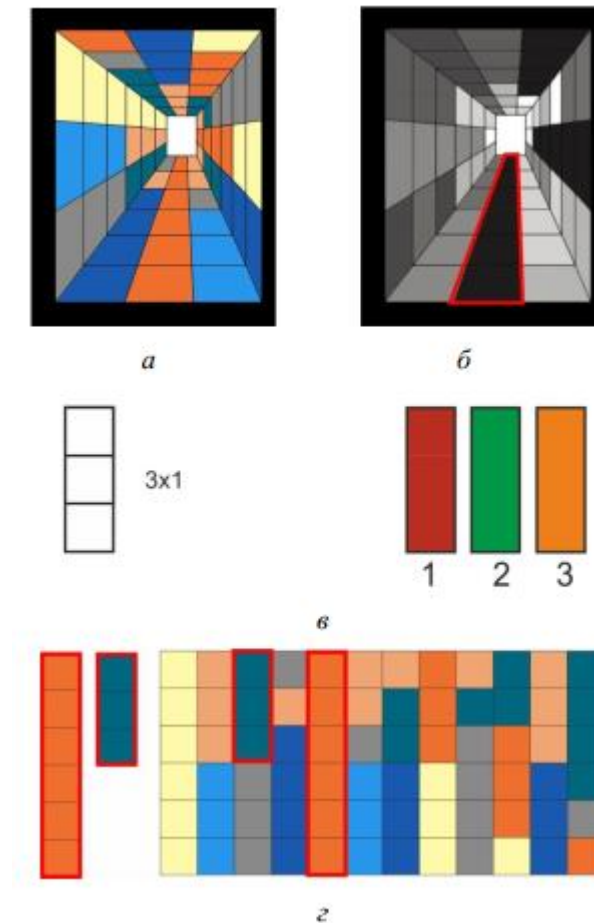


Рисунок 2.3 – Тип В: а – заповнений квазіпростір; б – шуканий елемент у квазі-пространстві (вертикальна лінії); в – варіанти шуканих елементів в області визначення 3x1; г – розгортка квазіпростору з елементами, що шукаються

Тип С характеризується присутністю елементів «горизонтальні лінії» (рисунок 2.4) та «широкі» або «вузькі» градієнти в областях визначення 3 x 3 та 2 x 3 (для широких) та 3 x 1, 4 x 1, 5 x 1, 6 x 1 (для вузьких) (рисунок 2.4). Лінії (від 3 осередків у довжину) можуть становити як кольори строго однакових координат, і кольору різних координат, але які стосуються одному тону (рисунок 2.4).

«Широкі градієнти» становлять смуги кольорів строго однакових координат по ширині і кольорів одного тонового діапазону за довжиною, наприклад – градієнти зеленого та помаранчевого кольорів в областях

визначення 3 x 3 і 2 x 3, ст. Загальна кількість відтінків може відрізнятись, але в основному є 4-5 відтінків для 2-3 тонів. Наприклад, на рисунку 2.4 г, на розгортці видно від трьох відтінків для червоного, зеленого, фіолетового та помаранчевого тонів.

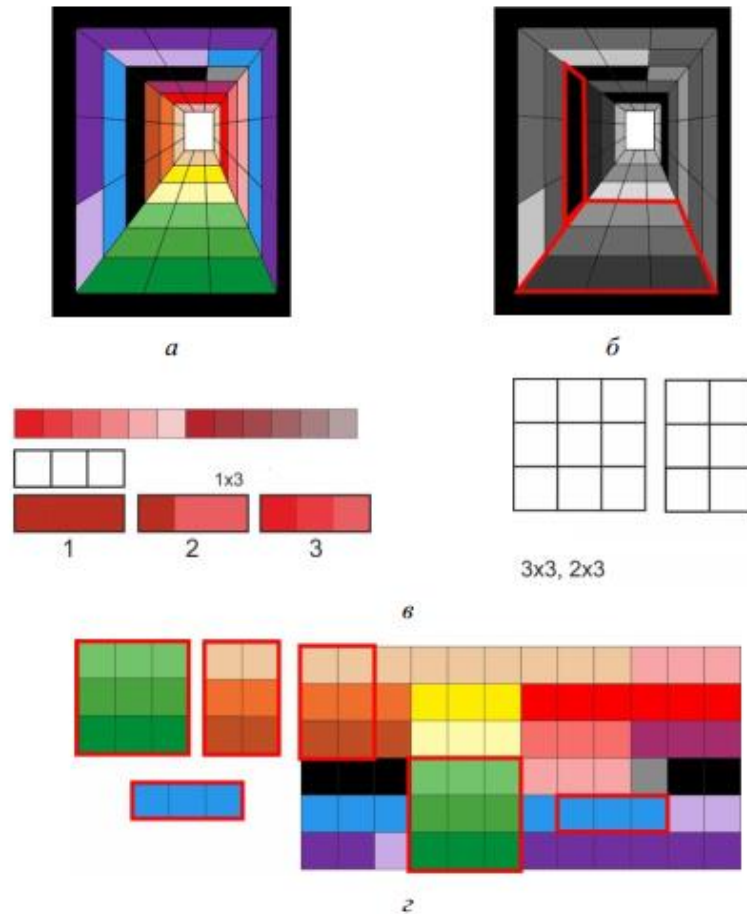


Рисунок 2.4 – Тип С: а – готова тестова форма; б – шуканий елемент (градієнти та лінії); в – варіанти шуканих елементів в областях визначення 1 x 3, 3 x 3; 2 x 3; г – розгортка квазіпростору з шукаються елементами

Ознаки домінуючого типу:

- одночасна наявність елементів «широких» або «вузьких» градієнтів та «ліній»;
- мінімальна кількість елементів для визначення типу 2 лінії та один широкий градієнт 3 x 3 та 2 x 3, або 2 лінії та два вузькі градієнти;
- відносно висока різноманітність та кількість відтінків.

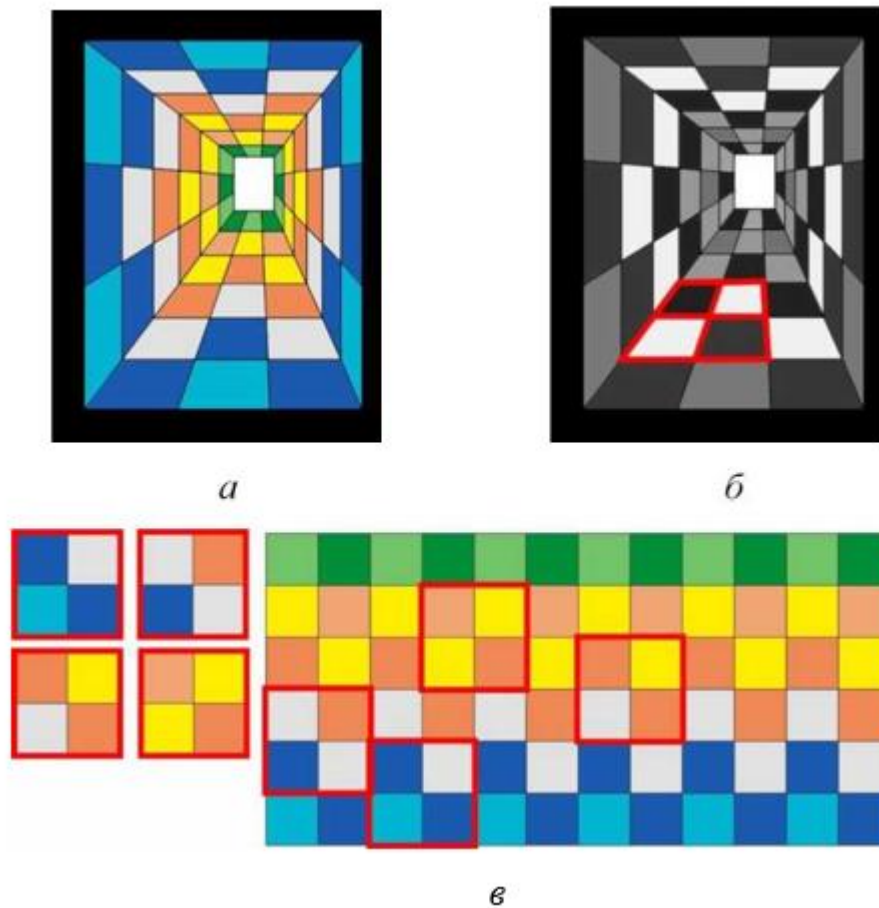


Рисунок 2.5 – Тип D: а – готова тестова форма; б – елемент, що шукається («шахи»); в – розгортка квазіпростору з елементами, що шукаються

Тип D характеризується присутністю елементів «шахи» або «подібність шахів» (рисунок 2.5). Шахові та шахові елементи вибудовуються за рахунок особливого розташування темних і світлих осередків в області визначення 2 x 2. Причому шахові елементи можуть містити будь-які кольори. Загальна кількість відтінків, як правило, невелика.

Ознаки домінуючого типу D:

- наявність елементів «шахи» на 50 і більше відсотках площі разів - вертушки;
- велика кількість осередків з кольорами, що повторюються, однакових координат, особливо чорним, білим, жовтим, червоним і синім.

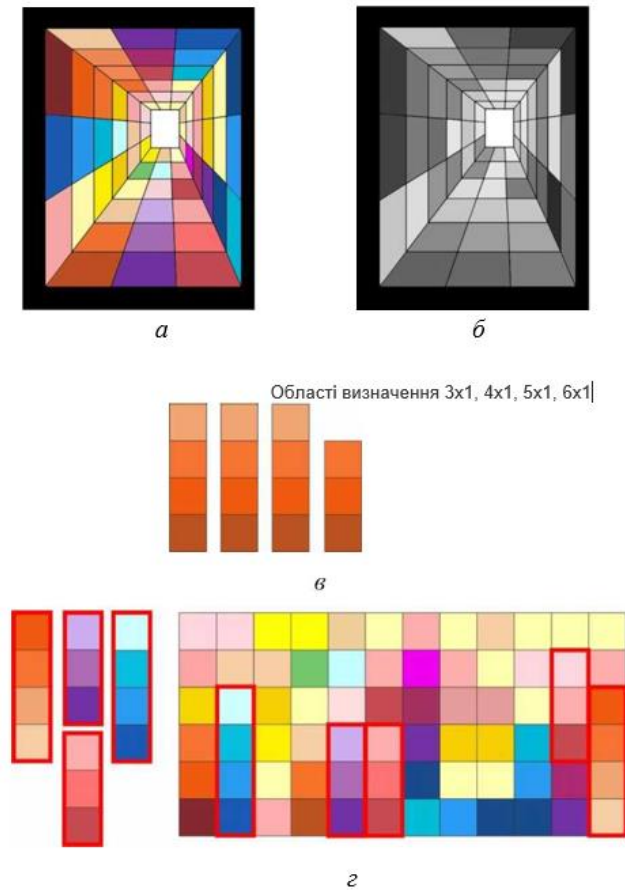


Рисунок 2.6 – Тип Е: а – готова тестова форма; б – елемент, що шукається (градієнти); в – варіанти шуканих елементів в областях визначення 3 x 1, 4 x 1, 5 x 1, 6 x 1; г – розгортка квазіпростор з елементами, що шукаються

Тип Е головним чином характеризується наявністю великої кількості відтінків кольорів, що вибираються, а саме – від 5 відтінків мінімум для двох тонів. Як правило, їх набагато більше. Також деяких представників цього типу відрізняє наявність переважно «вузьких» градієнтів, рідше зустрічаються «широкі».

Ознаки домінуючого типу Е:

- наявність великої кількості відтінків, а саме – від 5 відтінків мінімум для двох тонів;
- переважна наявність вузьких, іноді широких градієнтів.

2.3 Аналіз та класифікація зображень методами машинного навчання

Для розпізнавання, виявлення та класифікації об'єктів використовуються класичні методи та нейромережі. Доросла людина, за умови відсутності відповідних проблем зі здоров'ям і сприйняттям, розпізнає об'єкти навколо себе швидко і безпомилково, здатна назвати їх, відокремити від навколишнього середовища і розрізнити контури, віднести до конкретного класу, визначити розміри, місце розташування в просторі, колір та ін. Очі людини сприймають об'ємні пікселі, після чого мозок перетворює ці пікселі на різні форми: геометричні фігури, лінії, криві. Так людина розуміє, на що саме вона дивиться у той чи інший момент.

Програмні методи, спрямовані на моделювання процесів реального людського зору, ґрунтуються на використанні нейронних мереж. Так як нейромережі складаються з елементів, аналогічних елементарним функціям головного мозку, то вони можуть демонструвати деякі властивості, властиві людському мозку (здатність до навчання, узагальнення досвіду, вилучення істотних властивостей з даних, що відходять, відсікання зайвих даних). Нейронні мережі можуть використовувати навчання з учителем чи ні. Також можна класифікувати нейронні мережі за типом налаштування ваг:

- мережі із фіксованими зв'язками;
- мережі з динамічними зв'язками;
- аналогові мережі;
- двійкові мережі.

За способом подання інформації на входи:

- подача сигналів на синапси вхідних нейронів;
- подача сигналів на виходи вхідних нейронів;
- подача сигналів у вигляді ваг синапсів вхідних нейронів.

За способом отримання інформації з виходів:

- знімання з виходів вихідних нейронів;
- знімання з синапсів вихідних нейронів;

- знімання у вигляді ваг синапсів вихідних нейронів.

За видом функцій активації:

- розривні;
- ступінчасті;
- безперервні;
- від наявності зворотних зв'язків.

Існує безліч архітектур нейронних мереж, проте для виділення складних ознак з великого масиву пікселів використовують згорткові нейронні мережі. Незважаючи на те, що спочатку об'єкти даного дисертаційного дослідження володіють чіткістю, і користувач навмисно або несвідомо виділяє колірні ознаки, розгортка представляє собою досить складний масив різнокольорових осередків.

Згорткова нейронна мережа (CNN або ЗНМ) має схожість з зоровою корою головного мозку людини, на якій розташовані області клітин нейронів, пов'язаних з певними місцями зорового поля. Це відкриття призвело Д. Х'юбел та Т. Візель до Нобелівської премії з медицини у 1981 році. Експеримент 1962 показав, що деякі нейрони давали відгук суворо при спостереженні меж конкретної орієнтації, наприклад, деякі клітини відгукувалися на вертикальні грані, а інші – на горизонтальні. Так, вчені визначили, що дані нейрони є основою візуального сприйняття людини.

Нейронні мережі в цілому відрізняють високу ефективність та продуктивні. Незважаючи на популярність нейромереж, класичні методи аналізу та обробки зображень також удосконалюються і використовуються для різних цілей. Серед таких методів:

- розпізнавання за набором вимірів;
- розпізнавання за елементарними ознаками;
- структурні методи розпізнавання;
- матриця неточностей;
- дерево рішень;
- байєсовський метод прийняття рішень.

Під час сканування зображення (планігона), ковзне вікно покриває - кілька пікселів (осередків), траєкторія може бути лінійною або складнішою. Области визначення можуть бути різними залежно від форми шуканих елементів. Потрапляючи в область «вікна», осередки обробляються за певним правилом (пошук збігів по діагоналі, горизонталі, вертикалі). Деякі поширені види вікон представлені на рисунку 2.7.

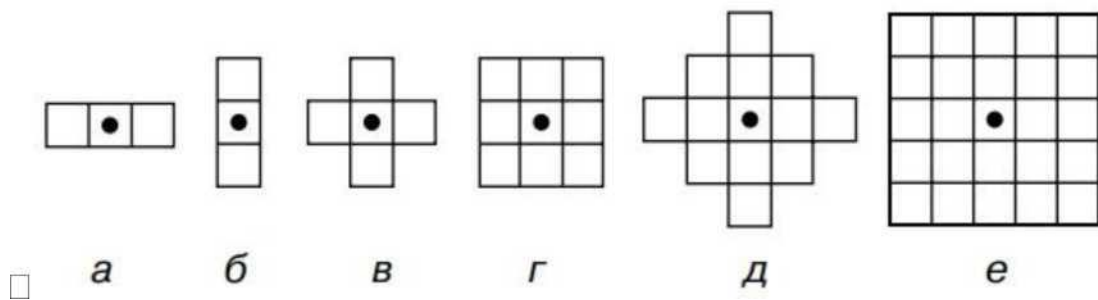


Рисунок 2. 7 – Деякі види «ковзаючих вікон»

Так як об'єктами зображення є прямокутні розгортки, що складаються з квадратних осередків, передбачається, що «ковзне вікно» рухатиметься лінійно і матиме форму прямокутника, або квадрата з областю визначення, обумовлену ознаками шуканого класу зображень.

Перший класифікатор був розроблений основі згорткової нейронної мережі, а в якості основи для другого класифікатора був обраний метод обробки зображень « вікном, що ковзає». Розглянемо перший класифікатор.

Процес навчання нейромережі складається з наступних етапів: навчання, валідація, тестування. Від загальної кількості даних виділяється 30%, у тому числі 15% йдуть на валідацію, 15% на тестування; інші 70 % використовуються на етапі навчання. Є три папки з даними для навчання, валідації та тестування, у кожній папці зображення розбиті за класами.

Для реалізації нейромережевого класифікатора були задіяні три архітектури: MobileNet, Inception_v3 та Une.

Inception V3. У 2012 році, А. Крижевським була опублікована мережа

AlexNet, що є більш поглибленою і розширеною версією мережі LeNet. Її архітектура налічувала 8 рівнів (5 convolutional та 3 fully-connected), але при цьому мережа містила близько 60 мільйонів параметрів. Ця умова не дозволяла здійснювати тренування мережі GPU з обсягом пам'яті в 3GB.

Так, після низки перетворень у 2015 році компанія Google представила мережу вже під назвою Inception V3. Архітектура мережі представлена на рисунку 2.8, вид модуля inception – на рисунку 2.9.

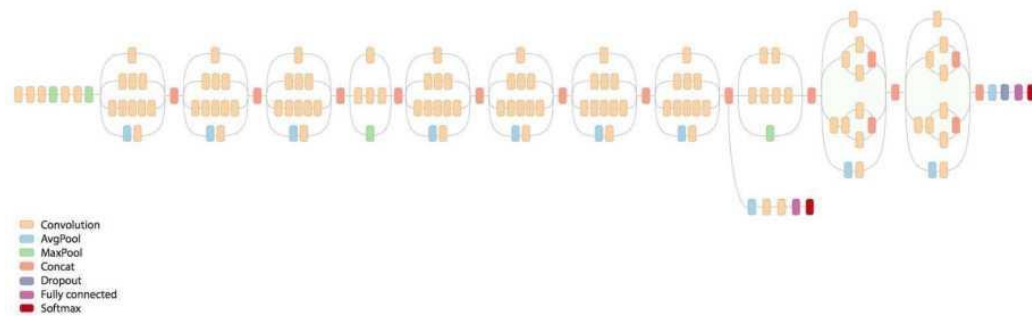


Рисунок 2.8 – Архітектура мережі inception V3

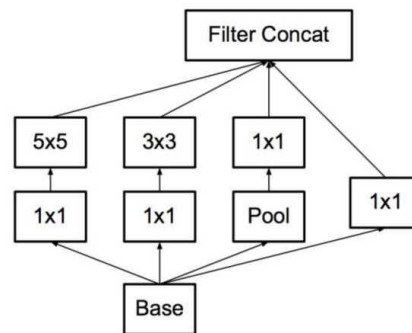


Рисунок 2.9 – Загальний вигляд модуля inception

Навчання. Inception_v3 навчається на 40 епохах. Кількість епох задається вручну, з результатів. Однак занадто велика кількість епох може спричинити перенавчання (ситуацію, коли мережа розпізнає не загальні ознаки, а локальні, наприклад, прив'язується до кольору, наявності або відсутності певної деталі тощо).

У процесі навчання мережа вважає помилку щокроку, потім змінює ваги зменшення кількості помилок. З цією метою використовується оптимізатор стохастичний градієнтний спуск (Stochastic gradient descent, SGD), так як на даний момент при роботі з мережею для профорієнтації він показав себе найбільш ефективно.

Таблиця 2.1 – Схема нейромережі на основі архітектури Inception_v3 для класифікації зображень

Layer (type)	Output Shape	Param #
inception_v3 (Mode 1)	(None, 3, 1, 2048)	21802784
Flatten (Flatten)	(None, 6144)	0
dense_one (Dense)	(None, 64)	393280
dense_two (Dense)	(None, 32)	2080
dropout_one (Dropout)	(None, 32)	0
dense_three (Dense)	(None, 16)	528
output (Dense)	(None, 5)	85

Розглянемо графіки точності та втрати на навчальних та валідаційних даних (рисунок 2.10). Як видно із графіка на рисунку 2.10 а приблизно на 12 епосі відбулося розходження в точності і на валідації та навчанні. Мережа почала перевчитися. Графік на рисунку 2.10 б також свідчить про перенавчання мережі.

Мережі Mobile Nets були створені для мобільних і вбудованих додатків, що використовують комп'ютерний зір. Дані мережі засновані на оптимізованій архітектурі, що використовує згортки, що розділяються по глибині для побудови глибоких нейромереж. Мережами Mobile Nets була продемонстрована висока продуктивність порівняння з іншими моделями класифікації.

MobileNet складається з одного звичайного згорткового шару з згорткою 3×3 на початку і 13 блоків, з числом фільтрів, що поступово збільшується, і за нижчою просторовою розмірністю тензора, як представлено на рисунку 2.11.

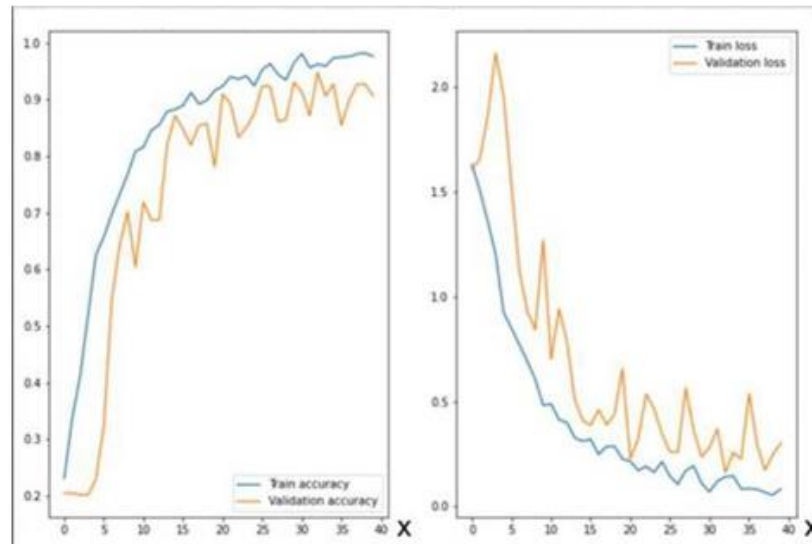


Рисунок 2.10 – Вісь X – кількість епох, вісь Y – точність від 0 до 1:
 а – графік точності на навчальних та валідаційних даних; б – графік втрат на
 навчальних та валідаційних даних



Рисунок 2.11 – Згортка: а – блок звичайної згорткової мережі;
 б – базовий блок мережі MobileNet

Архітектура відрізняється відсутністю max pooling-шарів, замість яких для зниження просторової розмірності використовується згортка з параметром stride, рівним 2. Двома гіперпараметрами архітектури Mobile Net є a (множник ширини) і r (множник глибини або множник роздільної здатності). Множник ширини відповідає за кількість каналів у кожному шарі.

Наприклад, $a = 1$ дає нам архітектуру, описану в статті, а $a = 0,25$ – архітектуру зі зменшеним у чотири рази числом каналів на виході кожного блоку. Чисельник дозволу відповідає за просторові розміри вхідних тензорів. Наприклад, $p = 0,5$ означає, що висота і ширина feature map, що подається на вхід кожному шару, буде зменшена в два рази. Обидва параметри дозволяють змінювати розміри мережі. Так, при зменшенні a і p знижується точність розпізнавання, але збільшується швидкість роботи і зменшується пам'ять, що споживається.

Mobile Net може застосовуватися для додатків на Android та iOS, збереження яких має суттєві обмеження, а кількість оперативної пам'яті в кілька разів менша, ніж на ноутбуках або ПК.

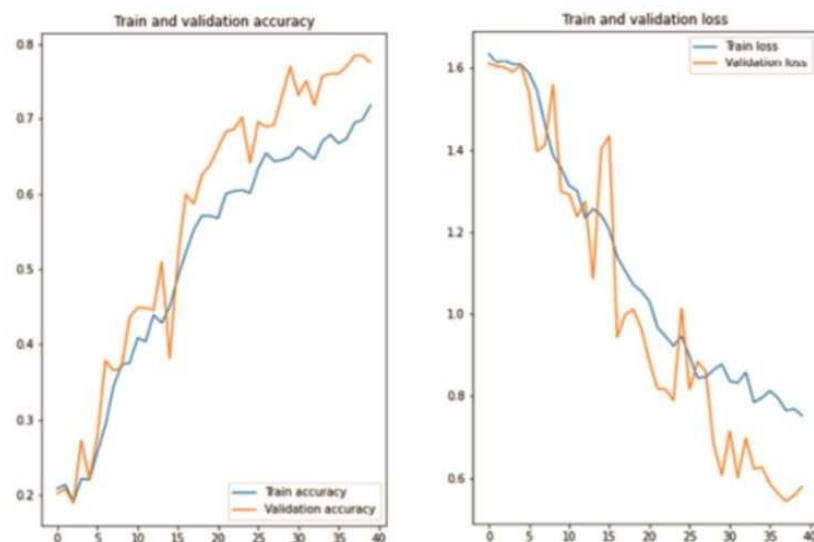


Рисунок 2.12 – Вісь X – кількість епох, вісь Y – точність від 0 до 1: а – графік точності на навчальних та валідаційних даних, б – графік втрат на навчальних та валідаційних даних

На графіках рисунку 2.12 наведено показники точності та втрати на навчальних та валідаційних даних для мережі, розробленої на основі архітектури MobileNet для профорієнтації.

U-net. Архітектура мережі U-net наведена на рисунку 2.13. Вона складається з шляху, що звужується (ліворуч) і розширюється шляху

(праворуч). Звуження шляху, у свою чергу, складається з повторного застосування двох згорток 3×3 , за якими слідує випрямлений лінійний блок (ReLU) і операція максимального об'єднання (2×2 ступеня 2) для зниження дозволу. На кожному етапі знижувальної дискретизації канали властивостей подвоюються. Кожен крок у шляху, що розширюється, складається з операції підвищуючої дискретизації карти властивостей, за якою слідують:

- згортка 2×2 , яка зменшує кількість каналів властивостей;
- об'єднання з відповідним чином обрізаною картою властивостей зі стягується шляху;
- дві 3×3 згортки, за якими слідує ReLU.

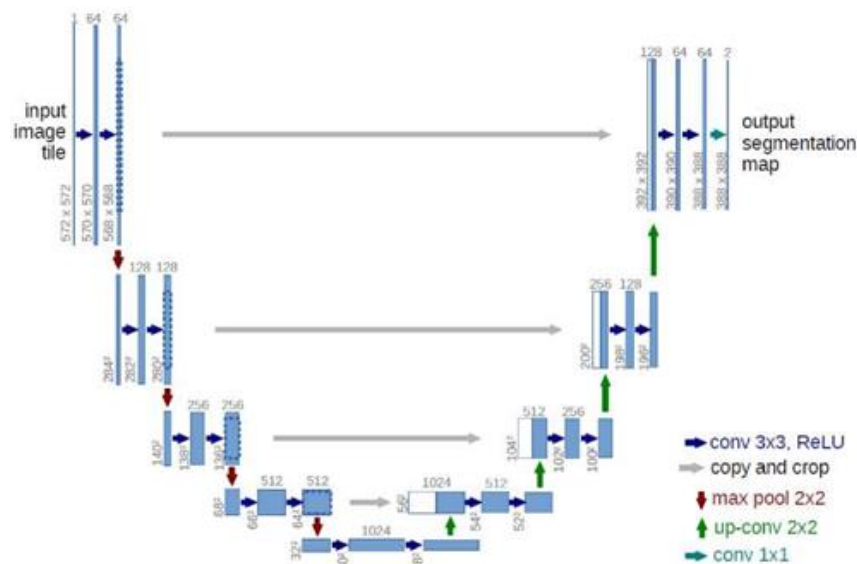


Рисунок 2.13 – U-net

На останньому шарі використовується згортка 1×1 для зіставлення кожного 64-компонентного вектора властивостей із бажаною кількістю класів. Усього мережа містить 23 згорткових шари.

Розглянемо графіки точності та втрати на навчальних та валідаційних даних (рисунок 2.13). Перший графік показує, що приблизно на 25 епосі точність мережі досягла 1, але точність на валідації знаходиться в діапазоні 0.6 – 0.65. Це означає, що на зразках, використаних для навчання, точність

становитиме близько 100%, але на зразках, які будуть використані для перевірки, точність становитиме 65%.

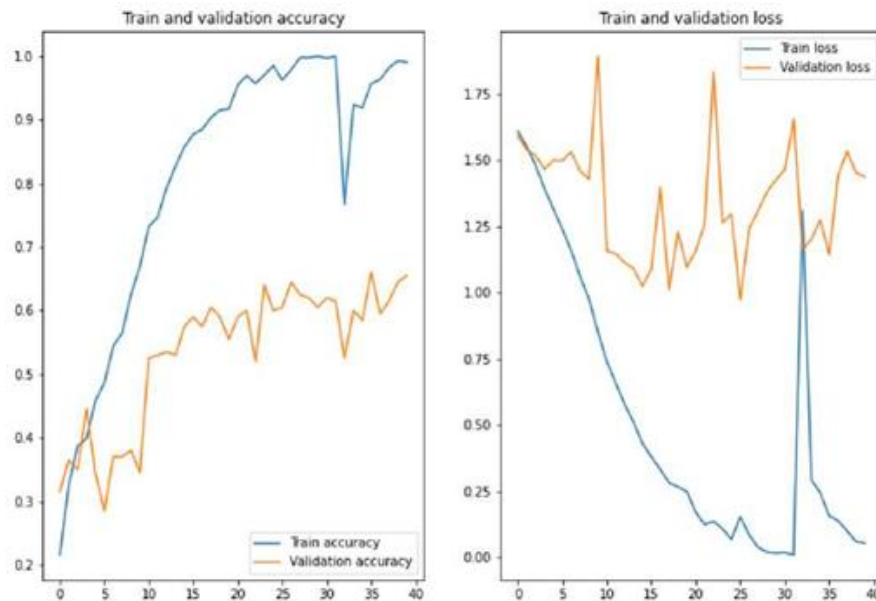


Рисунок 2.13 – Вісь X – кількість епох, вісь Y – точність від 0 до 1: а – графік точності на навчальних та валідаційних даних, б – графік втрат на навчальних та валідаційних даних

Три вибрані архітектури нейронних мереж на момент завершення дослідження показали такі показники точності (таблиця 2.2).

Таблиця 2.2 – Точність класифікації нейронних мереж

Inception_v3	MobileNet	Unet
91,12%	84,51%	71,05%

Графіки представлені на рисунку 2.14 також показують ефективність використовуваних мереж для класифікації зображень.

Так, найбільше ефективно показав себе класифікатор, заснований на архітектурі Inception_v3. Для поліпшення роботи мережі та подальших експериментів з розширенням вибірки доведеться змінити оптимізатор, створити нові шари класифікації, можливо – зменшити крок зміни ваг, так як

подібна форма кривої валідації може свідчити про те, що значення змінюються занадто сильно на кожній епосі.

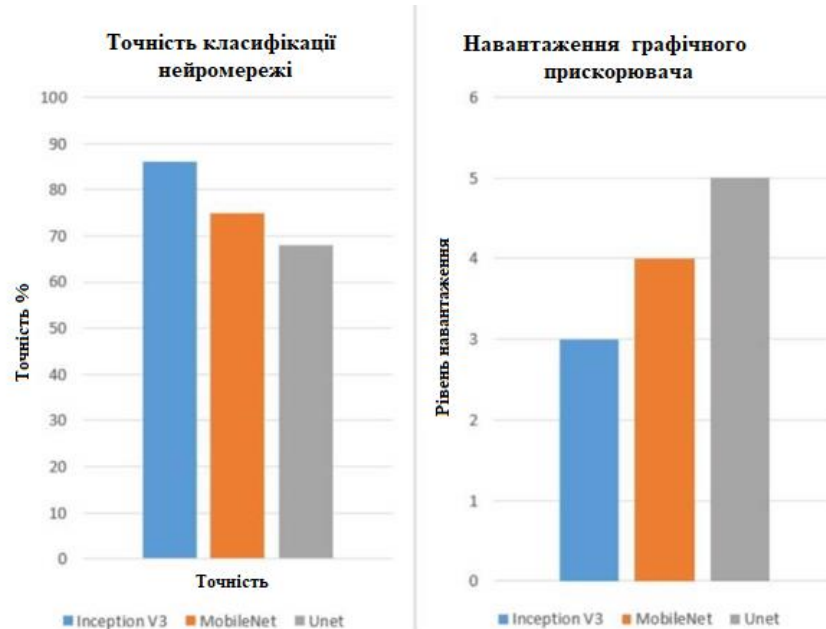


Рисунок 2.14 – Ефективність класифікації використовуваних мереж: точність класифікації; рівень навантаження відеокарти під час навчання

Валідація. Навчальна вибірка поділяється на ітерації. Мережа «дивиться» певну кількість зображень за ітерацію, підбирає ваги. Потім оцінює помилку. Далі, з цими ж вагами мережа проходить у валідаційну вибірку (15% даних від навчальної вибірки) та застосовує їх на зображення у валідаційній вибірці. Зменшення помилки одночасно на навчанні та валідації свідчить про успішність навчання мережі. Якщо ж помилка на навчанні падає, але в валідації зростає, отже, відбувається перенавчання.

Тестування . Мережі тестуються на даних, які вони раніше не бачили - (15% від навчальної вибірки). Тобто мережа не знає, який клас є на тому чи іншому зображенні та намагається визначити це самостійно. Отримані відповіді збираються, виводиться процент правильного визначення.

Бібліотека . Бібліотека Keras використовувалася для побудови самих нейромереж, бібліотека комп'ютерного зору OpenCV – для пошуку дублікатів.

Аугментація. Для першого тестування нейромережі використовувалися зображення, розділені за типами (класами). А-1, А-2... А-70; В-1, В-2 ... В-70 і т. д. до Е-70. Сумарна кількість 1530 зображень. Через те, що зображень для навчання не вистачало, було вирішено застосувати аугментацію. Через специфіку завдання деякі методи аугментації, наприклад, такі як розтягування, повороти на градус вуса, відмінні від 180, викривлення та подібні не застосовувалися.

Методи аугментації, застосовані у завданні (рисунок 2.15):

- поворот на 180° (для таких зображень додано назву 180);
- відображення по вертикалі (для таких зображень додано назву top_bottom);
- відображення по горизонталі (для таких зображень додано до назви - left_right).



Рисунок 2.15 – Методи аугментації

Метод фільтрації реалізується мовою програмування Python і заснований на застосуванні алгоритму обробки зображень ковзним вікном. Фільтри $F(3)$, $F(4)$, $F(5)$ засновані на використанні областей визначення $N \times N$, кількість осередків яких обумовлена особливостями елементів шуканих структур. У загальному алгоритмі на даний момент окремо не виділено фільтр для виявлення шуканої структури S , так дана шукана структури містить як елементи, характерні для типу E , так і елементи, характерні для

типу А. Тип С планується ідентифікувати на завершальній стадії аналізу готової тестової форми за умови вираженості структур обох названих типів (понад 50%).

Колір (у колористиці) – це властивість світла викликати певні зорові відчуття відповідно до спектрального складу відбитого випромінювання. Випромінювання з різною довжиною хвилі викликає різні колірні відчуття

Колірний тон – це якість кольору, що визначається довжиною світлової хвилі і прирівнювана до одного з спектральних і неспектральних (пурпурних) кольорів. Колірний тон (X) також визначає назву кольору, наприклад, «червоний», «помаранчевий», «жовтий» тощо.

В описі принципу роботи фільтра F(4) під осередками одного кольору маються на увазі осередки, що мають кольори строго однакових координат у системі RGB.

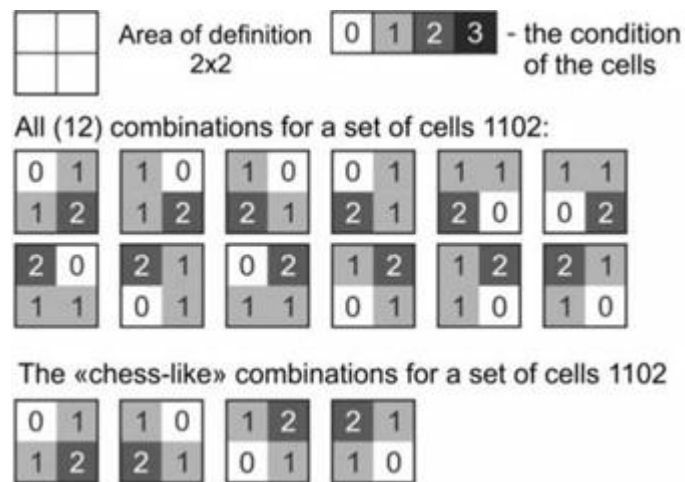


Рисунок 2.16 – Графічна схема застосовуваних комбінацій осередків 4-х різних станів без повторень для області визначення 2x2

Останній з фільтрів F(5) – єдиний фільтр, де потрібно перевод розгортки квазіпростору в градації сірого. Це пов'язано з тим, що «шахову» чи «шаховоподібну» структури можуть становити осередки різних кольорів. Автор методики припустила, що переведення планігонів у градації сірого

дозволить виділяти та зіставляти осередки у виділеній області 2x2 з еталонами (комбінаціями) «шахових» і «шахматоподібних структур».

Спочатку стани осередків визначалися наступними діапазонами колірної системи СМҮК для градацій сірого (0 - С:0, М:0, Y:0, К:0-24; 1 - С:0, М:0, Y:0, К:25 -49; 2 - С: 0, М: 0, Y: 0, К: 50-74; .

Таким чином, отримуємо чотири стани для осередків. Для моделювання еталонів (шуканих елементів «шахової» або «шахматоподібної структури») осередку потрібно розподілити в межах області визначення 2x2, де комбінації відрізнятимуться один від одного тільки порядком розташування осередків. Таким чином, отримуємо 24 комбінації для 4 осередків, де для заповнення області визначення 2x2 використовуються всі чотири осередки без дублювання. Але «шахова» або «шахова» структури можуть також являти собою комбінації, що складаються з набору осередків 2-х і 3-х станів, з повтореннями. Не всі отримані комбінації мають «шахову» або «шаховоподібну структуру, в деяких випадках осередки однакового ступеня світлоти розташовані поруч або один над одним, такі комбінації виключаються із зразків. Разом залишається 48 «шахових» і «шахів» елементів для фільтра F(5).

Далі, реалізуючи алгоритм, кожному осередку присвоюється сірий колір відповідно до наступної формулою: колір = $0,299 * R + 0,587 * G + 0,114 * B$. Далі всі сірі кольори (від 0 до 255) осередків діляться на 4 діапазони по 3 кольорах кордонів (С1, С2, С3), і осередку присвоюється номер (1, 2, 3 або 4), залежно від того, в який діапазон (0-С1, С1-С2, С2-С3, С3-255) потрапив її колір. Результат сканування комірки, що тирає, відображається чотирма кольорами, залежно від номера діапазону. Кожному номеру діапазону заздалегідь було надано колір (1: білий, 2: світло-сірий, 3: сірий, 4: темно-сірий). І, нарешті, виконується аналіз на наявність бажаних елементів класу D, які виявляють комбінації повторень двох станів, першого стану. Комбінації без повторень залишаються після виключення з їх розрахунку області визначення комбінацій 2x2, де всі кольори однакові, де три кольори однакові і де два однакові кольори знаходяться поруч один з одним.

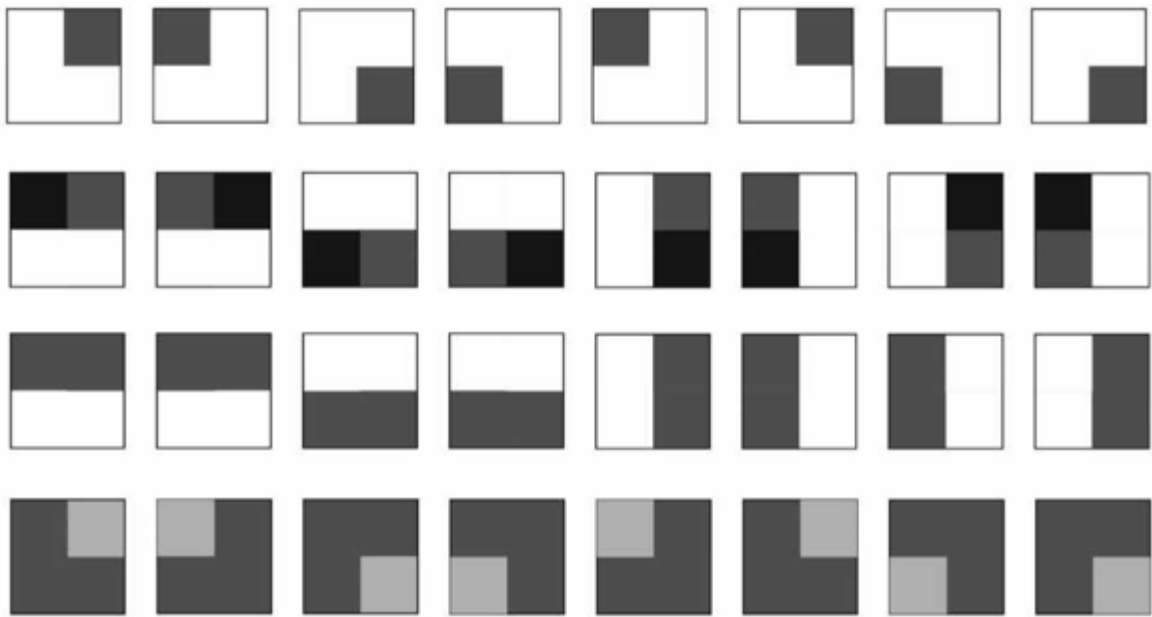


Рисунок 2.17 – Деякі фігури-виключення для області визначення 2x2

Розроблено гібридний класифікатор за параметрами зображень готових тестових форм, що відрізняється від відомих методів, спрямованих на виявлення неявних характерних елементів структур (неочевидних для експерта), що свідчать про збіг з одним із еталонів. Класифікатор відрізняється системою поетапного аналізу зображення, що включає аналіз - особливостей кольору, структури, форми, можливістю розділяти типи і підтипи з урахуванням неоднорідності ознак шуканих класів у процесі загальної інтерпретації результатів (порівняння результатів обох класифікаторів).

3 ПРОГРАМНА РЕАЛІЗАЦІЯ

У процесі реалізації та апробації класифікаторів насамперед зразки проходили тестування через нейромережевий класифікатор, потім – через «ковзне вікно». Другий класифікатор, одночасно з навчанням і тестуванням першого, аналізував зображення, що надходять, тільки від справжніх користувачів, так як для нього не була потрібна аугментація і процес налагодження не вимагав пред'явлення нових зображень. Так само другий класифікатор не схильний до перенавчання.

У процесі проведення тестів і зіставлення результатів класифікації зображень обома класифікаторами були отримані висновки, що свідчать про те, що обидва класифікатори досить ефективні для певних класів і для підвищення ефективності процесу тестування не слід відмовлятися від будь-якого одного на користь іншого.

3.1 Проведення експериментів

Тест 1. Кількість зображень для тестування: 150 (25 зображень на клас). Класифікатор з урахуванням архітектури мережі Insertion_v3. Експерт: якщо робітник методики тестування.

Правильно - збіг думки експерта та результату класифікації мережі.

Помилка класифікатора – класифікатор називає неправильний клас (признаки названого класу або відсутні, або є в малій кількості порівняно з іншими).

Помилка експерта – експерт надає зображенню неправильний клас, або поміщає розгортку в папку з вірним класом, але у своєму списку відзначає для даної розгортки найбільш ймовірним інший клас.

Результати тесту 1 (рисунок 3.1)

Найбільший відсоток вірних відповідей виявлено у класі Е (65%), найменший – у класі D (35%).

Найбільший відсоток помилок експерта виявлено у класі С (65%), найменший – у класі D (10%).

Точність мережі з урахуванням помилки експерта за класами склала: А – 64%, В – 70%, С – 75%, D – 45%, Е – 65%.

Загальна точність мережі становила 63,8%.

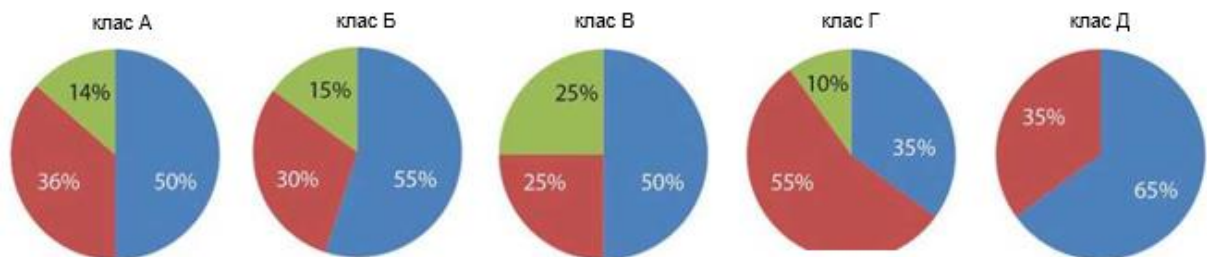


Рисунок 3.1 – Діаграми тестової вибірки: синій – вірно, червоний – помилка мережі, зелений – помилка експерта

Тест 2. Кількість зображень для тестування: 150 (25 зображень на клас). Класифікатор з урахуванням архітектури мережі Inception_v3. У цій тестовій вибірці були присутні «яскраві» (візуально обумовлені) і сумнівні (візуально неоднорідні) зразки у співвідношенні 50/50. Експерт: розробник методики тестування.

Результати тесту 2 (рисунок 3.2):

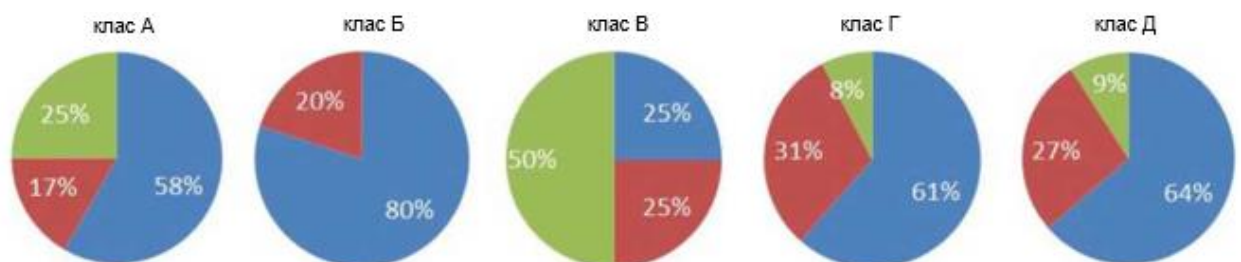


Рисунок 3.2 – Діаграми тестової вибірки: синій – вірно, червоний – помилка мережі, зелений – помилка експерта

У вибірці зразків з неоднорідністю ознак: найбільший відсоток вірних відповідей виявлено в класі В (80%), найменший – D (69%).

У вибірці зразків з візуально визначеними ознаками: найбільша точність мережі виявлена в класі А (100%), найменша – (75%) і С (75%).

Помилка експерта найбільш висока у класі С (50%), найменше – D(8%).

Результати тесту 3 (рисунок 3.3).

Відсоток вірних відповідей становив 71%.

Відсоток помилок мережі становив 15%.

Відсоток помилок експерта становив 14%.

Загальна точність мережі становила 85%.

Так, за підсумками тестів 1, 2 та 3 видно, що кількість вірних відповідей клас сифікатора збільшується, кількість помилок знижується, тоді як кількість помилок експерта збільшується на момент проведення тесту 2 і знову знижується на тесті 3.

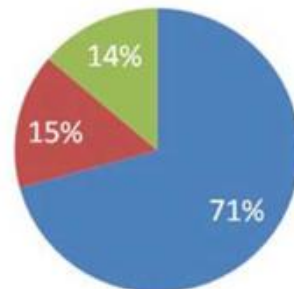


Рисунок 3.3 – Діаграма тестової вибірки

Точність мережі з урахуванням помилки експерта за класами склала:

- зразки з неоднорідністю ознак: А – 83%, В – 80%, С – 75%, D – 69%, Е – 73%;

- зразки з візуально визначеними ознаками: А – 100%, В – 75%, С – 75%, D – 87%, Е – 80%.

Загальна точність мережі становила 80%.

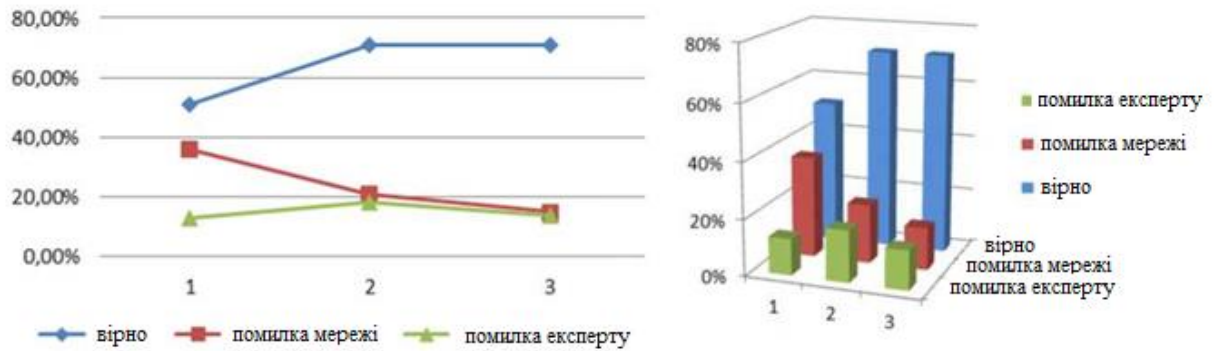


Рисунок 3.4 – Графік та гістограма динаміки точності класифікації між тестами 1, 2 та 3

За результатами аналізу зображень першої вибірки нейромережевим класифікатором стало ясно, що однією з головних проблем класифікатора є перенавчання мережі. Так, у ряді випадків мережа відносила розгортку до класу А якщо на зображенні були присутні горизонтальні смуги, що складаються з трьох осередків синього кольору, або помилково відносила розгортку до класу В, плутаючи розтяжку з вертикальною лінією строго одного кольору. Загалом перший класифікатор на основі згорткової нейронної мережі Inception_v3 ефективно виділяє розгортки типів А, В, С (у той час як саме в цьому класі існує найвища ймовірність помилки експерта) та Е, тому його не можна назвати неефективним.

Потенційно класифікатор, заснований на методі фільтрації з застосуванням «ковзного вікна», може бути налаштований на визначення ознак присутності всіх виділених класів, проте це представляє певні труднощі в класах С і при розділі лінії розгортки, що належать підтипам класів D і E. Так, перший класифікатор має розбіжність між прогнозованою точністю (91,12%) і реальною точністю (80-85%), що дорівнює 6-11% (5% відводиться на затруднення інтерпретації експертом результату машинного аналізу). Другий класифікатор дозволяє розділити типи за виділеними характеристиками, при спільній інтерпретації результатів обох класифікаторів можливий поділ типів і підтипів з урахуванням неявних

структур і неоднорідності ознак шуканих класів, а також індивідуалізація результатів за рахунок визначення наближеності зображення до групи типів, ознаки яких присутні в структуру.

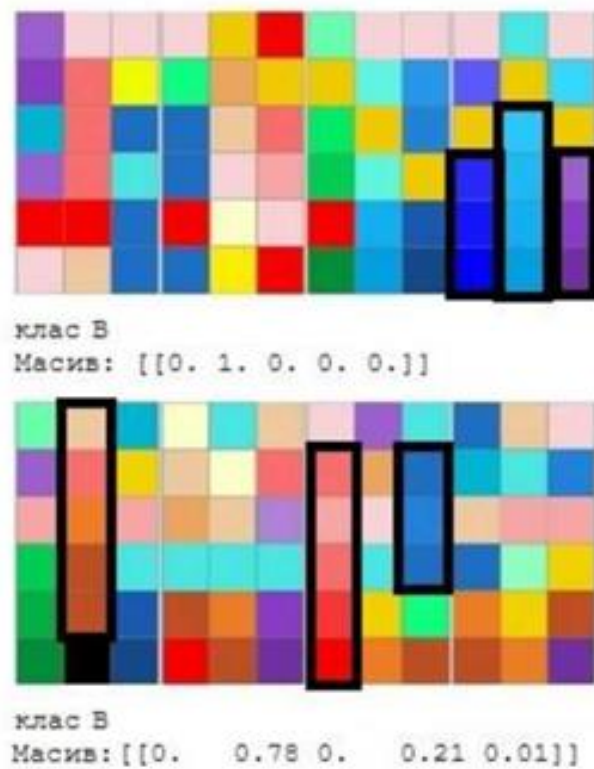


Рисунок 3.5 – Помилкова ідентифікація нейромережі розгортки як належать класу B, тоді як це розгортки класу E

3.2 Гібридний класифікатор

Перший класифікатор на основі згорткової нейронної мережі Insertion_v3 ефективно виділяє розгортки класу E як яскраві, так і з ознаками елементів інших шуканих типів, у тому числі – розгортки, що відносять до «неконтрастного» підтипу E. Саме всередині цього типу розмежування під типів представляє самі великі тонкощі та складності, оскільки в деяких випадках розгортки, що належать класу E, дуже близько підходять до розгортки класу D.

Гібридний класифікатор будувався з метою розмежування підтипів та аналізу розгортки з неоднорідністю ознак шуканих класів, і якщо при візуальному аналізі розгортки експерт відносив до чверті всіх зображень до «змішаного» типу без конкретних рекомендацій, то при використанні гібридного класифікатора «невизначений» тип має бути зведений до мінімуму.

Незважаючи на те, що другий класифікатор побудований на виявлення конкретних позначених ознак, його реалізація була пов'язана з вирішенням низки додаткових завдань:

- завдання виявлення елементів «шахових» та «шахів» структур у разі використання випробуванням кольорів, що мають близькі;
- ступеня насиченості, наприклад, червоно-зелений, червоно-оранжевий;
- завданням реалізації алгоритму є автоматичне розмежування візуально схожих підтипів, але мають різні характеристики ІСД;
- завдання виявлення та розмежування «демонстративної» і «прихованої» структур, підрахунок виразності всіх ознак у відсотковому співвідношенні, оцінка впливу наявності «прихованої» структури на ІСД індивіда.

У разі використання випробуванням кольорів, що мають близькі ступеня насиченості, наприклад, червоно-зелений, червоно-помаранчевий, при перекладі зображення в градації сірого, що передбачає реалізацію класифікатора за методом фільтрації, може відбуватися «втрата» шуканих «шахів» структур. У разі, коли тип D виражений досить яскраво, структура легко визначається як візуально, так і класифікатором, проте «чисті» типи зустрічаються все рідше, ніж менш «яскраві», також більшість розгортки містить ту чи іншу кількість ознак інших класів.

«Яскравих» і еталонних зразків, що легко визначаються візуально, зустрічається не так багато. Як правило, такі розгортки відносяться до класу А і містять досить мало тонів і відтінків.

Як тільки випробуваний починає використовувати колірні розтяжки або формувати «шахматоподібні» структури, візуальна ідентифікація затрудняється, особливо при оцінці квазіпростору - через перспективне спотворення більш і «близькі» до глядача (експерту) осередки здаються важливішими і займають більше місця в зорового поля. Але й за візуальної оцінки розгортки можливі труднощі. Наприклад, дана розгортка відноситься до класу С, але контрастні «шахматоподібні» області привертають увагу і створюють враження домінування.

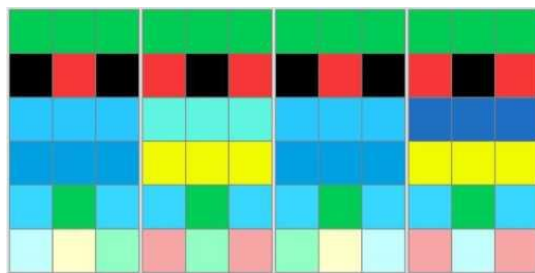


Рисунок 3.6 – Розгортка з ознаками кількох шуканих класів

Насправді ж «шахматоподібні» елементи зустрічаються тільки в нижньому ряду, верхній ряд заповнений темними осередками, що чергуються, з тими, що дають ілюзію наявності даної структури. Складніше розрізнити «неконтрастний» підтип Е і тип D з рисами типу Е. Згідно з описом ознак класу Е, у розгортці має бути від п'яти відтінків мінімум для двох тонів, а також має бути наявність «вузьких» градієнтів або вертикальних розтяжок. На рисунку 3.7 представлені квадрати простору типів Е, праворуч – еталонний, зліва – «неконтрастний».

В обох випадках колірні «розтяжки» присутні, у «неконтрастного» в цілому менше кольорів, що використовуються, але є два домінуючі тони з великою кількістю відтінків, які близько розташовані один від одного в просторі колірного кола.

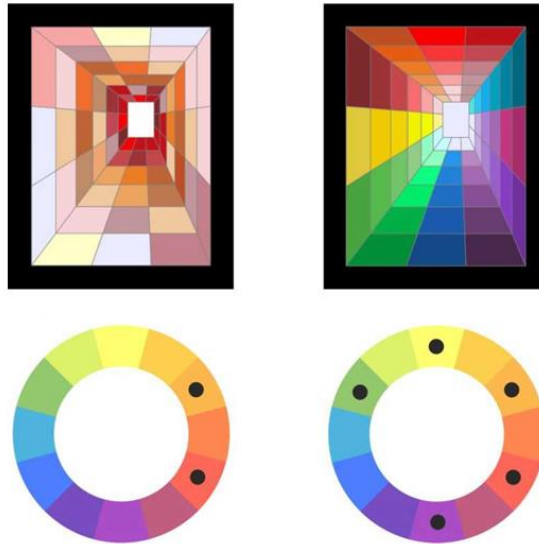


Рисунок 3.7 – Розгортка з ознаками кількох шуканих класів

Саме тому даний підтип отримав назву «неконтрастний», тому що його домінуючі тони, як правило, знаходяться або в одній із «родинних» чвертей колірної кола, або в одній із «родинно-контрастних» половин.

У розгортках «неконтрастних» підтипів E може бути присутнім і більша кількість тонів, але 2-3 домінуючі тони лежатимуть або в одній чверті колірної кола, або в одній його половині. У розгортках типу E зазвичай задіяні всі або майже всі тони, кожен з них має кількість відтінків від 4-5 серед менш задіяних і до 12 серед більш задіяних.

Але справа в тому, що, на відміну від еталонних E, «неконтрастні» підтипи піддослідних «розбивають» колірні розтяжки на комірочки і розміщують їх у квадраті простору на «шахматоподібний» манер, що властиво представникам типу D.

3.3 Аналіз отриманих результатів

Додаткове завдання номер один (втрата шахматоподібних структур при переведенні розгортки в градації сірого) ефективно вирішилася за рахунок

розробки кількох підфільтрів. Спочатку їх було 20, але після дослідження на тестовій вибірці (35 осіб) залишилося 11 найефективніших.

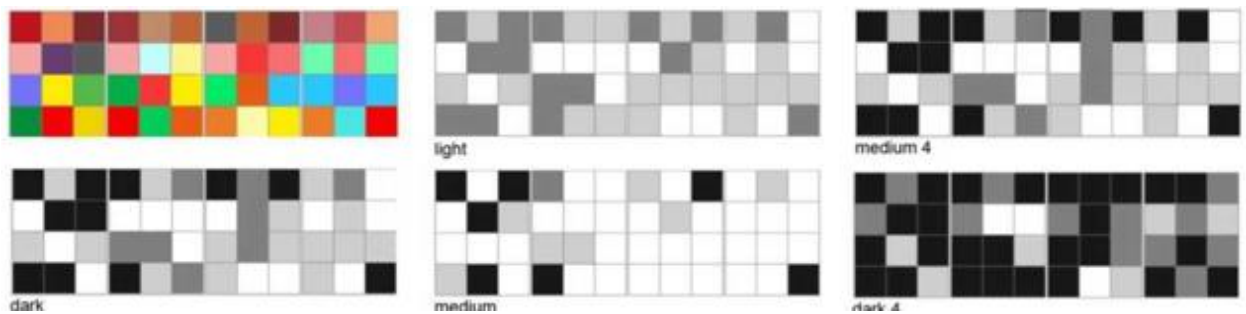


Рисунок 3.8 – Виявлення структур, що шукаються в різних галузях планігону

Таблиця 3.1 – Ефективність системи підфільтрів для 10 різних прикладів зображень

приклад фільтра	1	2	3	4	5	6	7	8	9	10
Dark	80%	18%	34%	17%	9%	19%	21%	20%	18%	13%
Dark 2	100%	72%	32%	33%	0%	15%	13%	20%	11%	29%
Dark 3	40%	0%	30%	4%	9%	19%	17%	0%	7%	0%
Light	80%	18%	29%	6%	0%	9%	23%	0%	15%	12%
Light 2	40%	0%	24%	4%	5%	27%	17%	0%	7%	0%
Light 3	20%	0%	7%	2%	11%	5%	15%	0%	11%	0%
Medium	40%	0%	26%	4%	9%	24%	17%	0%	7%	0%
Medium 2	40%	18%	22%	9%	9%	15%	20%	0%	18%	15%
Medium 3	80%	73%	29%	16%	9%	17%	18%	20%	18%	24%
Medium 4	80%	36%	31%	28%	9%	9%	19%	40%	15%	20%
Dark 4	80%	36%	25%	14%	0%	22%	9%	40%	11%	25%
The results	100%	73%	34%	33%	11%	27%	23%	40%	18%	29%

Зелений – найбільш ефективний фільтр, синій – другий за ефективністю фільтр, жовтий – зразок, що імовірно відноситься до підтипу класу об'єктів E – неконтрастний підтип E.

Після аналізу програма вважає ступінь вираженості типу D у відсотковому співвідношенні для кожного з підфільтрів, для виведення підсумкового відсотка береться найвищий результат, оскільки найчастіше в процесі фільтрації при виявленні втрачених структур в одній частині зображення вони «стираються» в іншій, і навпаки.

Друге з додаткових завдань (розподіл типів і підтипів, рисунок 3.9) ефективно вирішується спільним використанням нейромережевого класифікатора на основі мережі Inception_v3 і методом фільтрації з системою подфільтрів для класу D.

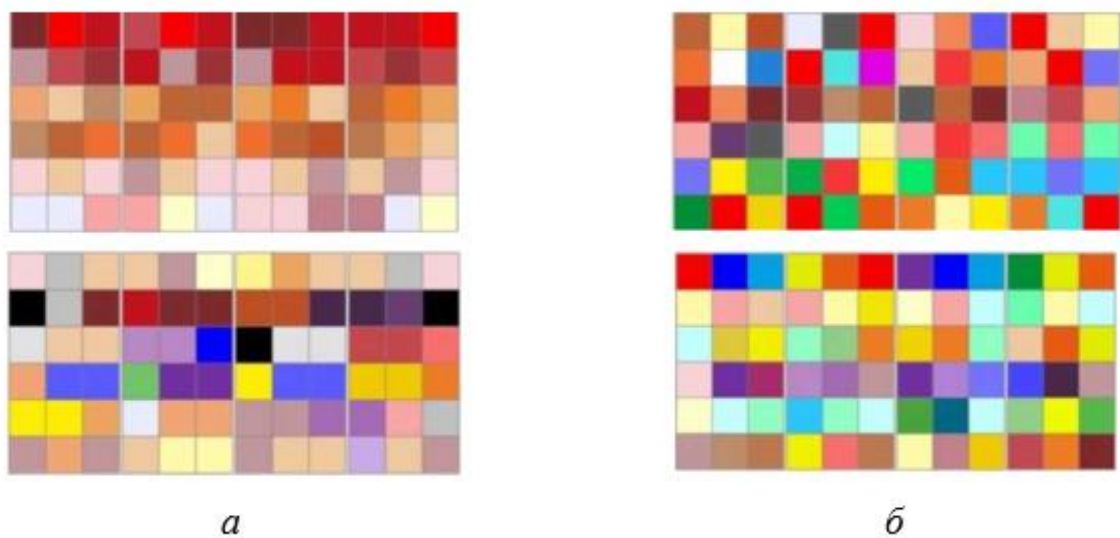


Рисунок 3.9 – Поділ типів і підтипів: а – зображення класу E підтипу «неконтрастний», б – зображення класу D з прихованими ознаками класу E

Так, якщо у зображенні присутнє до 45% структур класу D, але нейромережа визначає зображення до класу E, то перед нами «неконтрастний» підтип E, якщо нейромережа відносити зображення до класу D – перед нами тип D.

Для поділу підтипів було використано 35 зображень з ознаками класів D і E, 10 з них виявилось належать до «неконтрастного» підкласу E, 12 – класу D з рисами E і 13 зображень відносяться до класу E за наявності ознак класу D (від 20 до 75 %). В останньому випадку випробуваному рекомендується ознайомитися з характеристиками як для класу E, так і D, оскільки ознаки обох класів у таких розгортках виражені істотно (відтінки у кількості більше 5-ти для 3-х і більше кольорів за наявності «шаха топодібних» структур більше 50%).

ВИСНОВКИ

Досліджено методи аналізу та класифікації зображень з використанням машинного навчання. Розроблено гібридний класифікатор за параметрами зображень готових тестових форм, що відрізняється від відомих методів, спрямованих на виявлення неявних характерних елементів структур (неочевидних для експерта), що свідчать про збіг з одним із еталонів. Класифікатор відрізняється системою поетапного аналізу зображення, що включає аналіз особливостей кольору, структури, форми, можливістю розділяти типи і підтипи з урахуванням неоднорідності ознак шуканих класів у процесі загальної інтерпретації результатів (порівняння результатів обох класифікаторів).

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Дяченко Д. О., Кайда В. В., Левченко А. О., Міхаль О. П. Методи функціонування пристроїв IoT з використанням машинного навчання // Системи управління, навігації та зв'язку. 2024. № 2(76). С. 78-81.
2. Timmis J. An Overview of Artificial Immune Systems / J. Timmis, T. Knight, L.N. de Castro, E. Hart // *Computation in Cells and Tissues: Perspectives and Tools for Thought*, Natural Computation Series, Springer, 2004 – pp. 51-86.
3. Castro L. N. Artificial Immune Systems: A Novel Paradigm to Pattern Recognition / Castro L. N., Timmis J. I. // *In Artificial Neural Networks in Pattern Recognition*, SOCO-2002, University of Paisley, UK, pp. 67-84.
4. Борисов В.В., Круглов В.В., Федулов Ф.С. Нечеткие модели и сети. – М: Горячая линия. – Телеком, 2007. 7. Е.В. Бодянский, О.Г. Руденко Искусственные нейронные сети: архитектуры, обучение, применения // Харьков: ТЕЛІТЕХ, 2004. – 372с.
5. Растрингин Л.А. Адаптация сложных систем. Методы и приложения // Рига: Зинатне, 1981.– 375 с.
6. Цыпкин Я.З. Адаптация и обучение в автоматизированных системах // М., Наука, 1968.
7. Lauren M. Sompayrac *How the Immune System Works*. 4th Edition // Wiley-blackwell. 2012. – 153 p.
8. Ginzberg M. J., Stohr E. A. *Decision Support Systems: Issues and Perspectives*. // *Processes and Tools for Decision Support*. / Ed. by H. G. Sol. – Amsterdam: North – Holland Publ. Co., 1983. – p. 9-31.
9. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. arXiv:1704.04861v1, 2017. – <https://arxiv.org/pdf/1704.04861.pdf>
10. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net*:

Convolutional Networks for Biomedical Image Segmentation.
arXiv:1505.04597v1, 2015. – <https://arxiv.org/pdf/1505.04597.pdf>.