

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)

Метод обробки інформації у завданнях  
розпізнавання та аугментації зображень

(тема)

Виконав:

студент II курсу, групи СПМ-23-1  
Міщенко Д.О.  
(прізвище, ініціали)

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

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

Освітня програма Системне програмування  
(повна назва освітньої програми)

Керівник: проф. Можєв О.О.  
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2025 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-професійна \_\_\_\_\_  
(освітньо-професійна або освітньо-наукова)

Освітня програма \_\_\_\_\_ Системне програмування \_\_\_\_\_  
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

“ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ**

**НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студенту \_\_\_\_\_ Міщенко Дмитру Олександровичу \_\_\_\_\_  
(прізвище, ім'я, по батькові)

1. Тема роботи Метод обробки інформації у завданнях розпізнавання та аугментації зображень

затверджена наказом по університету від “ 22 ” листопада 2024 р. № 1236 Ст

2. Термін подання студентом роботи до екзаменаційної комісії \_\_\_\_\_ 20 січня 2025 р.

3. Вхідні дані до роботи операційна система –Windows або Linux,  
1Гб оперативної пам'яті ПК процесор на 1ГГц, 100Мб свобідної пам'яті на жорсткому диску

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

1) аналіз моделей у завданнях обробки зображень \_\_\_\_\_

2) дослідження методів внесення деформуючих спотворень у зображення \_\_\_\_\_

3) методи штучного розмноження даних шляхом внесення деформуючих спотворень для навчання алгоритмів розпізнавання осіб методом Віоли-Джонса \_\_\_\_\_

4) висновки \_\_\_\_\_

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 13

---

---

---

---

---

---

---

---

---

---

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	Аналіз предметної області	26.11.24-30.11.24	
2	Вибір та обґрунтування методики		
3	Вибір інструментальних засобів	02.12.24-05.12.24	
4	Розробка методу		
5	Проведення експериментів	11.12.24-21.12.24	
6	Оформлення матеріалів кваліфікаційної роботи	23.12.24-03.01.25 04.01.25-07.01.25	
7	Подання кваліфікаційної роботи її попередній захист	08.01.25-11.01.25	
8	Подання кваліфікаційної роботи на рецензування	13.01.25-17.01.25	

Дата видачі завдання 25 листопада 2024 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_  
(підпис)

проф. Можасєв О.О. \_\_\_\_\_  
(посада, прізвище, ініціали)

## РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 61 с., 16 рис., 4 табл., 1 дод., 19 джерел.

### ГАРМОНІЙНА ФУНКЦІЯ, ДЕФОРМУЮЧІ СПОТВОРЕННЯ, РОЗПІЗНАВАННЯ, МОРФІНГ, ОБРОБКА ЗОБРАЖЕНЬ

Метою кваліфікаційної роботи дослідження методу обробки інформації у завданнях розпізнавання аугментації зображень.

У ході виконання кваліфікаційної роботи було проведено аналіз відомих підходів до використання моделей деформуючих спотворень у завданнях обробки зображень.

Запропоновано, обґрунтовано та досліджено метод та алгоритми впливу деформуючих спотворень у рамках завдань обробки зображень.

Синтезовано та досліджено метод аугментації (штучного розмноження) даних шляхом внесення до них деформуючих спотворень у завданнях машинного навчання на прикладі алгоритмів пошуку та виявлення осіб шляхом Віоли -Джонса.

## ABSTRACT

Master's thesis: 61 pages, 16 figures, 4 tables, 1 appendices, 19 sources.

HARMONIC FUNCTION, DEFORMING DISTORTIONS,  
RECOGNITION, MORPHING, IMAGE PROCESSING

The purpose of the qualification work is to study the method of information processing in image recognition augmentation tasks.

During the qualification work, an analysis of known approaches to the use of deforming distortion models in image processing tasks was conducted.

The method and algorithms for the influence of deforming distortions within the framework of image processing tasks were proposed, substantiated and investigated.

The method of augmentation (artificial reproduction) of data by introducing deforming distortions into them in machine learning tasks was synthesized and investigated using the Viola-Jones search and face detection algorithms.

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ .....	8
ВСТУП .....	9
1 АНАЛІЗ МОДЕЛЕЙ У ЗАВДАННЯХ ОБРОБКИ ЗОБРАЖЕНЬ .....	12
1.1 Завдання зіставлення тимчасових рядів .....	14
1.2 Виправлення спотворень, що виникають при оптичній зйомці сцени .....	18
1.3 Виконання морфінгу .....	20
1.4 Облік деформуючих спотворень у моделях розпізнавання образів .....	23
1.5 Використання деформуючих спотворень для аугментації навчальних даних .....	26
2 ДОСЛІДЖЕННЯ МЕТОДІВ ВНЕСЕННЯ ДЕФОРМУЮЧИХ СПОТВОРЕНЬ У ЗОБРАЖЕННЯ.....	28
2.1 Математична модель внесення деформуючих спотворень .....	28
2.2 Внесення деформуючих спотворень у ґратчасті функції дискретних аргументів .....	29
3 МЕТОДИ ШТУЧНОГО РОЗМНОЖЕННЯ ДАНИХ ШЛЯХОМ ВНЕСЕННЯ ДЕФОРМУЮЧИХ СПОТВОРЕНЬ ДЛЯ НАВЧАННЯ АЛГОРИТМІВ РОЗПІЗНАВАННЯ ОСІБ МЕТОДОМ ВІОЛІ- ДЖОНСА .....	31
3.1 Обґрунтування та дослідження алгоритмів розпізнавання зображень .....	31
3.2 Реалізація паралельних версій окремих етапів роботи алгоритму .....	33
3.3 Моделі та алгоритми внесення деформуючих спотворень з метою аугментації навчальною вибірки під час розпізнавання осіб за методом Віоли – Джонса .....	34

3.4 Алгоритм внесення штучною деформації на основі аналізу оптичного потоку з використанням фільтра ентропії.....	42
ВИСНОВКИ .....	49
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ .....	51
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	54

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ  
І ТЕРМІНІВ

ГФ – гармонійних функцій

ДС – деформуючі спотворення

РБФ – радіально-базових функцій

DTW – Distance Time Warping

## ВСТУП

Актуальність. Розвиток сучасних систем управління та прийняття рішень в даний час неможливо без використання алгоритмів обробки інформації, заснованих на застосуванні методів машинного навчання та елементів штучного інтелекту [1-6].

Важливим класом подібних алгоритмів є алгоритми пошуку, виявлення та розпізнавання об'єктів, представлених породжуваними ними цифровими сигналами, зображеннями або багатовимірними ознаковими описами ситуаціях у ході цифровий обробки сигналів та зображень, поряд з наявністю адитивних спотворень різної природи (шумів вимірювання, індивідуальних флюктуацій форми), виникає специфічний вид спотворюючих впливів нелінійного характеру – деформуєчі спотворення (ДС). Зокрема, вплив ДС випадкового. характеру при розпізнаванні цифрових відеосигналів призводить до того, що їхня форма нелінійним чином спотворюється з локальним стиском або розтягуванням щодо тимчасової осі ситуація зустрічається у завданнях кластеризації та вилучення даних (data mining) [13], розпізнавання мови [14], жестів та рухів [15], дослідження біомедичних даних [18], а також під час аналізу потоків об'єктів у системах управління процесом сепарації реального часу .

Не викликає сумнівів також корисність використання моделей деформуєчих спотворень у завданнях аналізу цифрових зображень, наприклад, задачах пошуку та виявлення, а також розпізнавання осіб або інших об'єктів в умовах їх природною мінливості [8]. Світлі ДС можна сприймати як особливий вид випадкових перешкодових впливів , що надають вплив на форму розпізнаваних сигналів або зображень та мають високий рівень кореляції за часом або по простору [9] (у разі аналізу зображень).

Загалом слід відзначити, що специфіка прояви ДС у завданнях обробки інформації, що відображає ті або інші ознакові описи об'єктів, пов'язана з необхідністю побудови та використання моделей деформуєчих спотворень

під час проведення синтезу та аналізу алгоритмів обробки такої інформації та, зокрема, алгоритмів прийняття рішень у задачах пошуку, виявлення та розпізнавання об'єктів на основі сучасних методів статистичної теорії рішень та машинного навчання.

Метою роботи є дослідження методу обробки інформації у завданнях розпізнавання аугментації зображень.

Для досягнення мети в роботі розглядаються та вирішуються наступні завдання:

1 аналіз відомих підходів до використання моделей деформуючих спотворень у завданнях обробки зображень;

2 обґрунтування та дослідження моделей та алгоритмів формування деформуючих спотворень для зображень;

3 синтезувати алгоритми для розпізнавання зображень в умовах деформуючих спотворень;

4 обґрунтувати та дослідити методи штучного розмноження даних на основі внесення до них деформуючих спотворень у завданнях машинного навчання на прикладі алгоритмів пошуку та виявлення осіб шляхом Віоли - Джонса.

Об'єктом дослідження є процес обробки інформації та реалізовані в ньому процеси аналізу зображень.

Предметом дослідження є алгоритми обробки інформації у завданнях розпізнавання та аугментації зображень.

При вирішенні поставлених завдань використовувалися методи математичного аналізу, теорії ймовірностей та математичної статистики, теорії випадкових полів, лінійної алгебри, теорії цифровий обробки сигналів та зображень, а також методи оптимізації, машинного навчання та технології статистичного імітаційного моделювання.

Наукова новизна. Алгоритмічна реалізація методу Віоли-Джонса пошуку та виявлення осіб на зображеннях з використанням технології паралельних обчислень на графічних процесорах на всіх етапах обробки

інформації та її дослідження при використанні для навчання запропонованих моделей аугментації реальних зображень на основі внесення деформуючих спотворень. Моделі та алгоритми аугментації навчальних даних на основі базового набору реальних зображень, що відрізняються застосуванням при внесенні випадковий деформації наступних підходів: розкладання зображення з використанням гармонійних функцій; використання статистичних оцінок зміщень контрольних точок зображень; пошук оптичного потоку за методом Фарнебака з накладенням фільтра ентропії, застосування яких дозволило скоротити обсяг навчальних даних приблизно на порядок за достовірності виявлення осіб, порівнянної з достовірністю, одержуваної на основі повної навчальною вибірки.

## 1 АНАЛІЗ МОДЕЛЕЙ У ЗАВДАННЯХ ОБРОБКИ ЗОБРАЖЕНЬ

У багатьох практично важливих ситуаціях у ході цифровий обробки сигналів та зображень поряд з наявністю адитивного шуму виникає специфічний вид перешкод впливів нелінійного характеру – деформуючі спотворення.

Математичне подання ДІ для будь-якої вихідний функції  $f(x_1, \dots, x_n)$  від  $n$  змінних можна визначити наступним співвідношенням:

$$\begin{aligned} & f(x_1 + r_1(x_1, \dots, x_n), \dots, x_n + r_n(x_1, \dots, x_n)) = \\ & = f(u_1(x_1, \dots, x_n), \dots, u_n(x_1, \dots, x_n)) = g(x_1, \dots, x_n), \end{aligned} \quad (1.1)$$

де  $g(x_1, \dots, x_n)$  – результуюча деформована функція;  $r_i(x_1, \dots, x_n)$ ,  $i = \overline{1, n}$  – функції внесених деформуючих спотворень щодо кожної координаті можуть носити детермінований або стохастичний характер.

ДС можна сприймати як особливий вид випадкових перешкодових впливів, що надають вплив на форму розпізнаваних сигналів та зображень та мають високий рівень кореляції за часом або по простору. На рисунку 1.1 зображено приклад деформації одновимірною безперервний функції виду  $f(x) = \cos(x^2/100)$  за допомогою періодичної функції  $r(x) = 10\sin(\pi(x+13)/18)$  параметри підібрані для наочності уявлення.

Щодо обробки зображень, вплив ДІ відповідає природною мінливості вмісту сцени при нелінійному щодо один одного переміщенні окремих її елементів [2,8]. Такі ситуації виникають, наприклад, при вираженні різних емоцій на обличчі людини, зміні ракурсу під час зйомки або в умовах наявності аберацій оптичної системи, що використовується при реєстрації зображень [9]. Крім того, як деформуючі спотворення можна трактувати природні зміни форми всередині одного класу об'єктів.

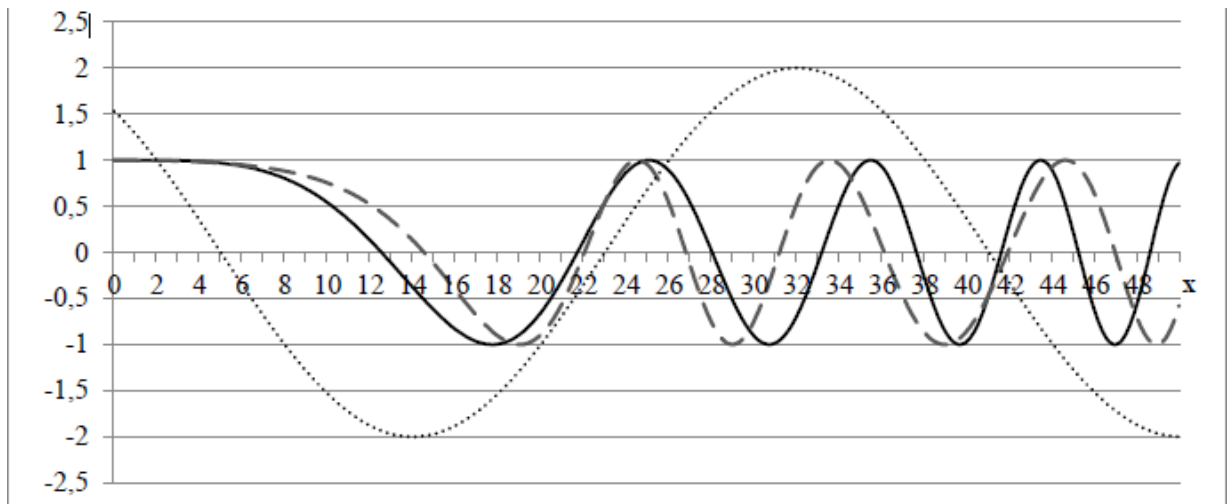


Рисунок 1.1 – Приклад деформації одновимірною безперервний функції

З урахуванням цих факторів переважна більшість попередніх робіт присвячено рішення різноманітних практичних завдань, що мають справу з сигналами або зображеннями, схильних, у тому чи іншому вигляді, що деформує спотворень. До них відносяться роботи, в яких наважувалися наступні завдання:

- зіставлення тимчасових рядів для різного роду сигналів [7];
- виправлення що виникають під час зйомки оптичних спотворень [3,5];
- розрахунок оптичного потоку при послідовній реєстрації зображень у процесі зміни положень та орієнтації об'єктів сцени у різні моменти часу , наприклад , при їх переміщення або при зміні ракурсу зйомки [7];
- виконання «морфінгу» (morphing) - процесу поступового заміщення одного зображення іншим шляхом плавних змін інтенсивностей пікселів, положень та форми схожих у якомусь сенсі частин цих зображень [5];
- облік відмінностей у формі аналізованих об'єктів у різноманітних моделях розпізнавання образів [8, 6], у тому числі при штучному розмноженні навчальних вибірок під час побудови сучасних алгоритмів машинного навчання [5, 7].

Дійсно, використання моделей ДС є важливим для побудови вирішальних правил, що базуються на концепціях машинного навчання. Як уже було зазначено у вступі, пошук та підготовка навчальних даних для

використання часто вимагають значних витрат часу та ресурсів. У ряді випадків збір необхідної кількості образів виявляється складний або навіть неможливий через специфічний характеру предметної області. Таким чином, з урахуванням необхідності забезпечення представницькості навчальних даних одним з можливих підходів тут представляється штучне розмноження обмеженого набору «опорних» образів на основі стохастичних або детерміністських моделей перетворення, що реалізують необхідні деформації [5, 7]. У подібній постановці ДС можуть бути істотним інструментом формування навчальною вибірки, що містить безліч образів аналізованих об'єктів.

Далі будуть більше докладно розглянуті представлені в сучасній літературі різні аспекти прояви деформуючих спотворень та їх обліку у завданнях обробки сигналів та зображень, у тому числі приклади застосування ДС при розмноженні елементів навчальних вибірок у завданнях навчання алгоритмів розпізнавання образів.

### 1.1 Завдання зіставлення тимчасових рядів

Тимчасовим рядом називається тимчасова послідовність дійсних чисел, тобто послідовність, значення якої розглядаються як функція часу  $f(t)$ . Конкретний тимчасовий ряд можна уявити або як реалізацію випадкового процесу із заданими характеристиками, або як детерміновану функцію часу. Із завданнями зіставлення тимчасових рядів, зокрема, їх класифікації, індексування та кластеризації стикаються в самих різних галузях науки і техніки, економіки та медицини – скрізь, де потрібно в тому чи іншому вигляді виробляти порівняння процесів, що розвиваються у часі [7].

Відмінною особливістю тимчасових рядів є те, що крім відмінностей в абсолютних значення у порівнюваних між собою елементів ряду, у тому числі під впливом шумів, для процесів, яких вони отримані, часто характерна різна швидкість протікання яка, більше того, сама може нелінійно

змінюватися в міру їх розвитку, що у багатьох ситуаціях вимагає окремого розгляду або модифікації наявних методів роботи зі статичними даними [9, 11].

Методи роботи з тимчасовими даними умовно можна розділити на два основні типу: методи, що мають справу з вихідними тимчасовими рядами під час проведення порівняння та методи, засновані на їх попередньому перетворення, наприклад, вилучення примітивів або поданні даних тимчасовий послідовності в рамках якоїсь відомої моделі та наступної оцінки та порівняння її параметрів [9, 13].

Методи, засновані на зіставленні вихідних тимчасових рядів засновані на крапковому розрахунку відстаней між вихідними значеннями тимчасових рядів. Розраховані відстані далі використовуються як базові значення при кластеризації, класифікації або індексування в залежності від вихідної завдання. Два часто використовувані методи зіставлення з цією категорії, що отримали найбільш широке висвітлення у літературі [9-11, 13]:

- розрахунок евклідова відстані (Euclidean Distance в англійській літературі є також модифікації, які використовують та інші норми) – заснований на прямому зіставленні точок тимчасових рядів без урахування їх можливих зрушень вздовж тимчасової осі;

- алгоритм динамічної трансформації тимчасовий шкали (Distance Time Warping або DTW) – метод, заснований на динамічному програмуванні та дозволяє знаходити відповідності між запізнюючими або випереджаючими один одного ділянками тимчасових рядів [8].

На рисунку 1.2 показано у порівнянні результати зіставлення двох однаковою форм, але зміщених щодо один одного вздовж тимчасової осі сигналів методами Euclidean Distance та DTW. Видно, що для методу, заснованого на розрахунку евклідова відстані, отримана значення заходи близькості (часто це сума відстаней між зіставленими точками) не враховує тимчасового зсуву і тому показує меншу схожість порівнюваних тимчасових рядів, ніж DTW. Стоїть відзначити, що проблема приведення один до одного

відліків двох нерівномірно дискретизованих тимчасових рядів легко вирішується за допомогою інтерполяції даних [8, 9, 19].

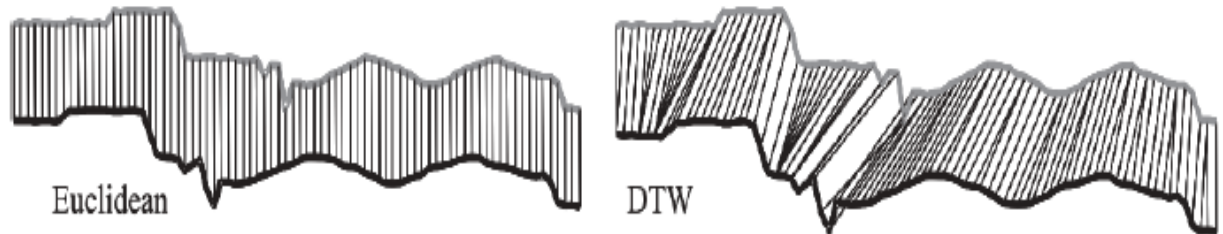


Рисунок 1.2 – Результати зіставлення 2-х тимчасових рядів методами Euclidian Distance (ліворуч) та DTW (праворуч)

Принцип роботи DTW (рисунок 1.3) у базовій своїй варіації заснований на побудові матриці відстаней між елементами двох тимчасових рядів  $Q$  і  $S$  та пошуку найкоротшого шляхи динамічної трансформації  $W$  у ній.

Такий найкоротший шлях часто шукається за допомогою динамічного програмування: рішення про те, який елемент шляху  $W$  буде наступним, приймається щоразу на основі локального порівняння відстаней між поточним елементом шляху  $W$  та найближчими до нього елементами матриці зіставлення. Існують різні модифікації даного методу, а також інші підходи, засновані на схожих принципах динамічного програмування, але дослідження показують, що краще всіх решти себе показують вихідний метод динамічної трансформації тимчасовий шкали, а також для великих тимчасових рядів, метод, заснований на простому розрахунку евклідова відстані, яка до того ж має набагато меншу обчислювальну складність, ніж DTW [10]. До недоліків обох методів можна віднести їх чутливість до шуму [11].

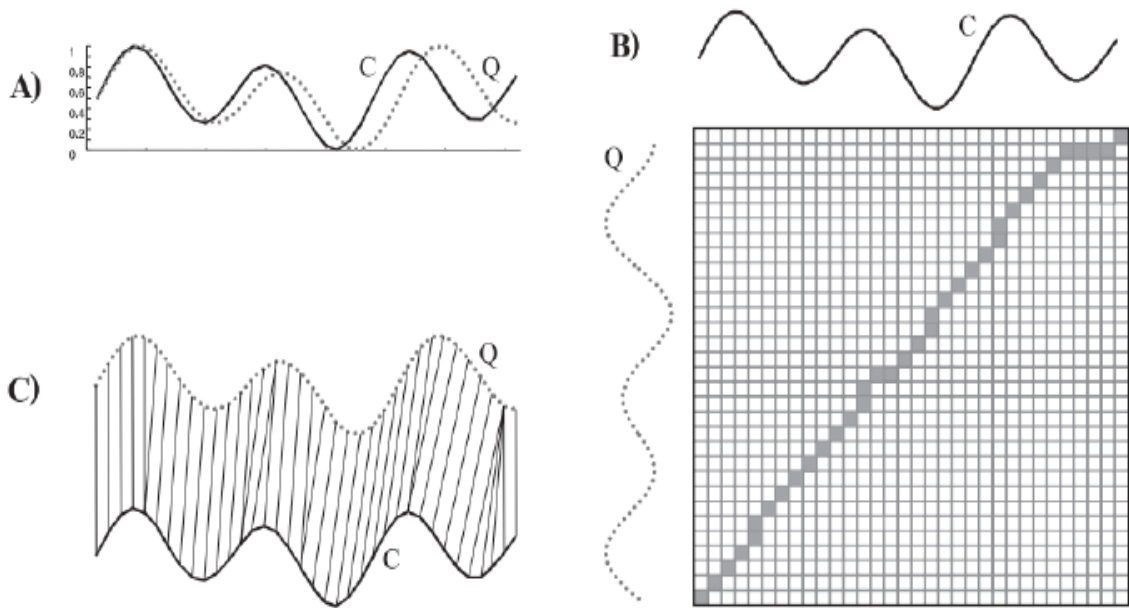


Рисунок 1.3 – Ілюстрація принципу роботи методу DTW: А) два вихідні тимчасових ряду, В) матриця зіставлення відстаней із виділеним у ній знайденим найкоротшим шляхом динамічної трансформації, С) отриманий результат зіставлення

Методи, засновані на попередньому перетворенні тимчасових рядів. В основі таких методів лежить перетворення даних вихідних тимчасових рядів з метою зменшити розмірність (наприклад, метод основних компонентів), виділити примітиви (вейвлет-перетворенн, масштабно інваріантне перетворення SIFT) або зробити перехід у частотну область [9-11, 13]. Також до таких методів можна віднести подання тимчасових рядів у рамках яких-небудь відомих моделей, при використанні яких замість самих вихідних даних виробляється оцінка та порівняння параметрів отриманих моделей (наприклад, приховані марківські або інші статистичні моделі) [11, 13, 20]. Можливі поєднання описаних підходів.

Поширеними також є методи, що здійснюють після виділення примітивів з подучастків тимчасових рядів побудова якогось їх «алфавіту» на основі результатів кластеризації та/ або показників їх частотного розподілу та виробляючи подальшу роботу вже з «символами» «алфавіту», що вийшов замість вихідних тимчасових рядів [10-13, 19]. До таких методів відноситься,

наприклад, SAX - метод символної апроксимації часового ряду.

Порівняльні дослідження та огляди [10, 11] показують, що вибір конкретного підходу до використання залежить від багатьох факторів, у тому числі від конкретної завдання (кластеризація, індексування або класифікація), самих даних, а також бажаного співвідношення швидкості та якості роботи вийшли алгоритмів і в цілому в середньому багато з представлених методів видають приблизно рівні показники точності своєю роботи.

Тим не менш, існуючі публікації, присвячені роботі з тимчасовими рядами, здебільшого своєму обходять стороною саме питання аналізу впливу деформуючих спотворень на процес їх розпізнавання і носять, як правило, суто евристичний характер.

## 1.2 Виправлення спотворень, що виникають при оптичній зйомці сцени

Через особливості побудови приладів зйомки з лінзової оптикою, напрямки пучків світла, що проходять через них, часто схильні різним відхиленням (аббераціям), та результати зйомки, таким чином, можуть містити в собі різноманітні геометричні спотворення (дисторсію), які у певних ситуаціях досягають великих величин та у ряді випадків і прикладних завдань вимагають обліку та виправлень [2, 9]. Серед інших часто виділяють дві великих категорії таких спотворень: радіальні (рисунок 1.4а) та тангенціальні (рисунок 1.4б) спотворення [3,5].

Причиною радіальних спотворень (рисунок 1.4а) є викривлення світлових променів на краях опуклих лінз в залежності від їх відстань до оптичного центру. Більшою ступеня даний ефект проявляється для ширококутних, а також лінз малого розміру. Це викривлення призводить до сумування або розсіювання світлових променів в залежності від відстані до оптичного центру і, таким чином, до стягування або розтягування зображення на його кутах (внутрішній та зовнішній контур відповідно на

рисунок 1.4а). Тангенціальні спотворення (рисунок 1.4б) на додаток до радіальних виникають у разі неколінеарності площин розташування лінз один щодо одного або лінз та оптичного сенсора.

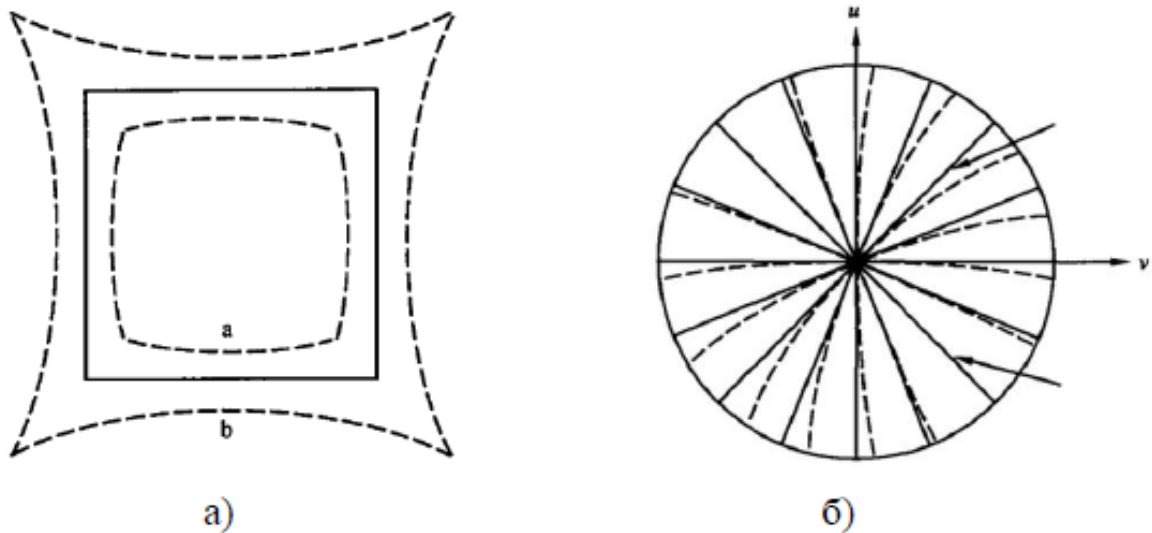


Рисунок 1.4 – Приклади можливих ефектів, що вносяться абераціями оптичних систем зйомки з лінзами, для радіальних (а) та тангенціальних (б) типів спотворень (переривчасті лінії показують можливі зміщені позиції відповідних суцільних ліній залежно від типу дисторсії)

Параметри даних спотворень поряд з іншими особливостями оптичного приладу оцінюються під час процедури калібрування полягає у використанні спеціальних зображень різного виду патернів (рисунок 1.5а) в оцінці параметрів відображення координат вихідної тривимірної сцени у двовимірні координати зафіксованого приладом зображення. Це, наприклад, на основі проставлення відповідностей між вихідними та отриманими координатами контрольних точок патерну (рисунок 1.5б).

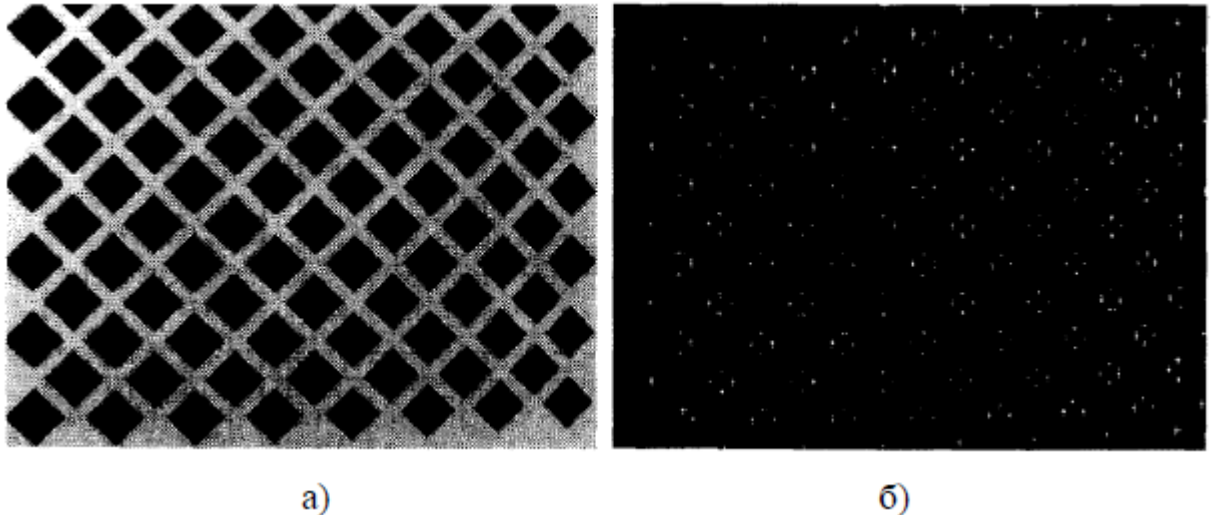


Рисунок 1.5 – Приклад патерну , який використовується під час процедури калібрування (а) та обрані на його основі контрольні точки (б)

Для оцінки параметрів дісторсії часто використовують представлені у вигляді поліномів різноманітні моделі спотворень, однієї з яких є модель Брауна-Конраді. Для визначення значень цих поліноміальних коефіцієнтів часто використовують різні методи оптимізації.

Таким чином, відмінною рисою виникнення деформуючих спотворень , що вносяться оптичними приладами зйомки, що базуються на використанні лінз популярність моделей, що вносяться спотворень, що засноване на самій природі їх прояви, пов'язаної з особливостями побудови даних приладів використовують при них обліку та виправленні ще на самих ранніх етапах отримання та обробки зображень, і, отже, дані спотворення в подальшому, як правило, не грають суттєвої ролі.

### 1.3 Виконання морфінгу

Морфінг (morphing) – процес поступового заміщення одного зображення іншим шляхом плавних змін значень інтенсивностей їх пікселів одночасно зі зміщеннями положень та модифікаціями форми схожих у якомусь сенсі частин цих зображень. Знаходить своє застосування в області створення ефектів анімації в комп'ютерній графіці, а також у завданнях

візуалізації, наприклад, у геоінформаційних системах [5] або медицині при побудові узагальнених результатів комп'ютерної томографії, зроблених на основі кількох розрізів [5-7].

На практиці, для створення ефекту плавного переходу до комп'ютерної анімації процес морфінгу полягає у побудові ряду проміжних зображень від вихідного до цільового, кожне з яких є сумішшю цих двох з більшою ступенем впливу того, до якого воно ближче розташований (рисунок 1.6).

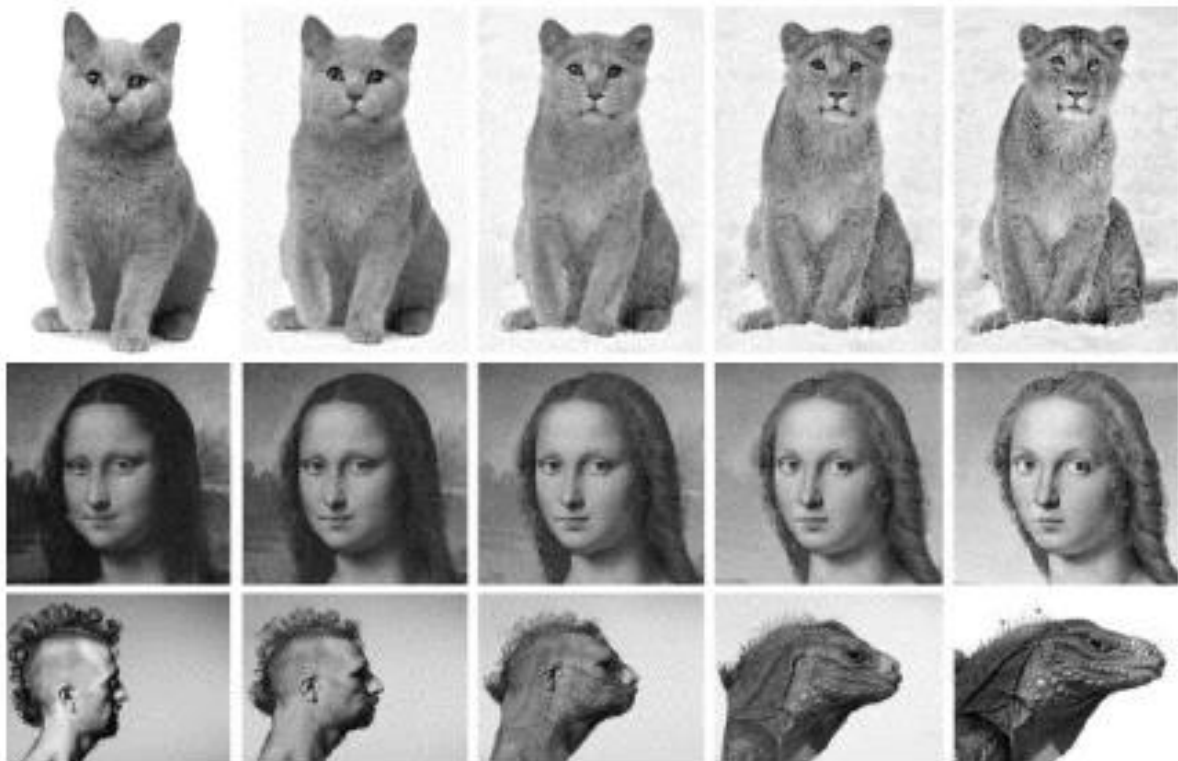


Рисунок 1.6 – Приклади морфінгу

При виконанні морфінгу деформація вихідних зображень здійснюється для того, щоб при змішуванні значень інтенсивностей їх пікселів під час побудови зображень проміжних етапів (рисунок 1.7) кожне з них виглядало природно, і не виникало ефектів «просвічування» одного зображення крізь інше. Така деформація виробляється для обох вихідних зображень з метою отримати відповідну поточному проміжному етапу деформовану версію кожного з них:

$$\begin{aligned} I_{\alpha 1}(x, y) &= I_1(x + \alpha v_x(x, y), y + \alpha v_y(x, y)), \\ I_{\alpha 2}(x, y) &= I_2(x - (1 - \alpha) v_x(x, y), y - (1 - \alpha) v_y(x, y)), \end{aligned} \quad (1.2)$$

де  $I_{\alpha 1}$  і  $I_{\alpha 2}$  – відповідні поточному проміжному етапу морфінгу деформовані зображення, отримані на основі вихідних зображень  $I_1$  і  $I_2$  ;

$v = (v_x, v_y)$  – розрахована тим чи іншим чином функція деформації;

$\alpha$  – ваговий коефіцієнт, що визначає близькість одержуваних деформованих зображень вихідним, на яких вони засновані.

Для отримання функції деформації, що забезпечує плавність та «природність» процедури морфінгу, спочатку необхідно визначити відповідності між ділянками вихідних зображень. Робиться це за допомогою різних методів, які можна поділити на ручні та автоматичні.

Серед ручних методів поширені: розмітка контрольних точок, накладення на кінцеві зображення сіток, вузли яких відповідають один одному, а також зображення на підданих морфінгу зображеннях пар прямих, що відображаються один в одного або кривих ліній та в принципі контурів вільною форми, які часто при подальшій обробці перетворюються на ті ж самі групи контрольних точок [4, 5, 6]. Плюсом останнього підходу є гнучкість у можливостях отримання необхідною послідовності кадрів при морфінгу на противагу до трудомісткості ручної вказівки відповідностей на вихідні зображеннях.

Частина автоматичних методів у певних ситуаціях (близькі за виглядом зображення одного і того ж чи схожих об'єктів) використовують техніки пошуку відповідностей на зображеннях на основі розрахунку оптичного потоку [8] та в цілому з області реєстрації зображень [7]. Інші – базуються на результатах роботи різноманітних алгоритмів сегментації або розпізнавання і за їх допомоги визначають позиції та кордони відповідних один одному ділянок підлягають деформації зображень з будь-якої конкретної предметний області, наприклад, під час роботи з особами або медичними зображення [4, 5]. Тим не менш, саме завдання знаходження загального та відповідного в

більшості випадків підходу до автоматичного пошуку відповідностей на підлягаючих морфінгу зображеннях і потім на цій основі розрахунку значень функції деформації ще чекає свого рішення .

Широко поширені також автоматизовані підходи, які на додаток до автоматичного пошуку відповідностей на зображеннях враховують необов'язкові ручні вказівки оператора [4, 8, 9].

#### 1.4 Облік деформуючих спотворень у моделях розпізнавання образів

Потенційна користь від обліку деформуючих спотворень у моделях розпізнавання образів пов'язана з тим, що об'єкти реального світу, відображені на зображеннях , мають тенденцію до того, щоб переміщатися, змінювати розміри, повертатися та спотворюватися від кадру до кадру [8]. Існуючі моделі розпізнавання образів , що враховують всередині себе або засновані на обліку таких спотворень, можна розділити на дві великі групи:

- еластичні моделі розпізнавання образів , що порівнюють зображення не тільки за вмістом, а й за близькістю розташування їх окремих частин до певних позиціям на зображенні або щодо один одного [6];

- статистичні моделі, що аналізують близькість зображень на основі знаходження найбільш ймовірною деформації, необхідної для перетворення їх одне одного [7, 8].

Еластичні моделі розпізнавання образів загалом випадку засновані на декомпозиції функції прийняття рішення, що аналізує якусь частину розпізнаваного зображення, як суми  $d(l)$  .

$$d(l) = d_c(l) + d_p(l), \quad (1.3)$$

де  $d_c(l)$  – функція , що здійснює порівняння близькості по вмісту (це може бути класифікатор будь -якого виду або пряме порівняння інтенсивностей пікселів);

$d_p(l)$  – функція, що здійснює порівняння близькості за розташуванням

(часто розраховується на основі різних моделей деформації, наприклад, сплайну тонкою пластини або інших обмежень, але це може бути і просто перевірка на входження в допустиму область щодо позиції, що приймається за еталон [12]), а представляє собою власне аналізовану область з безлічі підлягають дослідженню областей зображення.

Підсумкове рішення при цьому приймається на основі роботи всього набору функцій за кількістю аналізованих частин зображення, які можуть бути організовані у вигляді лінійної комбінації [14], прихованої марківської моделі [12], графа [15] чи дерева [16] трьох моделей характерно, що розрахунок виробляється на основі відносних позицій аналізованих частин зображень, що для графа та дерева відповідає параметрам їх ребер (рисунок 1.7).

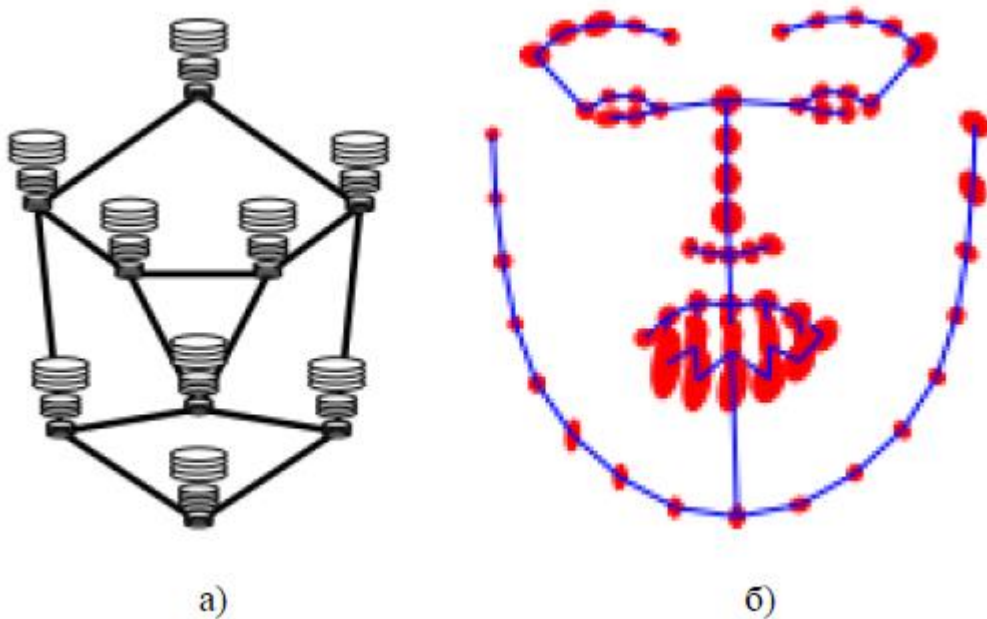


Рисунок 1.7 – Графічні уявлення еластичних моделей розпізнавання на основі  
(а) графа та (б) дерева

Завдання, які вирішуються такими класифікаторами можна розділити на зіставлення зображень один з одним та пошук об'єктів на зображеннях [12]. Для останніх часто характерна велика трудомісткість вихідний розмітки зображень навчальною вибіркою [5, 6], але є і моделі, які самі визначають позиції аналізованих підобластей зображень шляхом пошуку їх оптимального

в якомусь сенсі положення, в тому числі і на основі методів динамічного програмування [2,4].

Дані моделі засновані на оцінці ймовірностей деформацій , необхідні перетворення одних зображень на інші. Робиться це в основному параметричними методами на базі будь-якої моделі, у тому числі за допомогою обмежень , що накладаються на допустимі деформації. Для зменшення простору розмірності ознак серед інших використовується метод основних компонентів [2, 7, 8].

Такі підходи у своїй основі призначені для зіставлення зображень в умовах припущення сталості яскравості, але, проте, можливі зміни у освітленості у яких можуть бути представлені у вигляді адитивного шуму [2, 7].

Навчання алгоритмів розпізнавання тут виробляється шляхом вибору зображення, яке буде використано як референтного, та оцінки функції щільності ймовірності внутрішньокласових деформацій на основі відомих перетворень референтного зображення на всі інші зображення даного класу. При розпізнаванні, таким чином, шукається найбільш ймовірна деформація серед можливих деформацій розпізнаваного зображення у зображення, які були обрані як референтних [2, 7, 8].

Загалом загальним висновком для всіх методів розпізнавання, заснованих на обліку можливих деформуючих спотворень розпізнаваних зображень, є те, що вони націлені на імітацію властивості мінливості об'єктів реального світу і, таким чином, на практиці дозволяють досягти підвищення показників розпізнавання, а також в окремих випадках, знизити необхідний обсяг навчальних вибірок, проте, при цьому, такі засновані на обліку ДС моделі є більше трудомісткими для розрахунку.

## 1.5 Використання деформуючих спотворень для аугментації навчальних даних

Для аугментації – штучного розмноження елементів навчальних вибірок – поряд з іншими методами, що працюють як із вихідним поданням розпізнаваних даних (афінні трансформації, нарізка зображень, моделювання шуму та освітленості) [5, 8], так і в просторі отриманих на них основі ознак (Smote, DBSmote та інші) [9], також використовуються та деформуючі спотворення [7].

Сама ідея штучного збільшення обсягу навчальних вибірок пов'язана з частими вимогами великого обсягу навчальних даних у сучасних алгоритмів розпізнавання та складністю або навіть неможливістю в окремих випадках їх набору та підготовки у потрібному кількості [5, 6]. Зокрема, при роботі з нейронами мережами такий підхід дозволяє боротися з їх перенавчанням [5, 9]. Ще одна проблема, яку можна вирішити подібним чином – це незбалансованість навчальних вибірок [9]. Подібні ситуації особливо часто виникають при роботі з медичними зображеннями, де набір відповідних прикладів для навчання спочатку обмежений загалом або для конкретного класу образів [6]. Дана область досліджень останнім час отримала додатковий поштовх разом із поширенням глибоких нейронних мереж, для яких ідея аугментації навчальних даних є невід'ємною частиною їх процесу навчання [5, 6, 7].

Існують різні підходи до внесення деформуючих спотворень в елементи навчальною вибірки. У самих простих випадках вони ґрунтуються на застосуванні як ДС випадкових полів, згенерованих за якимсь законом розподілу, наприклад, при розпізнаванні рукописних символів з метою змодельовати випадкові осциляції м'язів при листі [5].

Під час навчання алгоритмів розпізнавання осіб для розмноження елементів навчальною вибірки також застосовуються методи, що базуються на формуванні тривимірних моделей голови на основі двовимірних

зображень осіб та побудову проєкцій отриманих моделей з метою змоделювати різні ракурси зйомки або повороти голови в кадрі (тут також застосовуються деформації самої тривимірній моделі з метою отримати більше відповідності конкретній особі) [5, 6].

Деформуючі спотворення використовуються у тому числі і при розмноженні навчальних вибірок у завданнях класифікації тимчасових рядів [5, 7].

Результати експериментів у цій області говорять про те, що розмноження навчальних вибірок на основі внесення до них деформуючих спотворень у багатьох випадках збільшує точність роботи алгоритмів розпізнавання [7]. Особливо зазначається, що найкращих результатів можна досягти тільки у випадку наявності інформації про параметри можливі деформуючих спотворень, що описують відмінності між елементами навчальною вибірки [5]. Таким чином, на даний момент перспективними є дослідження , присвячені знаходження параметрів таких спотворень, для чого представляється корисним використання підходів та розробок з областей морфінгу та розрахунку оптичного потоку.

## 2 ДОСЛІДЖЕННЯ МЕТОДІВ ВНЕСЕННЯ ДЕФОРМУЮЧИХ СПОТВОРЕНЬ У ЗОБРАЖЕННЯ

### 2.1 Математична модель внесення деформуючих спотворень

Як вже зазначалося раніше, математичне подання процесу внесення деформуючих спотворень для будь-якої вихідний функції  $f(x_1, \dots, x_n)$  від  $n$  змінних визначається через адитивне додавання значення функції деформації до вихідних значенням цих змінних:

$$\begin{aligned} f(x_1 + r_1(x_1, \dots, x_n), \dots, x_n + r_n(x_1, \dots, x_n)) = \\ = f(u_1(x_1, \dots, x_n), \dots, u_n(x_1, \dots, x_n)) = g(x_1, \dots, x_n), \end{aligned} \quad (2.1)$$

де  $g(x_1, \dots, x_n)$  - результуюча деформована функція;

$r_i(x_1, \dots, x_n)$ ,  $i = 1, n$  - безперервні функції внесених деформуючих спотворень щодо кожної координат можуть носити детермінований або стохастичний характер.

Раніше, на рисунку 1.1, було зображено приклад деформації одновимірною безперервний функції виду  $f(x) = \cos(x^2/100)$ , що ілюструє загальний характер виникаючих спотворень.

При реалізації процесу внесення ДС на основі (2.1) та (2.2) виникає ряд особливостей [6]. з них полягає в тому, що для того, щоб значення деформованою функції  $g(x) = f(u(x))$  існувало область її визначення не повинно виходити за межі області визначення вихідний функції  $f(x)$  Інакше, якщо для  $f(x)$  вектор  $x \in \Omega_x$ , і вектор  $u(x) \in \Omega_x$ .

Для забезпечення виконання умови  $u(x) \in \Omega_x$  застосовні різні підходи. Один із них полягає в тому, щоб при виході будь-якої компоненти координат в  $x + r(x) = u(x)$  з (2.2) за кордону вихідний області визначення  $\Omega_x$  як значення даної компоненти  $u(x)$  при розрахунку  $f(u(x))$  підставляти граничне значення відповідної координати вихідний функції  $f(x)$ . Перевага даного підходу полягає в тому, що при його застосуванні не модифікуються значення

функції деформації  $r(x)$  у разі, коли  $x + r(x) = u(x)$  знаходиться в межах  $\Omega_x$ .

## 2.2 Внесення деформуючих спотворень у ґратчасті функції дискретних аргументів

У випадку завдання вихідній функції як функції дискретних аргументів (як ґратчастої функції) головна особливість [6, 7] процесу внесення ДІ на основі (2.1) та (2.2) полягає в наступному. Якщо вихідна функція  $f(x)$  є ґратчастою функцією дискретних аргументів, тобто, задана на багатовимірній дискретній сітці (наприклад, у випадку внесення ДС в цифрові зображенн, пряме застосування (2.1) та (2.2) неможливо. У цьому випадку для забезпечення можливості модифікації аргументів  $f(x)$  при внесенні до них ДІ довільного характеру необхідно попередньо виконати її інтерполяцію та подання у вигляді функції  $f(x)$  речових змінних. При цьому загальна процедура буде відповідати схемою, запропонованої автором у роботі [4] та поданої на рисунку 2.1.

Слід також відзначити, що в даній схемою для вихідній функції  $f(x)$  не враховується можливий дискретний характер за рівнем, тобто, вважається, що розрядність уявлення даних велика та нею можна знехтувати, вважаючи, що значення функції мають речові значення. У деяких випадках це не так, наприклад, якщо функція описує напівтонові зображення з рівномірним квантуванням на 256 рівнів. Тоді описані у схемі перетворення здійснюються по відношенню до вихідної квантованою по рівнем функції, і з враховуючи те, що при деформації можливо отримання проміжних значень між вихідними рівнями, проводиться додатковою квантування деформованою функції та її приведення до стандартної сітці значень.

Відповідно до вищеописаних міркуваннями на користь реалізації алгоритму внесення ДС в ґратчасті функції дискретних аргументів було проведено порівняльне дослідження різних способів інтерполяції, у тому числі двох широко відомих і часто застосовуваних на практиці: лінійної та

кубічної інтерполяції [7], специфіка застосування яких для вирішуваної Завдання наведено автором в [6]. Крім того, для вирішення цієї завдання була досліджено можливість застосування радіально-базових функцій (РБФ). Даний підхід потенційно дозволяє вносити безперервні спотворення будь-якого виду.

В одновимірному у разі його реалізації вихідна гратчаста функція  $f(x_i)$ ,  $i = 1, N$  поділяється на відрізки, кожен довжиною  $n$ . При цьому крайні точки сусідніх один з одним відрізків є для них загальними. У середині відрізків використовуються локальні координати. Після цього для кожного з відрізків шукається апроксимація безперервний функцією виду

$$\tilde{f}(x) = \sum_{i=1}^n a_i \exp(-(x - \hat{x}_i)^2 / (2\sigma^2)), \quad (2.3)$$

де  $\sigma$ - Параметр впливу РБФ- функцій, а  $a_i$  - невідомі коефіцієнти. Елементів суми відповідають центри РБФ - функцій у точках  $x_i$  - по одній функції на точку вихідного дискретного набору відліків всередині даного відрізка.

### 3 МЕТОДИ ШТУЧНОГО РОЗМНОЖЕННЯ ДАНИХ ШЛЯХОМ ВНЕСЕННЯ ДЕФОРМУЮЧИХ СПОТВОРЕНЬ ДЛЯ НАВЧАННЯ АЛГОРИТМІВ РОЗПІЗНАВАННЯ ОСІБ МЕТОДОМ ВІОЛИ-ДЖОНСА

Побудова алгоритмів розпізнавання реальних зображень на основі методів машинного навчання за наявності необхідних навчальних даних зазвичай здійснюється за стандартними схемами, що диктуються парадигмою використовуваного підходу. Однак часто навчальних даних недостатньо, що може бути викликано не тільки об'єктивними обмеженнями при їх отриманні, але і великий трудомісткістю розмітки навчальних вибірок. У цій ситуації при побудові алгоритмів розпізнавання реальних зображень одним з перспективних представляється підхід, заснований на аугментації – штучному розмноженні навчального безлічі даних по кожному з класів.

Рішення завдання в такий спосіб має сенс, коли вплив ДС можна розглядати як реалізації деякої випадкової функції, що описує відмінності між елементами розмножується навчальною вибірки та які можна згенерувати за відомої статистичної моделі деформуючих спотворень. Можна очікувати, що подібний спосіб побудови описів класів добре вбудовується у завдання машинного навчання алгоритмів на основі навчальних даних, отриманих з набору реальних зображень.

Аналіз можливостей застосування такого підходу був виконаний нами на прикладі алгоритму розпізнавання осіб за методом Віоли - Джонса. Цей метод був обраний як інструменту дослідження як один з самих надійних та популярних на даний момент детекторів фронтальних осіб [8].

#### 3.1 Обґрунтування та дослідження алгоритмів розпізнавання зображень

Метод Віоли-Джонса докладно описаний та заснований на застосуванні цілого безлічі взаємопов'язаних підходів та технологій. У центрі цього методу лежить ідея побудови каскадів, що запускаються один за одним, що

реалізують точні або, як ще кажуть, сильні класифікатори шляхом комбінування їх на основі так званого «бустингу» з наборів інших менш точних або слабких класифікаторів-ознак Хаар. Застосування навченого детектора виробляється шляхом переміщення та масштабування скануючого вікна.

Для створення послідовною версії методу була використана середовище Matlab. Цей вибір обумовлений наявністю в ній вже безлічі реалізованих математичних операторів і функцій, часто відсутні в інших мовах програмування. Також середовище Matlab містить весь необхідний інструментарій, застосовний для проведення аналізу роботи реалізації методу.

Метод Віюлі-Джонса можна розділити на два незалежні алгоритми: навчання класифікатора та аналіз зображень готовим класифікатором. Сам алгоритм навчання складається з кількох наступних один за одним фаз (рисунок 3.1).

Підготовка ознак. собою фазу створення набору ознак Хаара. Перебір всіх типів ознак з урахуванням їх можливих розмірів та положень всередині скануючого вікна розміром  $24 \times 24$ .

Підготовка та аналіз позитивних екземплярів вибірки. Уявляє собою фазу підготовки навчальною вибіркою зображень осіб та отримання первинних результатів її аналізу повним набором ознак. Три фази, що залишилися виконуються в циклі до тих пір, поки не будуть отримані необхідні показники достовірності роботи алгоритму розпізнавання.

Збір та аналіз негативних екземплярів вибірки. На цій фазі [4, 7] проводиться збір та первинний аналіз повним набором ознак Хаара необхідної кількості негативних екземплярів вибірки. Як негативних екземплярів вибірки використовується вміст скануючих вікон, що викликали хибні спрацьовування останнього шару каскаду, що є на даний момент класифікаторів під час обходу їм у робочому режимі свідомо не містять особи зображень.

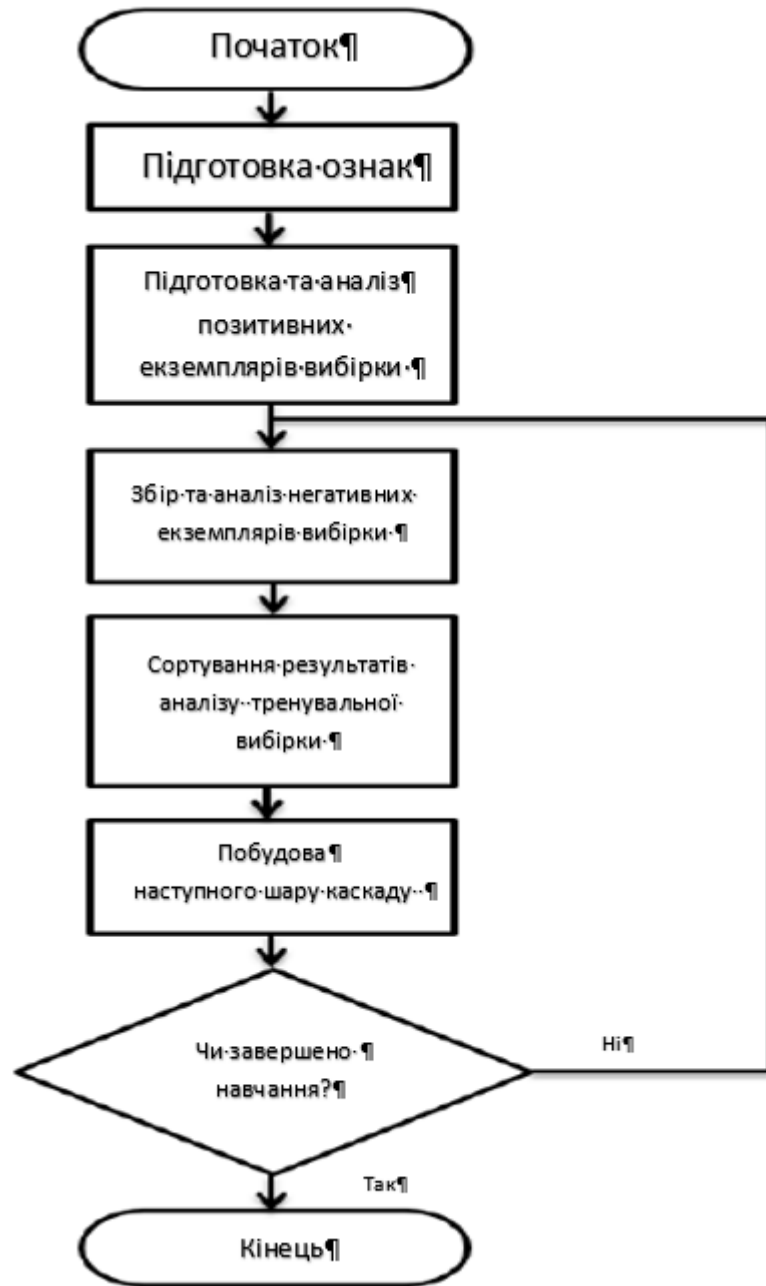


Рисунок 3.1 – Схема алгоритму навчання за методом Віюлі-Джонса

Сортування результатів аналізу тренувальній вибірки. На цьому кроці алгоритму навчання для кожної окремої ознаки Хаара виробляється об'єднання та сортування за зростанням первинних результатів обробки разом позитивних (зображень осі) та негативних екземплярів навчальною вибіркою. Сортування виробляється для прискорення процедури визначення порогового значення у процесі навчання слабкого класифікатора під час

бустингу на наступному фазі роботи алгоритму.

Побудова наступного шару каскаду здійснюється завдяки шару формування його в ході процедури бустингу у вигляді лінійної комбінації найкращих з слабких класифікаторів, навчених на поточному наборі позитивних та негативних екземплярів навчальною вибірки.

### 3.2 Реалізація паралельних версій окремих етапів роботи алгоритму

Для створення паралельних версій окремих етапів роботи алгоритму як механізму здійснення паралельних обчислень на графічному процесорі була обрано архітектура CUDA, як єдино доступна з середі Matlab. Під час проведення всіх експериментів з оцінки продуктивності алгоритмів як випробувального стенду використовувався комп'ютер зі наступними характеристиками: процесор Intel Core i5-3210M 2.5 ГГц; 6 ГБ DDR3 оперативної пам'яті; відеокарта – NVIDIA GeForce GT 640M з 2 ГБ пам'яті і 384 ядрами CUDA ( має можливість оперування з числами з плаваючою комою подвійний точності - double ); ОС - 64-х бітна версія Ubuntu 14.04.

На етапі навчання були розпаралелені алгоритм збору первинних результатів обробки ознаками екземплярів навчальною вибірки, як позитивних, і негативних, і навіть алгоритм навчання слабких класифікаторів.

Створення паралельною версії алгоритму збору первинних результатів обробки. Алгоритм розрахунку первинних результатів обробки ознаками екземплярів навчальною вибірки добре піддається розпаралелювання через те, що аналіз окремим ознакою окремого екземпляра вибірки виробляється незалежно від інших.

Через великий обсяг даних та обмеження ресурсів оперативною пам'яттю даний алгоритм вирішено було розбити на кілька етапів з поступовим збереженням результатів обробки на диск з метою наступного їх використання під час навчання (рисунок 3.2). При цьому кожен файл для

певної кількості ознак з загального набору (це число визначається доступним обсягом оперативною пам'яті) записуються результати первинного аналізу ними всіх наявних екземплярів навчальною вибірки. Виміри часу роботи показали, що паралельна версія алгоритму первинного аналізу навчальною вибірки повним набором ознак Хаара за швидкістю роботи перевершує послідовну його версію більше ніж у сто разів (таблиця 3.1).

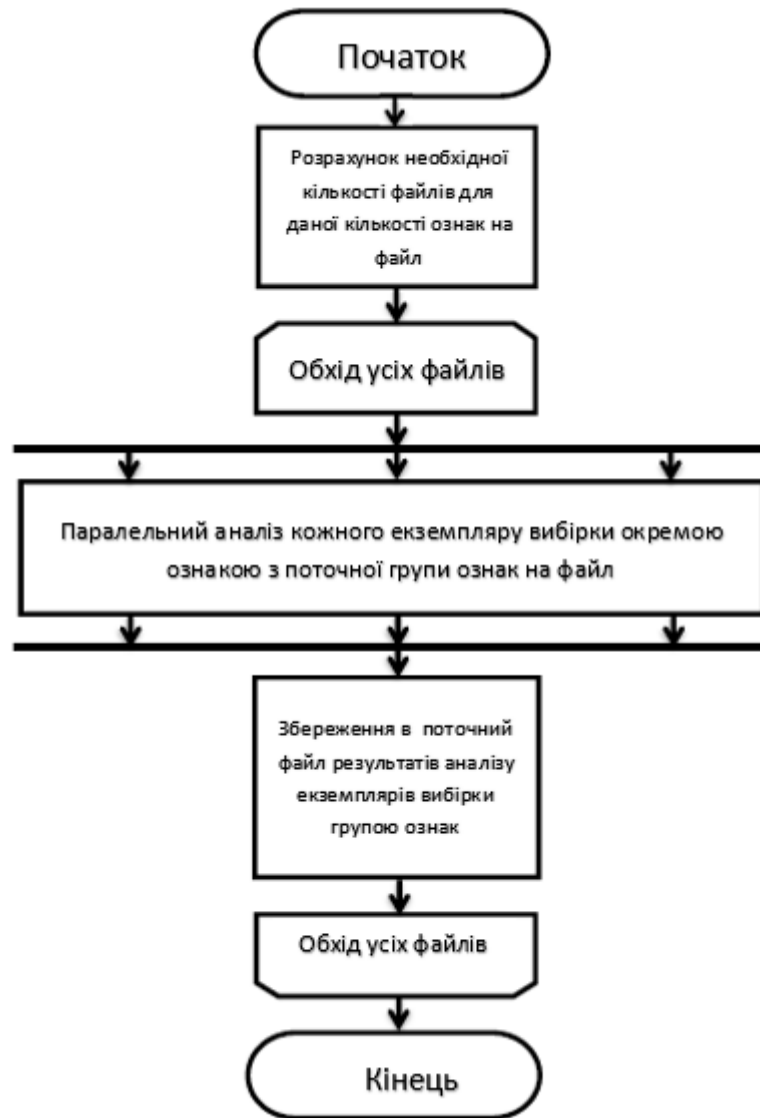


Рисунок 3.2 – Схема паралельної версії алгоритму збору первинних результатів обробки екземплярів навчальною вибіркою повним набором класифікаторів

Таблиця 3.1 – Показники часу роботи послідовною та паралельною версій алгоритму збору первинних результатів аналізу екземплярів навчальною вибірки повним набором ознак Хаара

Число оброблених екземплярів тренувальній вибірки	Послідовна версія	Паралельна версія
2000	3 год 23 хв	1 хв 33 с
4000	5 год 13 хв	3 хв 30 с
6000	8 год 38 хв	4 хв 26 с

Створення паралельною версії алгоритму навчання слабкого класифікатора. На цьому етапі прискорення було досягнуто за рахунок паралельного навчання слабких класифікаторів. Загальна схема роботи вийшов алгоритму наведено рисунку 3.3. У цьому випадку через обмеження пам'яті кількість одночасно працюючих потоків невелико і зі збільшенням обсягу навчальною вибірки їх кількість падає, що надає відповідне вплив на продуктивність алгоритму (таблиця 3.2).

Таблиця 3.2 – Показники часу виконання послідовної та паралельних версій алгоритму навчання слабких класифікаторів та вибору кращого з них у ході процедури бустингу

Число екземплярів навчальною вибірки	Послідовна версія	Паралельна версія
2000	0 хв 52 с	0 хв 24 с
4000	2 хв 04 с	1 хв 09 с
6000	2 хв 47 с	1 хв 53 с

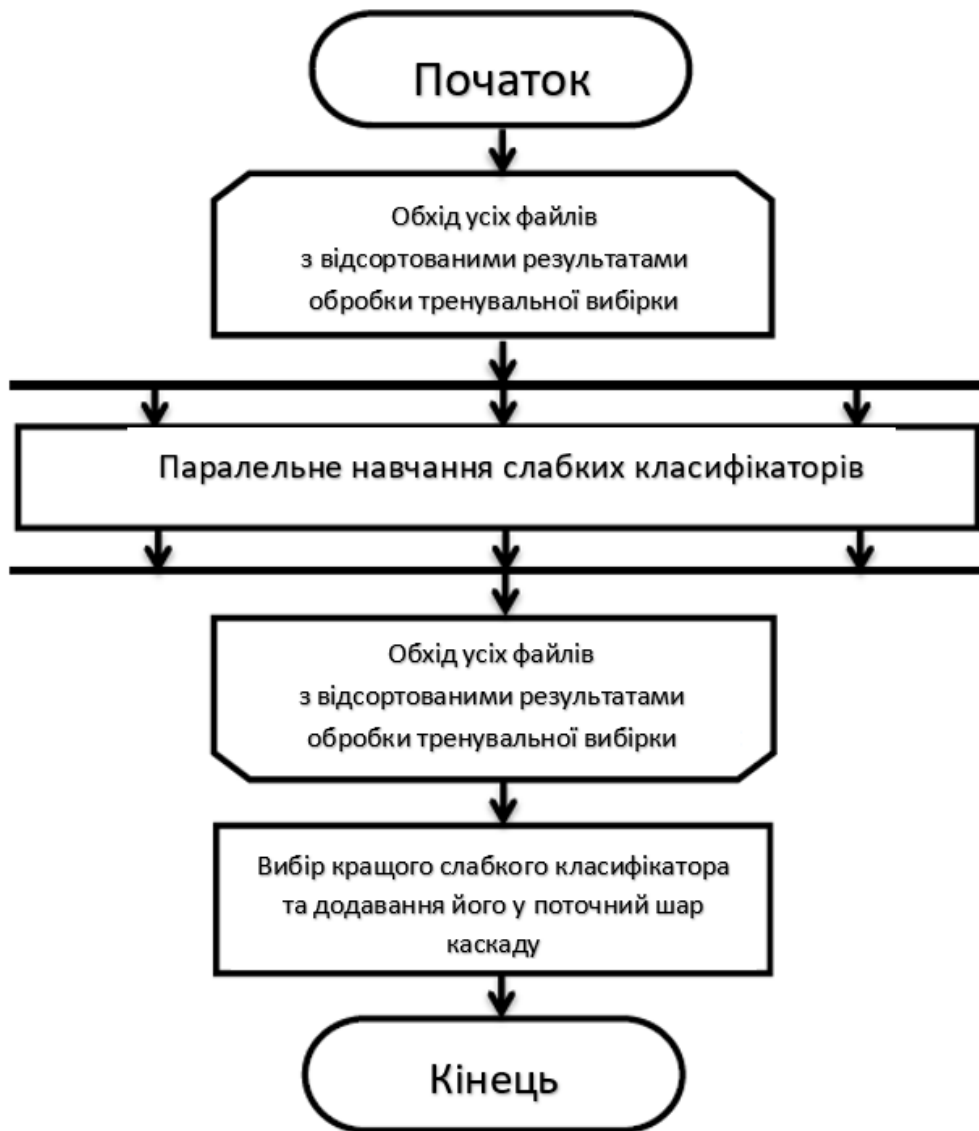


Рисунок 3.3 – Схема паралельна версії алгоритму навчання слабких класифікаторів перед додаванням кращого з них у поточний шар каскада

При цьому слід відзначити, що використований механізм збереження відсортованих первинних результатів обробки навчальною вибіркою малим числом ознак з наступним їх зчитуванням щоразу під час навчання кожного наступного слабого класифікатора дає вигоду тільки у випадку, якщо час, що витрачається на читання з диска, не перевищує час розрахунку та сортування первинних результатів обробки навчальною вибіркою даним числом ознак. В іншому випадку цей розрахунок вигідніше виробляти щоразу «на льоту» під час самого навчання.

Застосування раніше навченого класифікатора для розпізнавання осіб. На етапі використання раніше навченого класифікатора виробляється переміщення скануючого вікна по всьому зображенню з поступовим збільшенням масштабу аналізу та для кожного з положень скануючого вікна розраховується повертається отриманим під час навчання каскадом значення як результату класифікації. Також виробляється об'єднання кількох близьких позитивних результатів детектування на один.

При створенні паралельною версії алгоритму аналізу зображення був реалізований паралелізм скануючих вікон у ході виконання обчислювального процесу застосування каскаду.

Таблиця 3.3 – Показники вимірів часу роботи послідовною та паралельною версій алгоритму аналізу зображень готовим класифікатором

Розмір зображення у пікселях	Послідовна версія	Паралельна версія
160x120	0 год 14 хв	1.16 з
320x240	1 год 54 хв	4.15 з
640x480	9 год 32 хв	21.87 з

Через обмеження за обсягом пам'яті масштабування скануючих вікон проводиться у послідовному режимі. Виміри продуктивності показали в середньому більше чим тисячоразове перевага паралельною версії алгоритму (таблиця 3.3).

3.3 Моделі та алгоритми внесення деформуючих спотворень з метою аугментації навчальною вибірки під час розпізнавання осіб за методом Віюлі – Джонса

Під час проведення дослідження були розглянуто три моделі внесення ДІ з метою штучного розмноження навчальних даних. Результати дослідження представлено далі у роботі.

Алгоритм внесення штучною деформації на основі гармонійних функцій.

Перша модель внесення ДС заснована на застосуванні як функції деформації  $r(x) = r(x, y) = \{r_i(x, y), i = 1, 2\}$  гармонійних функцій (ГФ) виду

$$r_i(x, y) = A_i \sin(\omega_{xi}x + \omega_{yi}y + \psi_i), \quad (3.1)$$

де  $A_i$  – амплітуда зміщення при внесенні ГФ;  $\omega_{yi}$  – частоти функції деформації щодо кожної з осей координат зображення;  $\psi$  – значення амплітуди  $A_i$  обмежує максимальну величину усунення координат пікселів при внесенні ДС і визначається тим, наскільки можна незалежно один від одного змістити риси особи щодо їх вихідних позицій те щоб не зруйнувати його загальну структуру.

Частоти  $\omega_{xi}$  і  $\omega_{yi}$  визначають розміри деталей зображення, які будуть деформуватися. При малих значеннях здійснюється зміщення та плавна деформація зображення об'єкта в цілому зі збереженням його структурних властивостей, а при їх збільшенні – деформуються дедалі більше дрібні його частини. Виходячи з цих міркувань використовуються в експериментах значення  $\omega_{xi}$  і  $\omega_{yi}$  підбиралися так, щоб період функції деформації приблизно відповідав розмірам основних чорт особи на зображенні  $r_i(x, y)$

Фаза функції деформації визначає її зсув та для забезпечення можливості отримання кількох різних деформованих зображень на основі одного екземпляра вихідний навчальною вибірки, її значення генерується випадковим чином за рівномірним законом на всій області визначення функції деформації.

У результаті реалізований відповідно до даної моделі алгоритм полягав у тому, що для кожного з зображень вихідний навчальною вибірки задане число разів з урахуванням описаних вище міркувань генерувалися параметри функції деформації і на них основі вироблялося внесення ДС з метою отримати нові зображення .

На рисунку 3.4 показано приклади результатів внесення ДС такого типу у зображення осіб, які можуть використовуватися як навчальною вибіркою у методі Віоли-Джонса.



Рисунок 3.4 – Приклади внесення деформуючих спотворень у зображення осіб із застосуванням гармонійною функції деформації (перший стовпець зліва є вихідним)

Алгоритм внесення штучною деформації на основі аналізу зміщень контрольних точок. Друга модель внесення ДС заснована на часто застосовуваному під час реалізації морфінгу [14] підході, що полягає в ручній розміщення відповідностей на двох суміщаються зображення у вигляді контрольних точок або контурів. Далі, на основі відомих для зазначених елементів попарних зміщень або значень функції деформації в даних точках за допомогою інтерполяції або деякого іншого методу знаходять решта значення функції деформації. Одним з використовуються при такому підході способів інтерполяції є застосування радіально-базових функцій [8, 9].

Реалізований відповідно до даної моделлю алгоритм полягає в наступному. Для кожного з зображень підлягає розмноження навчальною

вибірки спочатку виробляється розміщення контрольних точок  $(x_k, y_k)$ ,  $k = 1, m$ , де число контрольних точок  $m=5$ : по одній на око, кінчик носа та куточки рота (рисунок 3.5).



Рисунок 3.5 – Приклади розміщення контрольних точок на зображеннях осіб

Далі розраховуються середні значення та їх матриця коваріації, які використовуються при генерації випадкових положень контрольних точок для нових штучно одержуваних даних. Після цього внесення ДС в одне з зображень вихідний навчальною вибіркою виробляється в такий спосіб.

Генерується випадковий набір координат контрольних точок  $(x_k, y_k)$  відповідно до отриманих раніше статистичними параметрами їх розподілу (при генерації замість середніх значень також можуть подаватися координати самих контрольних точок розмножується зображення) з цих точок на основі (2.1) та (2.2) розраховуються значення функції деформації  $r_i(x_k, y_k)$   $i = 1, 2$  (3.2) у даній точці так, щоб

$$\hat{f}(x_k + r_1(x_k, y_k), y_k + r_2(x_k, y_k)) = \hat{f}(u(x_k, y_k)) = g(x_k, y_k), \quad (3.2)$$

де  $f(x, y)$  – деформується зображення;  $g(x, y)$  - результат деформації, що відповідає набору випадково згенерованих положень контрольних точок.

У цій моделі кінцеві функції деформації  $r_i(x_k, y_k)$   $i = 1, 2$  представляються у вигляді сум РБФ- функцій.

На рисунку 3.6 показано приклади результатів внесення ДС на основі описаного підходу.



Рисунок 3.6 – Приклади внесення деформуючих спотворень у зображення осіб на основі розміщення контрольних точок та інтерполяції функції деформації за допомогою РБФ

3.4 Алгоритм внесення штучною деформації на основі аналізу оптичного потоку з використанням фільтра ентропії

Третя модель ґрунтується на ідеях визначення оптичного потоку (ВП) у процесі зміни положень та орієнтації об'єктів сцени у різні моменти часу або за різних умов зйомки [3, 7]. Обчислення оптичного потоку полягає в оцінці зміщень вмісту сцени за координатами  $x$  та  $y$ , які і пропонується безпосередньо використовувати як значення функцій деформації  $t_i(x, y)$ ,  $i = 1, 2$ .

Існує безліч механізмів визначення оптичного потоку [8]. Кожен з них обмежений своєю областю застосовності через закладені в нього припущення, на основі яких і проводиться його розрахунок. Для даної моделі через свої особливостей був обраний метод Фарнебака [4], короткий опис якого представлено у розділі 1.3.

Через специфіку предметний області – обробка зображень осіб, різницю між якими не можна ув'язати до зміщення положень інтенсивностей

їх пікселів, перед процедурою пошуку оптичного потоку потрібно провести додаткове перетворення. Їм є операція накладення фільтра ентропії [7], який для кожної координати зображення та її околиці з пікселів обчислюється в такий спосіб. Розраховуються ймовірності  $p_i$ ,  $i = 1, n$  появи кожного значення інтенсивності пікселя в даній області, і на них основі – ентропія її центрального пікселя:

$$e(x, y) = -\sum_{i=1}^n p_i \log p_i, \quad (3.3)$$

де – результат накладення фільтра, який і використовується для розрахунку оптичного потоку  $e(x, y)$ .

Таким чином, зображення осіб замінюються образами, що відбивають їх основну структуру. У них дрібні риси особи відповідають високим значенням ентропії, а фон і шкіра чола і щік – більше низьким (рисунок 3.7).



Рисунок 3.7 – Приклади накладення фільтра ентропії на зображення осіб

У результаті розмноження зображень навчальною вибіркою здійснювалося в такий спосіб. на кожне їх вироблялося накладення фільтра ентропії та обчислення. Після чого розраховувалися вибіркові параметри ентропії: середня та матриця підступи  $e(x, y)$ .

Далі, при розмноженні кожного зображення вихідний навчальною вибіркою вироблялася генерація випадкових значень ентропії та розрахунок функції  $e_g(x, y)$  деформації  $r_i(x, y)$ ,  $i = 1, 2$  шляхом обчислення оптичного

потоків між згенерованим значенням ентропії і ентропією розмножується зображення.

Отриманий результат  $r_i(x, y)$ ,  $i = 1, 2$  і використовувався при внесення ДС в якості функції деформації на рисунку 3.8 показано приклади роботи алгоритму, заснованого на даній моделі.



Рисунок 3.8 – Приклади внесення деформуючих спотворень у зображення осіб на основі накладення фільтра ентропії та розрахунку оптичного потоку

Експериментальні дослідження впливу штучного розмноження елементів навчальною вибіркою на якість розпізнавання осіб на основі методу Віоли-Джонса.

Для оцінки впливу аугментації або штучного розмноження елементів навчальною вибіркою на якість детектування були проведено експерименти з навчання детекторів осіб за методом Віоли -Джонса з наступними вимірами точності їх роботи [7]. Навчання було здійснено з використанням п'яти варіантів формування навчальних вибірок :

- оригінальна, що містить 4916 зображень осіб , зібраних та підготовлених вручну авторами оригінального методу Віоли-Джонса [3, 8];
- усічена, що містить 492 зображень осіб з числа зображень оригінальною вибіркою;
- три штучно розмножені вибірки, що містять по 4916 зображень осіб з

яких 492 зображення належали усіченому набору зображень, а решта (по 9 на кожне з вихідних) згенеровані штучно шляхом внесення ДС на основі кожною з описаних вище моделей.

Оцінка якості детектування проводилася в такий спосіб. Для кожного детектора модифікувався поріг його реагування, і відповідно до отриманих відгуків було побудовано залежність відносної частоти правильно знайдених осіб від числа помилкових спрацьовувань (рисунок 3.9). Як зображень для тестування був використаний спеціальний набір зображень з вибірки CMU+MIT [9], що часто застосовується для порівняння роботи детекторів фронтальних осіб та містить 42 зображення зі 170 особами.

У таблиці 3.4 представлені типові результати, отримані кожного з детекторів при такому числі помилкових спрацьовувань. Аналіз даних таблиці 3.4 та отриманих залежностей (рисунок 3.9) показує, що при такій кількості помилкових спрацьовувань детектор, навчений з використанням усіченої вибірки, стабільно видає результати гірше оригінального приблизно на 20%.

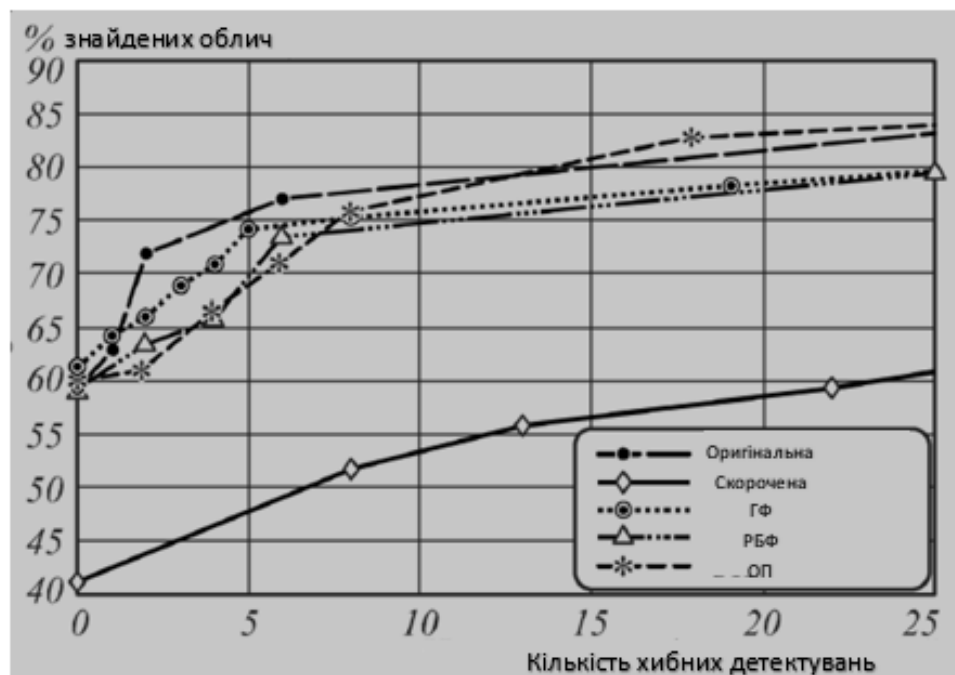


Рисунок 3.9 – Залежність відносної частоти правильно знайдених осіб від числа хибних спрацьовувань, отримані під час навчання за оригінальною та усіченою вибіркою

У той же час, детектори, засновані на трьох розглянутих вище моделях аугментації навчальних даних шляхом внесення в них ДС, дають порівняний з оригіналом результат з відхиленням від нього не більше 5%. При цьому найкращі результати (практично такі ж, як отримані на основі оригінальної версії) показав алгоритм, заснований на використанні розрахунку оптичного потоку із застосуванням ентропії.

Таблиця 3.4 – Відносні частоти правильно знайдених осіб при подібному числі хибних спрацьовувань, отримані під час навчання за оригінальною та усіченою вибіркою, а також з використанням трьох розглянутих моделей аугментації навчальних даних

	Оригінальна	Усічена	ГФ	РБФ	ВП
Відсоток знайдених осіб	83,53	59,41	78,23	79,41	82,94
Число помилкових спрацьовувань	26	22	19	25	18

Таким чином, показано значне збільшення точності роботи детекторів осіб при порівнянні їх із класифікатором, заснованим на усіченій навчальною вибірці (яка і була використана для розмноження). Отримані результати дозволяють говорити про застосовність даного підходу до завдань побудови детекторів виявлення об'єктів, що мають яскраво виражені структурні характеристики, для скорочення витрат на пошук та підготовку навчальних даних, а також в умовах неможливості збору достатнього для початку процесу навчання кількості елементів навчальною вибіркою.

В даному розділі описано реалізацію методу Віоли-Джонса з використанням технології паралельних обчислень на графічних процесори CUDA. Здійснено експерименти порівняно продуктивності паралельної та послідовної версій різних етапів виконання алгоритму. Результати проведених експериментів виявили помітні переваги у швидкості паралельною версії реалізації алгоритму Віоли -Джонса над послідовною та дозволили рекомендувати застосування паралельних обчислень на графічних процесорах як перспективного способу збільшення продуктивності методу Віоли-Джонса з урахуванням відносною простоти та низької вартості їх використання.

На користь вирішення завдань машинного навчання, пов'язаних з обробкою зображень на прикладі завдання фронтального розпізнавання осіб методом Віоли-Джонса розглянуто три моделі внесення деформуючих спотворень для аугментації навчальною вибірки : на основі застосування при деформації гармонійних функцій , на основі зміщення контрольних точок та на основі пошуку оптичного потоку за методом Фарнебака при накладенні фільтра ентропії .Розмножених відповідно до описаних моделей навчальних вибірок проведено експерименти з навчання алгоритмів розпізнавання зображень осіб методом Віоли-Джонса та дані результати порівняння точності роботи отриманих детекторів.

Аналіз отриманих результатів показує , що при подібному числі хибних спрацьовує детектор, навчений з використанням усіченою вибірки , стабільно видає результати гірше оригінального приблизно 20%. Водночас алгоритми , засновані на використанні трьох розглянутих моделей аугментації навчальних даних шляхом внесення до них деформуючих спотворень дають порівнянний з оригіналом результат з відхиленням від нього трохи більше 5%. При цьому найкращі результати (практично такі ж, як і отримані на основі оригінальною версії) показав алгоритм, заснований на використанні розрахунку оптичного потоку із застосуванням фільтра ентропії.

Представлені результати дозволяють говорити про застосовність

даного підходу до завдань побудови детекторів виявлення об'єктів, що мають яскраво виражені структурні характеристики, оскільки це забезпечує скорочення витрат на пошук та підготовку навчальних даних, а також в умовах неможливості збору достатнього для початку процесу навчання кількості елементів навчальною вибірки.

## ВИСНОВКИ

У ході виконання кваліфікаційної роботи було проведено аналіз відомих підходів до використання моделей деформуючих спотворень у завданнях обробки зображень.

Запропоновано, обґрунтовано та досліджено метод та алгоритми впливу деформуючих спотворень у рамках завдань обробки зображень.

Синтезовано та досліджено метод аугментації (штучного розмноження) даних шляхом внесення до них деформуючих спотворень у завданнях машинного навчання на прикладі алгоритмів пошуку та виявлення осіб шляхом Віоли -Джонса.

Описано алгоритмічну реалізацію методу Віоли-Джонса з використанням технології паралельних обчислень на графічних процесорах CUDA. Отримано кількісні результати порівняно швидкості роботи паралельної та послідовної версій окремих етапів роботи алгоритму, які виявили помітні переваги паралельних версій над послідовними та показали перспективність застосування останніх як способу збільшення продуктивності методу Віоли-Джонса.

В інтересах вирішення завдань машинного навчання, пов'язаних з обробкою зображень на прикладі завдання пошуку та виявлення фронтальних осіб методом Віоли-Джонса розглянуто три моделі внесення деформуючих спотворень для аугментації навчальною вибірки: на основі застосування при деформації гармонійних функцій; на основі зміщення контрольних точок; на основі пошуку оптичного потоку за методом Фарнебака при накладенні фільтра ентропії. Для вихідних та штучно розмножених відповідно до описаних моделей учнів вибірок проведено експерименти з навчання алгоритмів розпізнавання зображень осіб методом Віоли-Джонса та дані результати порівняння точності роботи отриманих детекторів. Отримані результати показали застосовність даного підходу до

завдань побудови детекторів виявлення об'єктів, що мають яскраво виражені структурні характеристики, для скорочення витрат на пошук та підготовку навчальних даних.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Д.О. Міщенко, О.О. Можасєв Облік деформуючих спотворень у моделях розпізнавання образів. Тези доповідей 12-ої міжнародної науково-технічної конференції «Проблеми інформатизації», 21-22 листопада 2024р. Баку: ІСУ – Харків: ХНУРЕ, 2024 – Т.2., С.140.
2. Viola, P. Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M. Jones // Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition 2001. – 2001. – Vol. 1. – P. 511-518.
3. Viola, P. Robust Real-Time Face Detection / P. Viola, M. Jones // International Journal of Computer Vision. – Netherlands: Kluwer Academic Publishers. – 2004. – № 57(2). – P. 137-154.
4. Papageorgiou, C.P. A general framework for object detection / C.P. Papageorgiou, M. Oren, T. Poggio // International Conference on Computer Vision. – 1998. – P. 555-562.
5. Freund, Y. A Short Introduction to Boosting / Y. Freund, R. Schapire // Japan: Journal of Japanese Society for Artificial Intelligence. – 1999. – № 14(5). – P. 771-780.
6. Jensen, O. H. Implementing the Viola-Jones Face Detection Algorithm / O. H. Jensen. – DK-2800 Kongens Lyngby, Denmark: Technical University of Denmark Informatics and Mathematical Modelling, 2008. – 36 p.
7. Steyvers M. Morphing Techniques for Manipulating Face Images / M. Steyvers // Behavior Research Methods, Instruments, & Computers. – 1999. – Vol. 31(2). – P. 359-369. – DOI: 10.3758/BF03207733.
8. Arad N. Image warping by radial basis functions: applications to facial expressions / N. Arad, N. Dyn, D. Reissfeld, Y. Yeshurun // CVGIP: Graph Models Image Processing. – 1994. – Vol. 56. – P. 161-172. – DOI: 10.1006/cgip.1994.1015.
9. Ciresan, D.C. Deep Big Simple Neural Nets Excel on Handwritten Digit

Recognition / D.C. Ciresan, U. Meier, L.M. Gambardella, J. Schmidhuber // Neural Computation Conference. – 2010. – Vol. 22(12). – P. 3207-3220. – DOI: 10.1162/NECO\_a\_00052.

10. Wong, S.C. Understanding Data Augmentation for Vlassification: When to Warp? / S.C. Wong, A. Gatt, V. Stamatescu, M.D. McDonnell // DICTA 2016 Conference, Computer Vision and Pattern Recognition (cs.CV). – 2016. – P. 1-6. – DOI: 10.1109/DICTA.2016.7797091.

11. Wang, J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning / J. Wang, L. Perez // arXiv: Computer Vision and Pattern Recognition (cs.CV). – 2017. – P. 1-8.

12. Wolberg G. Image Morphing Survey / G. Wolberg // The Visual Computer. – 1998. – Vol. 14. – P. 360-372. – DOI: 10.1007/s003710050148.

13. Zope, B. A Survey of Morphing Techniques / B. Zope, S.B. Zope // International Journal of Advanced Engineering, Management and Science (IJAEMS) – 2017. – Vol. 3(2). – P. 81-87. – DOI: 10.24001/ijaems.3.2.15.

14. Ruprecht, D. Deformed Cross-dissolves for Interpolation in Scientific Visualization / D. Ruprecht, H. Müller // J. of Visualization and Computer Animation – 1994. – Vol 5(3). – P. 167–181. – DOI: 10.1002/vis.4340050304.

15. Gao, P. A Work Minimization Approach to Image Morphing / P. Gao, T.W. Sederberg // The Visual Computer. – 1998. – Vol. 14(8-9). – P. 390-400. – DOI: 10.1007/s003710050150.

16. Liao, J. Automating Image Morphing Using Structural Similarity on a Halfway Domain / J. Liao. R.S. Lima, D. Nehab, H. Hoppe, P.V. Sander, J.Yu // ACM Transactions on Graphics (TOG). – 2014. – Vol. 33(5). – P. 1-12. – DOI: 10.1145/2629494.

17. Le Guennec, A. Data Augmentation for Time Series Classification Using Convolutional Neural Networks / A. Le Guennec, S. Malinowski, R. Tavenard // ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, Sep. 2016. – 2016. – P. 1-8.

18. Rabiner, L. Fundamentals of Speech Recognition / L. Rabiner, B.H.

Juang. – Englewood Cliffs, NJ: Prentice Hall PTR, 1993. – 507 p.

19. Corradini, A. Dynamic time warping for offline recognition of a small gesture vocabulary / A. Corradini. // Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems. – 2001. – P. 82-89. – DOI: 10.1.1.200.2035.