

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет

Інфокомунікацій

(повна назва)

Кафедра

Інформаційно-мережної інженерії

(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти другий(магістерський)

Сумісне розпізнавання віку людини за його голосом та зображенням
(тема)

Виконав:

студент 2 курсу, групи ІМІМ-19-2

Єрмолаєв А.А.

(прізвище, ініціали)

Спеціальність 172 «Телекомунікації та
радіотехніка»

(код і повна назва спеціальності)

Тип програми освітньо-наукова

(освітньо-професійна або освітньо-наукова)

Освітня програма «Інформаційно-
мережна інженерія»

(повна назва освітньої програми)

Керівник доц. Омельченко С.В.

(прізвище та ініціали)

Допускається до захисту

Зав. кафедри

(підпис)

Безрук В.М.

(прізвище, ініціали)

2021 р.

Не містить відомостей, заборонених до відкритого публікування

Студент _____ /Єрмолаєв А.А./

Керівник _____ /Омельченко С.В./

Харківський національний університет радіоелектроніки

Факультет інфокомунікацій
Кафедра інформаційно-мережної інженерії
Рівень вищої освіти другий (магістерський)
Спеціальність 172 «Телекомунікації та радіотехніка»
(код і повна назва)
Тип програми ОВІТНЬО-НАУКОВА
(освітньо-професійна або освітньо-наукова)
Освітня програма «Інформаційно-мережна інженерія»
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри ІМІ _____
(підпис)

«____» _____ 2021 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студентові Єрмолаєву Артему Андрійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Сумісне розпізнавання віку людини за його голосом та зображенням обличчя

затверджена наказом по університету від «15» березня 2021 року № 350 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 21 травня 2021р.

3. Вихідні дані до роботи Розглянуті методи розпізнавання віку людини за її голосом та зображенням обличчя, були розглянуті техніки вилучення ознак віку, та розглянуті алгоритми розпізнавання віку людини за голосом та зображенням обличчя

4. Перелік питань, що потрібно опрацювати в роботі

Вступ та висновки

1 Огляд сучасних підходів до вирішення задачі розпізнавання віку людини по голосу та зображенню обличчя

2 Техніки вилучення ознак віку

3 Алгоритми оцінки віку

4 Опис практичної частини

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

Слайди у форматі Power Point (мета роботи, розглянуті методи розпізнавання, метод опорних векторів, база еталонів дикторів, спектрограма чоловічого голосу, спектрограма жіночого голосу, Результати досліджень алгоритмів з застосуванням ознак MFCC, Алгоритм розпізнавання віку за зображеннями обличчя на основі вибору ієрархічних ознак, дослідження методу KNN – K-найближчих сусідів, результати досліджень методу KNN – K-найближчих сусідів, Висновки).

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів атестаційної роботи	Термін виконання етапів роботи	Примітка
1	Ознайомлення із завданням. Уточнення ТЗ.	15.03.21	
2	Підбір літератури за темою роботи.	17.03.21-22.03.21	
3	Виконання розділу 1	22.03.21-08.04.21	
4	Виконання розділу 2	09.04.21-15.04.21	
5	Виконання розділу 3	16.04.21-22.04.21	
6	Виконання розділу 4	23.04.21-28.04.21	
7	Виконання розділу 5	28.04.21-03.05.21	
8	Оформлення пояснювальної записки	03.05.21-07.05.21	
9	Оформлення презентаційного матеріалу, підготовка до захисту у		
	ЕК	07.05.21-13.05.21	

Дата видачі завдання 15 березня 2021рр.

Студент _____
(підпис)

Керівник роботи _____
(підпис)

доц. Омельченко С.В.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка 64 с., 11 рис., 9 табл., 60 джерел, 2 додатки

Об'єкт роботи – Сумісне розпізнавання віку людини за його голосом та зображенням обличчя.

Мета роботи – дослідження та впровадження методів сумісної ідентифікації віку людини за голосом та зображенню обличчя.

Розглянуті основні методи розпізнавання віку людини за голосом та зображенням обличчя. Були також розглянуті основні техніки вилучення ознак віку. В практичній частині були проведені експериментальні дослідження розпізнавання віку людини по голосу та зображенню обличчя з використанням формантих частот та мел-частотних кепстральних коефіцієнтів.

РОЗПІЗНАВАННЯ, ВІК, ЗА ГОЛОСОМ ,ЗА ЗОБРАЖЕННЯМ ОБЛИЧЧЯ,
ОЗНАКИ ВІКУ, МЕТОДИ, АЛГОРИТМИ.

THE ABSTRACT

Explanatory note: 64p., 11 fig., 9 tabl., 60 sources, 2 app.

Object of research – Joint recognition of a person's age by his voice and facial image.

The purpose of this work is research and implementation of methods of joint recognition of human age by voice and facial image

The main methods of recognizing person's age by voice and face image are considered. The basic techniques for extracting signs of age were also discussed. In the practical part, experimental studies of human age recognition by voice and facial image were performed using frequency formants and mel-frequency cepstral coefficients.

RECOGNITION, AGE, BY VOICE, BY FACIAL IMAGES, SIGNS OF AGE, METHODS, ALGORITHMS

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ.....	7
ВСТУП.....	8
1 ОГЛЯД СУЧАСНИХ ПІДХОДІВ ДО ВИРІШЕННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ ВІКУ ЛЮДИНИ ПО ГОЛОСУ ТА ЗОБРАЖЕННЮ ОБЛИЧЧЯ.....	11
1.1 Оцінка віку за зображеннями обличчя на основі вибору ієрархічних ознак.....	11
1.2 Виділення ознак низького рівня.....	12
1.3 Вибір рис обличчя.....	16
1.4 Метод опорних векторів (SVM).....	19
2 ТЕХНІКИ ВИЛУЧЕННЯ ОЗНАК ВІКУ	25
2.1 Фільтри Габору	25
2.2 Лінійний дискримінантний аналіз	26
2.3 Локальний двійковий шаблон	27
3 АЛГОРИТМИ ОЦІНКИ ВІКУ	31
3.1 Класифікація	31
3.2 Регресія.....	32
3.3 Гібридний підхід.....	33
4 ОПИС ПРАКТИЧНОЇ ЧАСТИНИ.....	36
4.1 Експериментальні дослідження алгоритму розпізнавання віку за зображеннями обличчя на основі вибору ієрархічних ознак	36
4.2 Експериментальні результати з використанням алгоритму KNN (К-Найближчих сусідів).....	39
ВИСНОВКИ	44
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	45
ДОДАТОК А	51
ДОДАТОК Б.....	55

ПЕРЕЛІК СКОРОЧЕНЬ

LBP – Local Binary Pattern

LDA – Linear discriminant analysis

MFCC – мел-частотні кепстральні коефіцієнти

CCR – коефіцієнт правильної класифікації

РБФ – радіальна базисна функція

ANN – штучна нейронна мережа

ВСТУП

Людське обличчя передає надзвичайну кількість сприйнятливої інформації та таких рис, як емоційний стан, етнічна приналежність, вік та стать. Ця інформація відіграє ключову роль у очній взаємодії між людьми [1]. Більшість людей здатні з легкістю визначити такі людські риси, як емоційний стан, де вони можуть зробити висновок, коли хтось сумний, щасливий чи злий [8] суто з обличчя.

Аналогічно, стать чоловіка чи жінки можна розпізнати спонтанно з рис обличчя. Навпаки, пізнання віку людини за допомогою перегляду або аналізу зображень часто є складним завданням навіть для самої людини. Методи оцінки віку за допомогою комп'ютерних методів зору мають на меті автоматично наблизити віковий діапазон або точний вік людини з її обличчя [2].

Автоматична оцінка віку за даними про обличчя нещодавно привернула багато уваги дослідницького співтовариства, яка виникла як ключова технологія із численними додатками, починаючи від взаємодії людини та машини, контролю доступу, біометрії та індексації та пошуку даних на основі зображень.

Типові сценарії для щойно згаданих категорій включають контроль за безпекою та моніторинг з обмеженим віком. Додаток для оцінки віку може генерувати попереджувальний звук або сигнал тривоги, коли особа, яка не досягла повноліття, заходить у бари чи інші заборонені зони, такі як казино, або навіть купує тютюнові вироби у саморобних машинах, де оцінка віку на основі обличчя може бути використано як основний пункт пропуску.

Цікаво, що автоматизована оцінка віку може бути застосована в зоні взаємодії людини з комп'ютером для визначення віку користувача з метою автоматичного повторного налаштування або коригування графічного інтерфейсу відповідно до потреб та вимог користувача на основі їх вікової групи [3]. Наприклад, значок або орієнтований на зображення інтерфейс

можуть бути ввімкненим для молодших користувачів, тим часом текст із великим шрифтом може відображатися для літніх користувачів.

Системи оцінки віку в цілому розроблені для того, щоб складатися з двох фаз: вилучення ознак та класифікація. Перший крок є центральним для успіху етапу класифікації, оскільки витягнуті ознаки значною мірою впливають на ефективність процесу вікового прогнозування [2]. Враховуючи зображення обличчя людини, віковий діапазон оцінюється на основі набору функцій низького рівня, які автоматично витягуються або коментуються вручну.

Однак автоматизована без маркерна оцінка віку за статичними зображеннями виявляється громіздким процесом. Труднощі випливають із ряду складних факторів, що стосуються як особистості, так і середовища придбання. Прогрес старіння є неконтрольованим і незворотним, так що різні люди зазнають різної швидкості старіння, яка керується не тільки людським геном, а й багатьма іншими зовнішніми факторами, що впливають на процес старіння, включаючи умови здоров'я, спосіб життя та робоче середовище [4], [5], [6]. Проблеми, пов'язані із середовищем придбання, можуть включати безлад у фоновому режимі, освітлення, рух камери та точки зору, а також оклюзію.

Через ключову важливість прогнозування віку в різних додатках, починаючи від взаємодії людини з комп'ютером і закінчуючи інтелектуальними програмами безпеки, ми описуємо в цьому дослідженні підхід на основі бачення для оцінки вікового діапазону від зображення обличчя. Локальний двійковий шаблон застосовується для вилучення гібридного набору ознак, включаючи місцеві та глобальні характеристики з обличчя. Гістограма ознак побудована на основі об'єднання векторів гістограми, що виробляються місцевим шляхом, взятих з клітин сітки.

Вибір ієрархічного ознаки описаний для процесу класифікації, де вікові діапазони групуються по дереву. Вибір ознак базується на близькості та близькості даних, що належать до подібних класів, і застосовується для отримання найбільш виразних ознак на кожному рівні визначеного вікового діапазону. Експериментальні результати, проведені на загальнодоступному

наборі даних, підтвердили можливості запропонованого методу для кращої оцінки вікового діапазону для різних зображень обличчя.

Метою даної роботи є дослідження та впровадження методів сумісної ідентифікації віку людини за голосом та зображенню обличчя.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

1. Дослідити існуючі методи вирішення проблеми ідентифікації віку людини по голосу та зображенню обличчя, оцінити якість, а також існуючі проблеми та обмеження в їх роботі.
2. Впровадити один з кращих алгоритмів ідентифікації віку людини за голосом та зображенню обличчя.
3. Провести експериментальне дослідження розробленого алгоритму.

1 ОГЛЯД СУЧАСНИХ ПІДХОДІВ ДО ВИРІШЕННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ ВІКУ ЛЮДИНИ ПО ГОЛОСУ ТА ЗОБРАЖЕННЮ ОБЛИЧЧЯ

1.1 Оцінка віку за зображеннями обличчя на основі вибору ієрархічних ознак

Порівняно зі значним обсягом досліджень, пов'язаних з дослідженням особи, як біометрія та синтез обличчя, існує порівняно мало публікацій, присвячених автоматизованому оцінюванню віку. Чой та ін. досліджували різні типи ознак, що використовуються для прогнозування віку за зображеннями обличчя, групуючи їх у дві основні категорії; локальні та глобальні особливості.

Для першого типу особливостей вони беруться з глибини або кількості зморшок, волосся на обличчі або шкіри на додаток до геометричних властивостей елементів обличчя. Місцеві характеристики, як відомо, краще класифікують людей за віковими групами, оскільки вони містять певні характеристики, які розрізняють різні вікові групи. На відміну від місцевих особливостей, Чой [2] стверджував, що глобальні аналоги краще оцінюють точнішу вікову інформацію і містять не лише вікові ознаки, а й інші, пов'язані з індивідом риси, як ідентичність, емоції та етнічне походження. Встановлено, що гібридні функції, створені завдяки поєднанню місцевих та глобальних функцій, пропонують чудову продуктивність для різних програм, пов'язаних з обличчям.

Це через неефективність, виявлену для кожного типу ознак, яку можна компенсувати, що призводить до висновку, що гібридні ознаки бажані для точної оцінки віку. У літературі було проведено ряд ранніх досліджень щодо вікового прогресування шляхом моделювання ефектів старіння, що вважається зворотним процесом оцінки віку [7]. Сюди входить робота [9,10], де вони

штучно моделювали зміни старіння, піддаючи зображення обличчя типовим змінам форми та кольору. Є й інші наукові дослідження, які частково пов'язані з оцінкою віку, але спрямовані на вивчення картографії між біометрією обличчя та віком. Челлаппа [12] описав підхід до перевірки біометрії обличчя за віком на основі байєсівського класифікатора. Перше дослідження, знайдене в літературі, яке стосується вікової класифікації зображень обличчя за допомогою методів обробки зображень, опубліковане Квоном. [14]. Вони виділили та використали природні зморшки для вікової класифікації зображень обличчя на три основні групи: немовля, дорослий та дорослий. Перша велика робота була запропонована Ланітісом [1], який стверджує, що мав першу спробу автоматизованого оцінювання віку. Статистичні моделі обличчя будуються шляхом застосування основного компонентного аналізу на наборі зображень, які згодом використовуються як основа для створення компактного параметричного зображення обличчя. Для оцінки показників продуктивності представленого методу розглядаються різні класифікатори, такі як квадратична функція, класифікатор найкоротшої відстані, нейронні мережі та самоорганізуюча карта. Для кожної вікової групи використовується інший класифікатор на основі іншої процедури вибору найбільш підходящого класифікатора. Ланітіс стверджував, що отримані результати є свідченням того, що машини можуть оцінити вік людини майже настільки ж надійно, як і люди. Чой [2] працював над оцінкою віку за даними обличчя в ієрархічному порядку, використовуючи машинний класифікатор опорних векторів. Вікові ознаки побудовані як поєднання локальних та глобальних ознак, отриманих за допомогою локального оператора двійкового шаблону разом із фільтрами Габора для вилучення малюнка зморшок. Юлініус [15] використовував комбінацію бінарних шаблонів. Метод був апробований на зображеннях, записаних в необмежених умовах.

1.2 Виділення ознак низького рівня

Оператор Local Binary Pattern (LBP) вперше був представлений для аналізу текстури Ожалай та іншими співавторами [17] у 1996 р. У своїй редакції “Порівняльне дослідження фактурних мір з класифікацією на основі ознак розподілу”. LBP можна ефективно та швидко обчислити за допомогою одного сканування зображення, що пропонує можливості розпізнавання обличчя навіть для зображень із нижчою роздільною здатністю. Оператор встановлює пікселі даного зображення, встановлюючи порогове значення кожного сусіднього пікселя щодо центрального пікселя в межах матриці 3×3 , і, отже, отримує ряд послідовних значень 1 або 0, як показано на рисунку (1.1). Читаючи в тому ж напрямку стрілку, показану на рисунку (1.1), формулюється двійкове число, яке перетворюється в десяткове число, тобто мітка, де двійкове число: 11010011 перетворюється на 211. Гістограма з отриманими мітками в 256 бінів обчислюється і використовується як дескриптор текстури для додатків на основі обличчя.



Рисунок 1.1 – Основний локальний оператор двійкового шаблону (LBP)

Основним недоліком основного локального двійкового шаблону є його невелика зона сусідства ($3 * 3$), внаслідок чого він може ігнорувати або ігнорувати видатні особливості більших структур. Розширена версія оператора LBP викладена в останніх дослідженнях Ожали та інших співавторів [17] для використання районів різних розмірів. Розширений оператор LBP представлений круговою околицею, записаною як (P, R) , де P відноситься до кількості пікселів у круговій околиці, тоді як R - радіус кругової області, як показано на рисунку (1.2).

Значення LBP для точки пікселя, що має координату (x_c, y_c) , обчислюється, як показано в рівнянні (1.1):

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c)2^i, \quad (1.1)$$

де g_i – значення градації сірого для точки пікселя i . с-центр пікселя. Функція $s(a)$ є функцією порогового значення, що повертає 1 для випадку коли $a \geq 0$ та 0 при значенні $a \leq 0$.

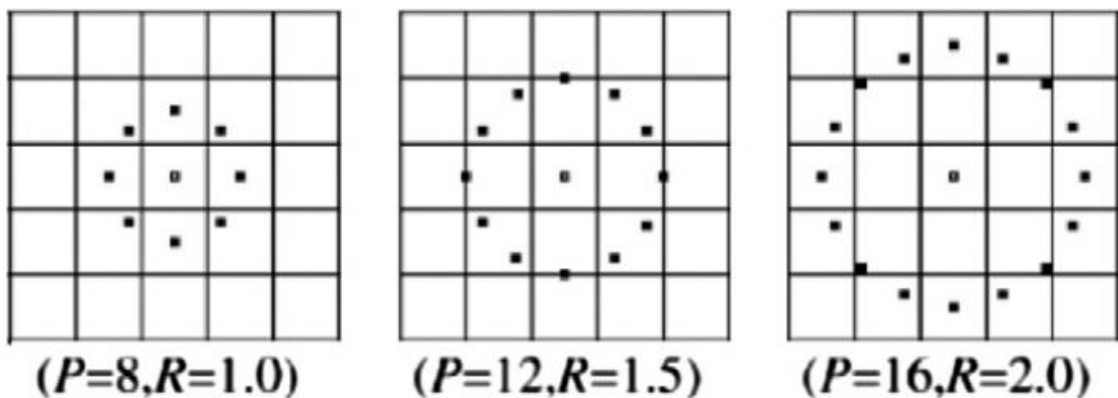


Рисунок 1.2 – Розширений локальний оператор двійкового шаблону.

LBP (P, R) може виробляти 2^P різних вихідних значень відповідно до 2^P різних шаблонів, утворених P , що є кількістю точок у вибраному круговому сусідстві. Дослідження показали, що деякі закономірності передають більш характерну інформацію, ніж інші.

Подальше розширення LBP було введено з урахуванням лише однакових шаблонів, які визначаються як шаблони, що містять щонайбільше два переходи від 0 до 1 або навпаки, тобто: кількість разів, коли цифра змінюється від 0 до 1 або навпаки.

Як приклад, наступні двійкові числа 00000000 та 00111000 та 11100001 є єдиними шаблонами. Оператор LBP, записаний як (LBP^{U2P}, R) , що означає, що

цей оператор LBP використовує (P, R) кругову область сусідства з урахуванням лише однакових шаблонів.

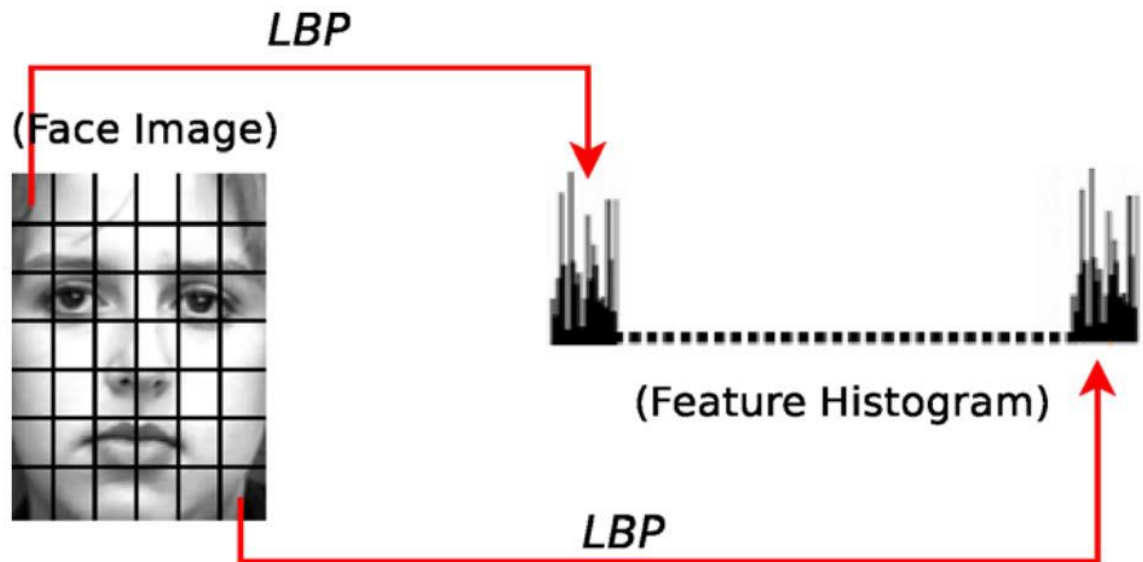


Рисунок 1.3 – Гістограма Local Binary Pattern

Гістограма з n -біном для оператора Локальний двійковий шаблон отримана з позначеного зображення fl , як визначено у рівнянні (1.2):

$$H_i = \sum_{x,y} b(f_i(x,y) == i) \quad i = 0,1,\dots,n-1, \quad (1.2)$$

де b - булева функція, що повертає 1 для справжніх випадків і 0 для помилкових умов.

У роботі [9] описано вдосконалену просторово засновану гістограму, яка ділить зображення на t менших областей R з метою збереження просторових ознак, як показано на рисунку (1.3), де гістограма обчислюється, як встановлено у рівнянні (1.3):

$$H_{ij} = \sum_{x,y} b(f_i(x,y) == i) b((x,y) \in R_j), \quad (1.3)$$

де R_j – це j – та область зображення, розділена на m -областей.

1.3 Вибір рис обличчя

Вектор ознак для оцінки віку обличчя спочатку будується шляхом виявлення обличчя на одному статичному зображенні за допомогою методу Віюли та Джонса. Реалізація передбачена в наборі інструментів комп'ютерного зору Matlab. Для подальшого вдосконалення та налаштування точності виявлення ми застосовуємо симетричний аналіз, щоб неіснуючі риси обличчя, такі як волосся, ігнорувались. На рисунку (1.1) показані кроки, які виконуються від виявлення грані до формулювання вектора ознак за допомогою побудови гистограми на основі локального двійкового оператора, обговореного в попередньому розділі. У цьому дослідженні розглянуто методику вибору ознак для вилучення найбільш характерних рис та фільтрування зайвої та неактуальної інформації про обличчя, яка може погано впливати на точність оцінки. Процедура адаптивного послідовного плаваючого вибору вперед розгортається з метою зменшення початкового простору об'єктів. Метод вибору ознак успішно застосовується в біометричних та спостережних програмах [13,19,16].

Процедура вибору підмножини ознак залежить від критерію оцінки або цільової функції, яка досліджує відмінність виміру ознаки або групи ознак для отримання ідеальної підмножини для завдання класифікації. У цій роботі використано два критерії перевірки.

Перша функція полягає в апроксимації різних кластерів вікових груп. Система, впроваджена в цьому дослідженні, використовує варіацію виміру відстані Бхаттачарія. Метрика відстані Бхаттачарія є мірою оцінки поділу $S_{i,j}$ між класом i та j заданою формулою:

$$S_{i,j} = (m_i - m_j) \left(\frac{\Sigma i + \Sigma j}{2} \right)^{-1} * (m_i - m_j)^T, \quad (1.4)$$

такі, що m_i та Σi є середнім значенням та коваріацією класу i .

Для випадку з N класами оцінка поділу обчислюється з використанням наступної формули:

$$J = \frac{1}{N^2} \sum_{a=1}^N \sum_{b=1}^N S_{a,b}, \quad (1.5)$$

Для оцінки вікової групи для даного зображення на основі кластеризованих груп пропонується критерій оцінки на основі валідації, щоб вивести підмножину ознак, яка зменшила б класифікаційні помилки, а також гарантувала вищу міжкласову подільність між різними групами. На відміну від схеми голосування, що використовується класифікатором KNN , цільова функція використовує коефіцієнти ω , що відображають важливість найближчих сусідів, що належать до того самого класу. Значення оцінки для екземпляра S_c належності до вікового кластеру c наведено у наступному рівнянні (1.6):

$$f(S_c) = \frac{\sum_{i=1}^{N_c-1} z_i \omega_i}{\sum_{i=1}^{N_c-1} \omega_i}, \quad (1.6)$$

де N_c – кількість екземплярів у кластері c , i коефіцієнт ω_i для i -го найближчого екземпляра є обернено пов'язаним з близькістю як у рівнянні:

$$\omega_i = (N_c - i)^2, \quad (1.7)$$

Значення z_i визначається як:

$$z_i = \begin{cases} 1 & \text{якщо найближче } (S_c, i) \in c \text{ в іншому випадку,} \end{cases} \quad (1.8)$$

Така, що найближча (S_c, i) функція повертає i -й найближчий зразок екземпляру S_c . Евклідова відстань обчислюється для виведення найближчих сусідів з того ж класу.

Замість того, щоб виконувати процес вибору ознак один раз для всіх підсумкових класів, тобто певних вікових груп, підсумкові класи ієрархічно перегрупуються у вищі класи, виробляючи подібну класифікацію на основі дерев, де листя є кінцевими класами. Вузли відповідають перегрупованим класам вищого рівня або віковим діапазнам. Тому вибір функції застосовується рекурсивно для генерації на кожному рівні відповідної підмножини ознак.

Перегруповання вікових діапазонів виконується поступово на кожному рівні, щоб скласти два відмінних кластери віків, що містять сусідні класи, що мають вищу відокремленість. На рисунку (1.4) показано ієрархічне перегруповання класів, розгорнутих для оцінки віку, де для початку використовуються два вищі класи вікових груп.

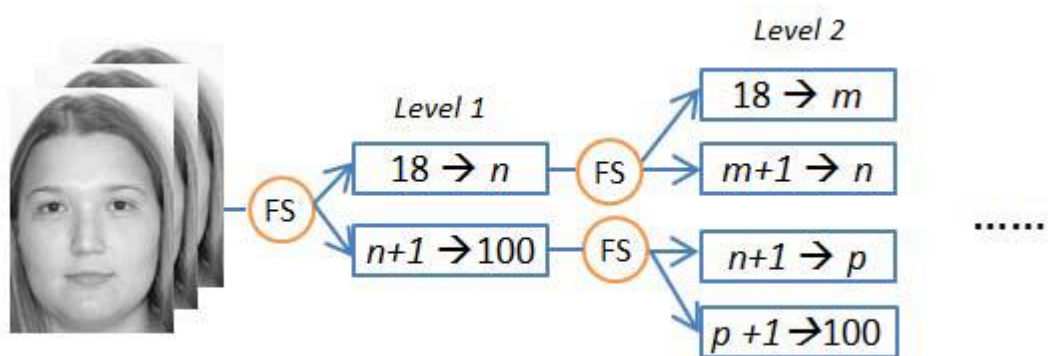


Рисунок 1.4 – Процес вибору ієрархічної підмножини для оцінки вікових груп

1.4 Метод опорних векторів (SVM)

Support vector machine (метод опорних векторів) – це двійковий класифікатор, який будує в просторі ознак розділяючу функцію, що задає гіперплощину (рис. 1.5), виду:

$$f(x) = w \cdot x + b, \quad (1.9)$$

де w – вектор коефіцієнтів

b – константа

Нехай задана послідовність точок простору ознак $X = \{x_1, x_2, \dots, x_n\}$ з мітками $Y = \{y_1, y_2, \dots, y_n\}$, $y_i \in \{-1, 1\}$, $1 \leq i \leq n$, які відповідають двом класам. У разі лінійної роздільності даних умови для знаходження функції $f(x)$ записуються у вигляді:

$$\begin{cases} w \cdot x_i + b \geq 1, y_i = 1 \\ w \cdot x_i + b \leq -1, y_i = -1 \end{cases} \Leftrightarrow y_i(w \cdot x_i + b) - 1 \geq 0, 1 \leq i \leq n, \quad (1.10)$$

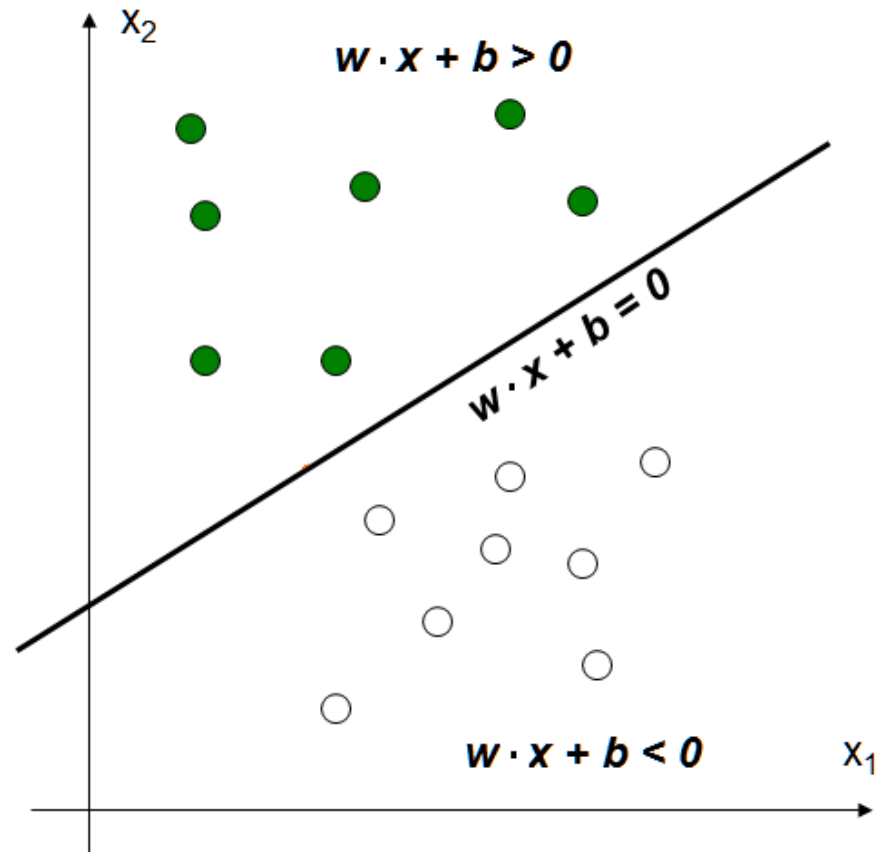


Рисунок 1.5 – Метод опорних векторів

Для надійного розділення класів необхідно, щоб відстань між розділяючими гіперплощинами була якомога більшою. Відстань розраховується як $\frac{2}{\|w\|}$, отже задача пошуку розділяючої гіперплощини може бути зведена до

мінімізації $\|w\|^2$ при заданих умовах у формулі (1.9). Цю задачу можна вирішити за допомогою методу множників Лагранжа.

У випадку лінійно нероз'ємних множин вводиться функція ядра. Основна ідея в тому щоб відобразити вихідний простір в простір більшого розміру в якому множини можуть бути лінійно відокремлені. В той же час в силу того що у всьому алгоритмі ознаки використовуються не окремо, а у вигляді скалярних добутоків, немає необхідності будувати дане перетворення в

явному вигляді. Достатньо задати функцію ядра, яка буде визначати скалярний добуток в новому просторі.

Загальні функції ядра - це ядро радіальної базисної функції (РБФ)

$$K(x, y) = \exp\left(-\frac{\|x - y\|_2^2}{2\sigma^2}\right), \quad (1.11)$$

та ядро полінома

$$K(x, y) = (x \cdot y + c)^d, \quad (1.12)$$

В якості збереженої моделі диктора в методі опорних векторів виступають параметри поділяючої функції $f(x)$, а також параметри функції ядра. Параметри ядра, як правило, визначаються шляхом перебору деякої множини значень та оцінкою застосовуючи метод крос-валідації.

Після того як вирішальна функція $f(x)$ знайдена, приналежність вектору x' до відповідного класу визначається знаком виразу $f(x')$.

Для застосування методу опорних векторів в задачі многокласового розпізнавання застосовується стратегія «один проти всіх». Для цього будується q класифікаторів, кожен з яких вчиться відрізняти один конкретний клас від всіх інших. В процесі розпізнавання об'єкт призначається тому класу, класифікатор якого дає найбільше значення розділяючої функції $f(x)$.

В цілому 246 даних про сигнали були зібрані з голосу 8 дикторів, 4 жінок і 4 чоловіків. Кожному голосу було присвоєно кілька слів з двох складів, а демографічне зведення наводиться в таблиці нижче.

Таблиця 1.1 – Кількість дикторів

Акцент	Стать		
	Жіноча	Чоловіча	Всього
Британський	65	57	122
Американський	59	65	124
Всього	124	122	246

Хоча тривалість звукових доріжок становить всього близько 5 секунд з частотою дискретизації 44100 Гц, кожен вектор звукової доріжки в часовій області містить понад 30000 записів. відповідь дається

$$y_i = \begin{cases} 0 & \text{якщо стать чоловіча} \\ 1 & \text{якщо стать жіноча} \end{cases}$$

Перед проведенням будь-якого формального аналізу були створені дві спектрограми, одна для чоловічого голосу, а інша для жіночого голосу, щоб візуалізувати різницю між статями (рис.1.6), (рис.1.7).

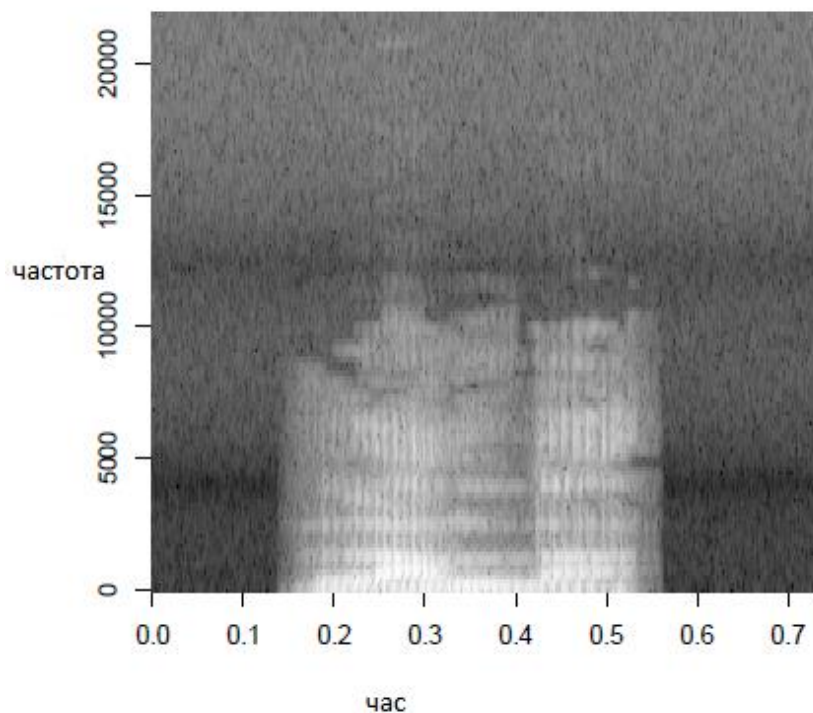


Рисунок 1.6 – Спектрограма чоловічого голосу

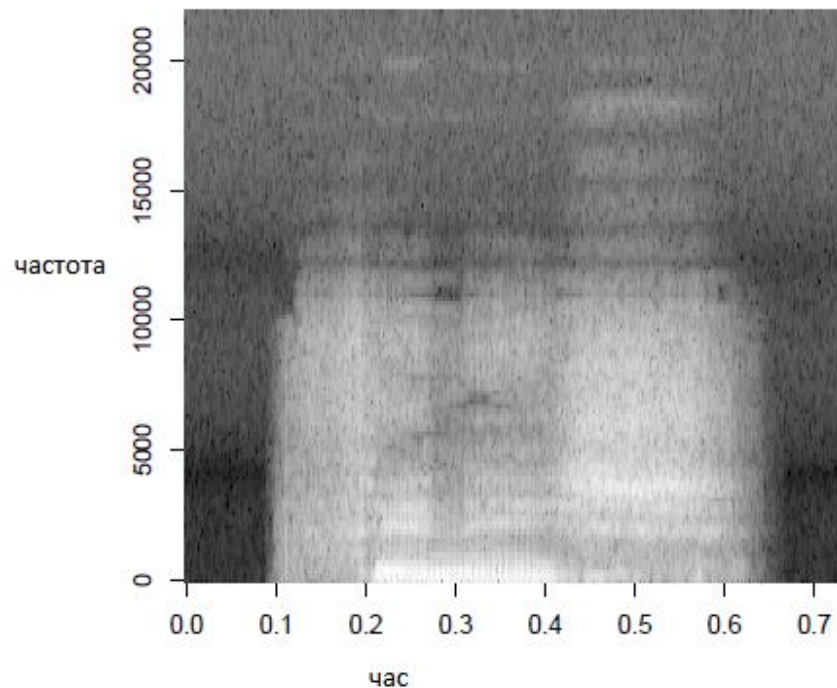


Рисунок 1.7 – Спектрограма жіночого голосу

Хоча важко відрізнити різницю безпосередньо від спектрограми, графіки все ще дають уявлення про те, що жіночий голос має більш високу частоту, ніж чоловічий голос.

Потім MFCC були обчислені для кожної звукової доріжки. Обидва ядра РБФ та ядро полінома були використані, а потім порівнювалися один з одним. Щоб апроксимувати здатність моделі прогнозувати, була проведена перехресна валідація розміру 500 на основі стратифікованої випадкової вибірки. Точність кожного прогнозу – це співвідношення між правильним прогнозом, який є сумою істинно позитивного (TP) і істинно негативного (TN), і загальної кількості звукових доріжок (N). А загальна помилка прогнозування (ERR) – це середня помилка прогнозування перехресного підтвердження розміром 500.

$$ERR = \frac{\sum \left\{ 1 - \frac{TP + TN}{N} \right\}}{500}, \quad (1.13)$$

де ERR – ймовірність помилки

TP – істинно позитивний (позитивні випадки, які правильно визначені)

TN – істинно негативний (негативні випадки, які були неправильно класифіковані як позитивні)

N – кількість звукових доріжок

Результати дослідження наведені в таблиці нижче

Таблиця 1.2 – Результати досліджень

# MFCC	ERR(РБФ)	ERR(Поліном)
12	0.0748	0.0549
19	0.0232	0.0418
26	0.0072	0.0131
34	0.0023	0.0051
40	0.0019	0.0063

Зверніть увагу, що коли число MFCC збільшується з 12 до 26, середня помилка прогнозування падає з 7,5% до 0,7% для ядра РБФ, але не так сильно з 26 до 40. Це не дивно, оскільки інформації про функції, що надається MFCC, більше, так як число MFCC збільшується.

Але оскільки розмірність продовжує зростати, це може вплинути на складність обчислень. Також зверніть увагу, що модель з поліноміальним ядром дуже близька до моделі з ядром РБФ. SVM з ядром РБФ і 26 MFCC є найкращим для прогнозування, хоча інші моделі також генерують результати з високою точністю.

Метод опорних векторів дає високу точність класифікації, має теоретичне обґрунтування, дозволяє застосовувати різні методи до класифікації відповідно до вибору функцій ядра. Серед недоліків необхідно відмітити проблему вибору ядра, а також повільне навчання у випадку задачі многокласового розпізнавання.

2 ТЕХНІКИ ВИЛУЧЕННЯ ОЗНАК ВІКУ

2.1 Фільтри Габора

Спочатку запроваджені Денисом Габором у 1946 р. [21], фільтри Габора широко використовувались для вилучення зморшок, країв та текстур завдяки здатності визначати орієнтацію та величину зморшок [22]. Фільтр Габора вважали найкращим дескриптором текстури у розпізнаванні об'єктів, сегментації, відстеженні руху та реєстрації зображень [23]. Характеристики Габора були використані при оцінці віку [24] і продемонстрували, що вони є ефективним дескриптором текстури порівняно з LBP. Оскільки зморшки виглядають як крайоподібні компоненти з високою частотою, для аналізу особливостей зморшок зазвичай використовується метод аналізу краю Габора. Фільтр Собеля [25, 26], перетворення Хафа [27] та активні контури [28] є одними з найбільш часто використовуваних дескрипторів країв текстури. Хоча краї зображення на обличчі також складаються з таких шумів, як бороди, вуса, волоски та тіні, для зменшення ефекту цього шуму [29] пропонує використовувати переважну орієнтацію зморшок, що слід враховувати при видаленні особливостей зморшок. 2D просторовий домен Габора визначається як:

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j \mathcal{W}x \right], \quad (2.1)$$

де σ_x^2 та σ_y^2 - стандартні відхилення розподілу по осях x та y відповідно, а \mathcal{W} - синусоїдальна радіальна частота.

Загальне рівняння для створення банку фільтрів Габора можна виразити як:

$$g_b(x, y) = a^{-m} g(\bar{x}, \bar{y}), \quad (2.2)$$

де $\bar{x} = x \cos \theta + y \sin \theta$ та $\bar{y} = -x \sin \theta + y \cos \theta$ де $\theta_k = \pi \frac{(k-1)}{n}$, $k = 1, 2, 3 \dots n$,

де n – кількість використаних орієнтацій, та a^{-m} є шкалою фільтра для $m = 0, 1, 2, \dots S$ для S шкал. Надмірність у частотній області запобігається шляхом проектування вейвлетів Габора як:

$$\sigma_u = \frac{\left(\left(\frac{U_h}{U_l} \right)^{\frac{1}{(s-1)}} - 1 \right) U_h}{\left(\left(\frac{U_h}{U_l} \right)^{\frac{1}{(s-1)}} + 1 \right) \sqrt{2 \ln 2}}, \quad (2.3)$$

$$\sigma_v = \tan \left(\frac{\pi}{2k} \right) \left[U_h - 2 \ln \left(\frac{\sigma_u^2}{U_h} \right) \right] \left[2 \ln 2 - \frac{(2 \ln 2)^2 \sigma_u^2}{U_h^2} \right]^{0.5}, \quad (2.4)$$

де U_l та U_h – позначають нижчу та вищу середні частоти відповідно, а $W = U_h$.

2.2 Лінійний дискримінантний аналіз

Лінійний дискримінантний аналіз (LDA) [30, 31] - це техніка вилучення ознак, яка здійснює пошук ознак, які найкраще розрізняють класи. Враховуючи набір незалежних ознак, LDA створює лінійну комбінацію цих ознак таким чином, що досягаються найбільші середні відмінності між класами. LDA визначає два показники: у межах матриці розсіювання класів, заданою наступною формулою:

$$S_w = \sum_{j=1}^c \sum_{i=1}^{N_j} (x_i^j - \mu) - (x_i^j - \mu_j)^T, \quad (2.5)$$

де x_i^j - i -й зразок класу j , μ_j – середнє значення класу j , c – кількість класів, N_j - кількість зразків у класі j і матриця розсіювання між класами, задана наступною формулою:

$$S_b = \sum_{j=1}^c (\mu_j - \mu) (\mu_j - \mu)^T, \quad (2.6)$$

де μ – середнє значення всіх класів. Основною метою LDA є максимізація матриці розсіювання між класами при мінімізації матриці розсіювання всередині класу.

Одним із способів цього є максимізація коефіцієнта $\frac{\det|S_b|}{\det|S_w|}$. Враховуючи, що S_w є неособливим, було доведено [40], що це відношення максимізується, коли вектори стовпців матриці проєкції є власними векторами $S_w^{-1} S_b * S_w$. S_w – максимальний ранг $N - c$ з N вибірками та c класами. Отже, для цього потрібні вибірки $N = t + c$, щоб гарантувати, що S_w не стане одиничним числом, де t - розмірність вхідних даних. Кількість зразків N майже завжди менше t , що робить матрицю розсіювання S_w особливою. Щоб вирішити цю проблему, Белхумеор [32] та Світс та Венг [33] пропонують проєктувати вхідні дані на підпростір PCA, щоб зменшити розмірність до $N - c$ або менше, перед застосуванням LDA. PCA та LDA - широко використовувані методи вилучення зовнішніх ознак при розпізнаванні зразків [34]. Отже, ми приймаємо LDA для вилучення загальних особливостей зовнішнього вигляду обличчя для оцінки вікових груп.

2.3 Локальний двійковий шаблон

Особливості текстури широко використовуються в методиках оцінки віку [35]. Локальний двійковий патерн (LBP) являє собою опис текстури, який може виявити мікроструктури моделей, такі як плями, краї, лінії і плоскі ділянки на

шкірі [36]. LBP використовується для опису текстури для розпізнавання обличчя, гендерної класифікації, оцінки віку, виявлення обличчя та відстеження компонентів обличчя. Гуней та Набієв [37] використовували LBP для характеристики особливостей текстури для оцінки віку. Вони повідомили про точність 80% для набору даних FERET [38] з використанням класифікатора найближчих сусідів та про точність 80–90% для наборів даних FERET та PIE за допомогою класифікатора AdaBoost [39]. На рисунку 2.1 показаний зразок операції 3×3 LBP.

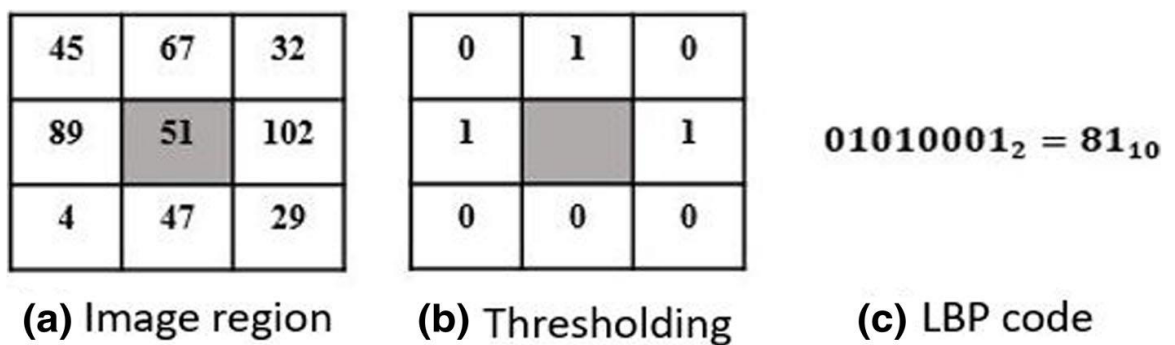


Рисунок 2.1 – Операція LBP з $P=8$, $R=1$

Об'єднання всіх 8 бітів дає двійкове число. Отримане двійкове число перетворюється в десяткове і присвоюється центральному пікселю як його LBP-код.

Ояла та ін. [41] виявив, що при використанні восьми сусідів та радіуса 1 90% усіх шаблонів складаються з однакових шаблонів. Початковий оператор LBP мав обмеження у фіксації домінуючих рис із великими масштабами структур. Останній оператор був розширений для фіксації текстурних особливостей із сусідством різних радіусів [41]. Сусідство визначає набір пікселів вибірки, розподілених рівномірно по колу кола з центром у позначеному пікселі. Білінійна інтерполяція точок, які не потрапляють у пікселі, виконується, щоб дозволити будь-який радіус і будь-яку кількість пікселів вибірки. Рівномірні шаблони можуть представляти мікроструктури у

вигляді лінії, плями, краю або плоскої області. На рисунку 2.2 показано зображення мікроструктури.

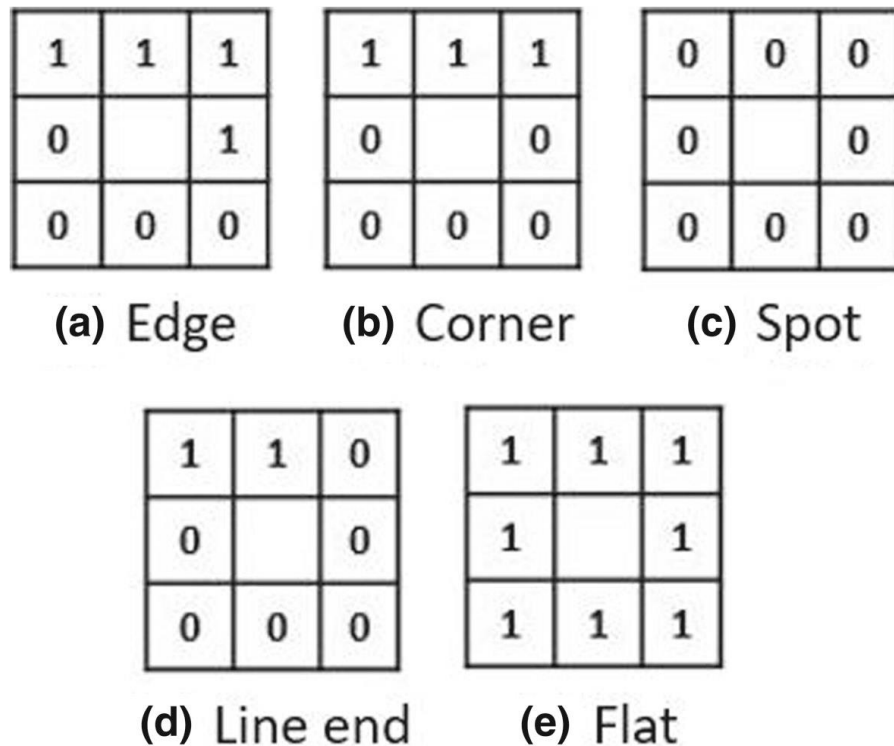


Рисунок 2.2 – Шаблон мікроструктури LBP з $P = 8$, $R = 1$

Ояла та ін. [36] далі класифікував коди LBP як однорідні та нерівномірні шаблони. Шаблон LBP з граничним двобітовим переходом від 0 до 1 або 1 до 0 класифікується як єдиний шаблон. Наприклад, шаблони 00000000, 00010000 та 11011111 є однорідними, тоді як 01010000, 11100101 та 10101001 є нерівномірними шаблонами. Для представлення n -бітового шаблону існує $n(n-1) + 2$ рівномірних шаблонів. На рисунку 2.3 показані коди LBP для зразків рівномірних шаблонів LBP (8, 1). Для вилучення інваріантних обертальних функцій за допомогою LBP, згенерований код LBP обертається по колу, доки не буде отримано його мінімальне значення [36].

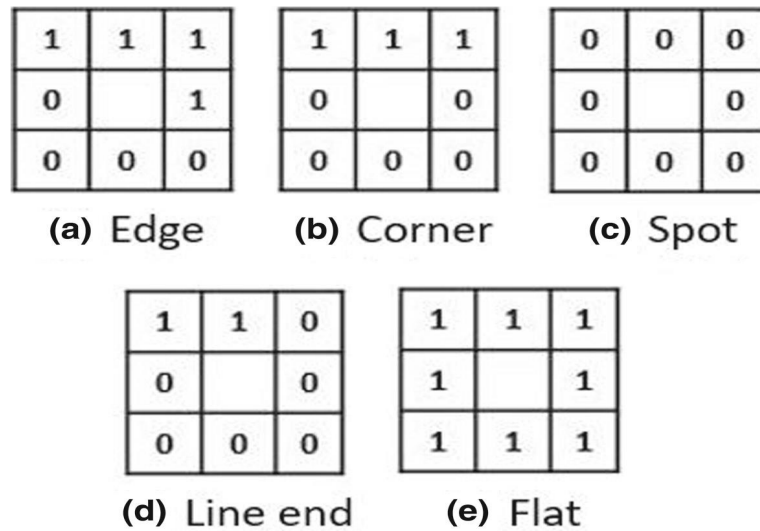


Рисунок 2.3 – Шаблон мікроструктури LBP з $P = 8$, $R = 1$

Розширений оператор LBP може захопити більше елементів текстури на зображенні, але все ж він не може зберегти просторову інформацію про ці функції. Ахонен та ін. [42] запропонував техніку поділу зображення обличчя на n клітин. Гістограми генеруються для кожної комірки, а потім об'єднуються в одну просторову гістограму. Просторова гістограма зберігає як просторові, так і текстурні описи зображення. Особливості текстури зображення нарешті представлені гістограмою кодів LBP. Гістограма LBP містить детальний дескриптор текстури для всіх структур на зображенні обличчя, таких як плями, лінії, та рівні ділянки.

3 АЛГОРИТМИ ОЦІНКИ ВІКУ

Після вилучення та представлення ознак віку наступним етапом є оцінка віку. Оцінка віку - це спеціальне завдання розпізнавання зразків, де мітки віку можна розглядати як клас або набір послідовних значень. Коли мітки віку розглядаються як класи, оцінка віку підходить як проблема класифікації, тоді як коли мітки віку розглядаються як послідовні хронологічні ряди, для оцінки віку використовується підхід регресії. Гібридний підхід також може бути використаний для оцінки віку, де як класифікаційні, так і регресійні методи інтегровані, переважно ієрархічно, для виявлення взаємозв'язку між витягнутими векторами ознак та віковими мітками.

3.1 Класифікація

Ланітіс та ін. [43] дослідив ефективність найближчого сусіда, використовуючи штучну нейронну мережу (ANN) та квадратичну функцію в задачах оцінки віку. Незважаючи на те, що квадратична функція, яка використовується для відношення подань граней до міток граней, є функцією регресії, автори називали її класифікатором квадратичних функцій [43]. Квадратична функція повідомляла MAE 5,04, що перевищувало MAE, про які повідомляв найближчий сусід. ANN та самоорганізуючі карти (SOM) повідомляють про кращу продуктивність порівняно з квадратичною функцією. Автори запропонували кластеризацію та ієрархічну оцінку віку для підвищення ефективності. Рівень помилок у розширених методах зменшився, хоча оцінки проводились на невеликих наборах даних. Також було проведено порівняння між людьми та комп'ютерами для оцінки віку, і було встановлено, що комп'ютери можуть оцінювати вік майже настільки ж надійно, як і люди. Уекі та ін. [44] побудував 11 моделей Гауса в низьковимірному просторі характеристик 2DLDA та LDA, використовуючи максимізацію очікувань (EM).

Оцінка вікової групи визначалася шляхом підгонки зображення зонда до кожного кластеру та порівняння ймовірностей. Вони повідомили про більш високу точність, 82% чоловіків та 74% жінок, із широкими віковими групами 15 років у порівнянні з 50% чоловіків та 43% жінок у вікових групах 5-річного віку. Це демонструє, що такий підхід може оприлюднювати лише точніші випадки, коли вікові групи мають широкий діапазон, а отже, не застосовуються в оцінці вікових груп вузького діапазону.

Зливна текстура та місцевий вигляд, Хуерта та ін. [45] використовував класифікацію глибокого навчання для оцінки віку. Використовуючи LBP [41], прискорені надійні функції (SURF) [46] та гістограму орієнтованих градієнтів (HOG) [47], він оцінив ефективність глибокого навчання на двох великих наборах даних і досяг рівня MAE 3,31. Ху та ін. [48] використовував інтенсивність Куллбека-Лейблера для представлення обличчя перед використанням згорткової нейронної мережі (CNN) для оцінки віку. Їхній підхід досяг MAE 2,8 на FGNET та 2,78 на MORPH II. Це демонструє, що глибоке навчання (глибокі нейронні мережі або CNN) забезпечує кращий рівень MAE порівняно з традиційними методами класифікації.

3.2 Регресія

Використовуючи 50 необроблених параметрів моделі, Ланітіс та ін. [49] досліджував лінійну, квадратну та кубічну формулювання функції старіння. Генетичний алгоритм використовується для вивчення оптимальних параметрів моделі на тренувальних зображеннях обличчя різного віку. Функція квадратичного та кубічного старіння досягла кращих MAE 0,86 та 0,75 відповідно, порівняно з 1,39 лінійної функції. Це свідчить про те, що квадратична функція пропонує найкращу альтернативу, оскільки її MAE суттєво не відрізнялася від кубічної функції, і вона не є обчислювально інтенсивною, як кубічна функція. Ян та ін. [50] сформулював задачу регресії для оцінки віку з використанням напіввизначеного програмування (SDP).

Регресору навчилися на непевних невід’ємних ярликах. Вони повідомили про MAE 10,36 та 9,79 років для чоловіків та жінок відповідно на YGA. Вони також продемонстрували, що оцінка віку за формулюванням СДП досягає кращих результатів порівняно з ANN. Обмеженням SDP є те, що це обчислювально дорого, особливо коли навчальний набір великий.

Нгуєн та ін. [51] використовував регресійну модель для оцінки віку. Зображення обличчя було представлено багаторівневим локальним двійковим шаблоном (MLBP). Їхній підхід досягнув MAE 6,6. Го і Му [52] досягли рівня MAE 4,0 за допомогою VIF для моделювання регресійної моделі для оцінки віку. Використовуючи багатоманітність інтенсивності пікселів для представлення зображення обличчя, Лу та Тан [53] оцінили свою регресійну модель на наборі даних MORPH II та отримали MAE 5,2 для білої етнічної групи та 4,2 для чорної етнічної групи. Оніфейд та ін. [54] застосував посилений регресор для локальних двійкових шаблонів *agerank* (arLBP). Вони повідомили про MAE 2,34 на FG-NET з використанням протоколу LOROvalidation.

Їхній підхід продемонстрував, що віковий рейтинг із співвідношенням моделей старіння за віковими групами покращує результати оцінки віку. Використовуючи необроблені піксельні функції, Акінемі та Оніфейд [55] досліджували класифікацію вікових груп за етнічною ознакою для оцінки віку.

Цей підхід вивчає етнічні параметри на додаток до параметрів, вивчених у [54]. Вони оцінили цю методику на наборах даних FG-NET та FAGE та повідомили про MAE 3,19 року. Їхні результати показують, що врахування етнічних параметрів покращує ефективність підходів до оцінки віку. Це можна пояснити тим, що люди різних етнічних груп старіють по-різному.

3.3 Гібридний підхід

Як обговорювалося в попередніх розділах, до завдання оцінки віку можна підходити як до класифікації, так і до проблеми регресії. Щоб вибрати між

ними, можна провести експеримент, вибравши репрезентативні класифікатори та регресори, щоб порівняти їх ефективність на одному і тому ж наборі даних, використовуючи ті самі функції. Го та ін. [57, 58] порівнював класифікатор SVM із регресором SVR. Цей експеримент показав, що SVM працює краще порівняно з SVR на наборі даних YGA, коли SVM досягає MAE 5,55 для жінок та 7,00 для чоловіків, тоді як SVR досягає 5,52 для жінок та 7,47 для чоловіків. Також повідомлялося, що SVM погано працював на FG-NET порівняно з SVR (MAE 7,16 проти 5,16 років). Цей експеримент показує, що класифікаційний підхід до оцінки віку може виконуватись краще або гірше, ніж підхід регресії, залежно від інших аспектів, таких як якість зображень у використуваному наборі даних, використувані методи вибору об'єктів та вилучення об'єктів, а також розподіл зображень за віком серед інших факторів. Поєднання класифікації та регресії може призвести до надійних та точніших систем оцінки віку. Го та ін. [56, 57] тому запропонував оцінити вік, використовуючи локально скориговану стійку регресію (LARR). LARR спочатку виконує регресію, використовуючи всі існуючі старіння зображень. Потім результати регресії використовуються для обмеження класифікатора з малим діапазоном пошуку. Вони продемонстрували, що кращих результатів оцінки віку можна досягти, поєднуючи схеми класифікації та регресії. Поєднуючи регресію та класифікацію, MAE покращився до 5,30 та 5,25 років для жінок та чоловіків відповідно на наборі даних YGA та 5,07 на наборі даних FG-NET. Обмеження методу LARR [58] полягає в тому, що він не може автоматично визначати локальний діапазон пошуку для класифікатора. Діапазон визначається евристичним випробуванням різних діапазонів і вимагає від користувача експериментального вибору найкращого рішення. Щоб автоматично визначити обмежений діапазон пошуку, Го та ін. [45] запропонував підхід на основі правдоподібності для поєднання результатів класифікації та регресії. За допомогою рівномірного розподілу результати регресії перетворюються на ймовірності, тоді ймовірності від результату класифікації обрізаються рівномірним розподілом. Це ще більше покращило точність завдяки

досягненню MAE 5.12 та 5.11 для чоловіків та жінок відповідно на YGA та 4.97 на FG-NET.

Гюне та ін. [59] представляв старіння обличчя, поєднуючи функції AAM, LBP та Gabor. Вони використали ансамбль із трьох SVM, розташованих в ієрархічному порядку, для побудови моделі оцінки віку. Першим кроком їхньої моделі було проведення оцінки вікових груп за класифікацією SVM. Потім проводили лінійну регресію для оцінки віку у віковій групі. Їхній підхід досягнув MAE 4,13 на FG-NET. Ці результати показують, що поєднання особливостей та рішень, що використовується в гібридній ієрархічній оцінці віку, може покращити помилки оцінки порівняно з класифікаційними підходами.

Хан та ін. [60] здійснив ієрархічну демографічну оцінку та порівняв продуктивність машини та людини. Вони видобували функції BIF та демографічні інформативні функції за допомогою алгоритму підсилення. Потім вони виконують ієрархічну оцінку віку, використовуючи класифікацію між групами з подальшою регресією в межах групи.

Оцінюючи цю техніку на MORPH II та FG-NET, вони досягли MAE 3,6 та 3,8 на наборах даних MORPH II та FG-NET, відповідно. Чой та ін. 2011 [29] використовував AAM, Gabor та LBP для представлення зображення обличчя. Їх гібридна модель оцінки віку досягла MAE 4,7 на FG-NET, 4,3 на PAL та 4,7 на наборах даних BERC.

Гібридний підхід до оцінки віку демонструє кращі показники порівняно з регресією та класифікацією, якщо застосовувати їх окремо. Щоб поєднати класифікацію та регресію, можна протестувати виділені ознаки на обох техніках окремо, перш ніж комбінувати їх.

4 ОПИС ПРАКТИЧНОЇ ЧАСТИНИ

4.1 Експериментальні дослідження алгоритму розпізнавання віку за зображеннями обличчя на основі вибору ієрархічних ознак

Щоб дослідити ефективність запропонованого способу для автоматичного прогнозування вікового діапазону від рис обличчя, експерименти проводяться з використанням бази даних Б. В. Кеннеді [20], наданої Техаським університетом, Даллас. Набір даних містить 180 зображень обличчя з рівномірним розподілом

як чоловіків, так і жінок. Набір даних побудований для включення людей з різним віком, статтю та етнічною приналежністю. Зображення обличчя поділено на 3 групи по 60 зображень, кожна з яких однаково збалансована за статтю, віковою групою та етнічною приналежністю. Зображення обличчя - це все з фото з розширенням 300×450 пікселів. Набір даних містить анотації, проведені 108 студентами магістрантів для збору даних, що стосуються семантичних атрибутів, таких як сприйманий вік особи з її обличчя, сприйнятий настрій та відмінність запам'ятовуваності зображення.



Рисунок 4.1 – База даних Б. В. Кеннеді для оцінки віку

Після запуску процедури вибору ознак на отриманих вихідних ознаках, яка містить 2550 ознак, вікові підписи будуються на кожному рівні двійкового дерева після автоматизованої кластеризації вікових груп. Коефіцієнт правильної класифікації (CCR) вимірюється за допомогою правила К-найближчого сусіда (KNN) з $k = 3$, використовуючи процедуру перехресної перевірки. Використовуючи метод оцінки сукупного балу відповідності (CMS), запропонований Філіпсом під час протоколу FERET, досягається поганий коефіцієнт правильної класифікації (CCR) 48,3% при використанні звичайної процедури вибору ознак одночасно для всіх класів. Цікаво, що під час запуску процесу ієрархічної класифікації досягається високий CCR у 96,1% для успішного розмежування між двома віковими групами: 18-43 та 44-100. Загальний коефіцієнт правильної класифікації для ієрархічного процесу сягає

87,8% для рангу $R = 1$. На рисунку 4.2 показана крива CMS для процесу класифікації для двох випадків. Досягнуті результати є багатообіцяючими, оскільки оцінка віку базується виключно на інформації на основі текстур, і це може бути посилено шляхом додавання геометричних властивостей обличчя.

Автоматизована оцінка віку за допомогою зображень обличчя нещодавно залучила значну кількість уваги з боку спільноти комп'ютерного зору, яка стала ключовою темою дослідження з численними додатками, починаючи від контролю доступу та індексування та отримуючи дані на основі зображень. У цьому дослідженні ми досліджуємо базований на бачення метод кластеризації та оцінки вікових груп за рисами обличчя. Локальний двійковий шаблон застосовується для видалення гібридного набору ознак, включаючи місцеві та глобальні характеристики з обличчя. Вибір ієрархічної ознаки описаний для процесу класифікації, де вікові діапазони групуються по дереву. Експериментальні результати, проведені на загальнодоступному наборі даних, підтвердили потенціал для запропонований метод для кращої оцінки вікового діапазону для різних зображень обличчя.

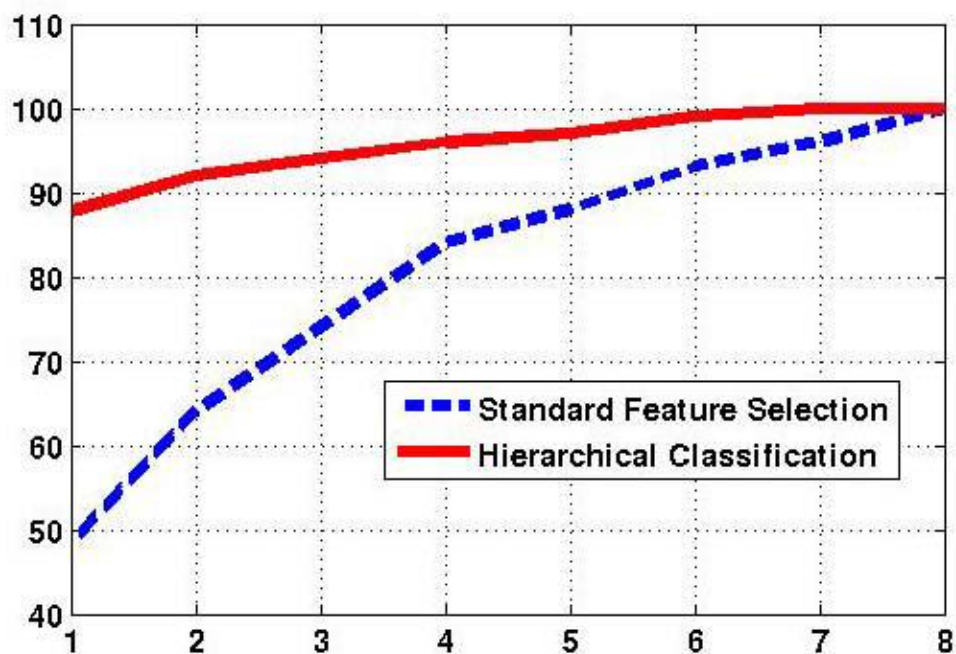


Рисунок 4.2 – Сукупний бал відповідності для оцінки віку

Таблиця 4.1 – Ієрархічна кластеризація та оцінка вікових груп

Level 1		Level 2		Level 3	
18-43	96.1%	18-32	93.3%	18-24	91.5%
				25-32	
		33-43		33-37	92.1%
				38-43	
44-100	96.1%	44-56	97.2%	44-48	90.7%
				49-56	
		57-100		57-72	94.6%
				73-100	

4.2 Експериментальні результати з використанням алгоритму KNN (К-Найближчих сусідів)

Експериментальні результати, отримані для прогнозування віку та статі, використовували 40 зразків англійською мовою. У модельних класах (літні люди, дорослі та діти обох статей) система визначення статі досягла 66% точності розпізнавання. Тоді як визначення віку досягло 55% точності розпізнавання.

Таблиця 4.2 – Процент правильного розпізнавання за віком та статі лише за допомогою мел-частотних кепстральних коефіцієнтів MFCC

MFCC	Стать вірно	Вік вірно
MFCC 1 2 3	55,55%	44,44%
MFCC 4 5 6 7 8 9	66,66%	44,44%
MFCC 10 11 12	22,22%	33,33%
Усі MFCC	66,66%	44,44%

Таблиця 4.3 – Процент правильного розпізнавання за віком та статтю разом за допомогою мел-частотних кепстральних коефіцієнтів MFCC

MFCC	Рекрівень пізнання (вік і стать разом)
MFCC 1 2 3	33,33%
MFCC 4 5 6 7 8 9	44,44%
MFCC 10 11 12	11,11%
Усі MFCC	44,44%

За допомогою функції MFCC найкращий результат був отриманий при використанні MFCC від 4 до 9. Наступні таблиці - це результати, отримані за допомогою формант:

Таблиця 4.4 – Розпізнавання віку та розпізнавання статі лише за допомогою форматних частот

Форманти	Процент правильного розпізнавання віку	Процент правильного розпізнавання статі
F3, F4	22,22%	22,22%
F3	38,88%	27,77%
F4	27,77%	44,44%
F1, F4	38,88%	55,55%
F1, F2, F3	50%	55,55%
F1, F2, F4	44,44%	50%
F1, F2	55,55%	61,11%

Продовження таблиці 4.4

F1, F3, F4	33,33%	44,44%
F1, F3	50%	55,55%
F1	38,88%	38,88%
F2, F3	55,55%	50%
F2, F3, F4	38,88%	50%
F2, F4	50%	38,88%
F2	38,88%	38,88%
F1,F2, F3, F4	50%	66,66%

Таблиця 4.5 – Розпізнавання статі та віку разом використовуючи лише формантні частоти

Форманти	Процент правильного розпізнавання (вік та стать разом)
F1, F2, F3, F4	50%
F1, F4	33%
F1, F3	39%
F1, F2	55%
F1	28%
F1, F2, F3	39%
F1, F2, F4	39%
F1, F3, F4	28%
F2, F3, F4	33%
F2, F4	33%

Продовження таблиці 4.5

F2	33%
F4	17%
F3	22%
F2, F3	39%
F3, F4	22%

Таблиця 4.6 – Процент правильного розпізнавання лише з використанням формантних частот та мел-частотних кепстральних коефіцієнтів

Формантні частоти	MFCCs	Вік і стать	Гендер	Вік
Усі	Усі	50%	66,66%	50%
F1 F2	Усі	50%	55,55%	50%
F1 F2	MFCC 1 2 3	50%	55,55%	50%
F1 F2	MFCC 10 11 12	50%	55,55%	50%
F1 F2	MFCC 4 5 6 7 8 9	50%	55,55%	50%

При використанні функції формантних частот найкращий результат був отриманий з формантами 1 і 2, найкращий результат для розпізнавання статі отриманий, коли всі форманти використовувались разом.

Таблиця 4.7 – Коефіцієнт розпізнавання статі за допомогою формантних частот (без врахування дітей)

Процент правильного розпізнавання (без дітей)	
Формантні частоти	Стать
F1 F2 F3 F4	83,33%
F1 F2	83,33%
F3 F4	91,66%
F3	66,66%
F4	66,66%
F1 F4	66,66%

Використовуючи формантові частоти, найкращий результат було отримано з використанням частот 1 і 2, оскільки основна інформація мовного сигналу зосереджена на низькочастотній частині, те саме для MFCC, але результат не покращився з MFCC, оскільки MFCC не є активним із записаними даними. в різних місцях, на різних комп'ютерах і мікрофонах, для поліпшення результатів MFCC слід поєднувати з іншими акустичними характеристиками, особливо основною частотою.

ВИСНОВКИ

В моїй атестаційній роботі необхідно було дослідити алгоритми сумісного розпізнавання віку людини за її голосом та зображенням обличчя. В ході виконання атестаційної роботи мною були розглянуті методи сумісного розпізнавання віку людини по голосу та зображенню обличчя. Також були розглянуті техніки вилучення ознак віку. Серед алгоритмів, був обраний алгоритм гібридного підходу. В практичній частині було розглянуто алгоритм KNN (K-Найближчих сусідів), а також алгоритм розпізнавання віку за зображеннями обличчя на основі вибору ієрархічних ознак.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. A. Lanitis, C. Draganova, and C. Christodoulou, “Comparing different classifiers for automatic age estimation,” *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 34, no. 1, pp. 621–628, 2004.
2. S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim, “Age estimation using a hierarchical classifier based on global and local facial features,” *Pattern Recognition*, vol. 44, no. 6, pp. 1262–1281, 2011.
3. Y. Fu, G. Guo, and T. S. Huang, “Age synthesis and estimation via faces: A survey,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 11, pp. 1955–1976, 2010.
4. D. S. Berry and L. Z. McArthur, “Perceiving character in faces: the impact of age-related craniofacial changes on social perception.” *Psychological bulletin*, vol. 100, no. 1, p. 3, 1986.
5. G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang, “Imagebased human age estimation by manifold learning and locally adjusted robust regression,” *Image Processing, IEEE Transactions on*, vol. 17, no. 7, pp. 1178–1188, 2008.
6. G. Guo, G. Mu, Y. Fu, and T. S. Huang, “Human age estimation using bio-inspired features,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009*, pp. 112–119.
7. X. Geng, Z.-H. Zhou, and K. Smith-Miles, “Automatic age estimation based on facial aging patterns,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 12, pp. 2234–2240, 2007.
8. N. Harrati, I. Bouchrika, A. Tari, and A. Ladjailia, “Automating the Automated classification of facial expressions using bag of visual words and texture-based features,” in *16 International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2015*.
9. D. M. Burt and D. I. Perrett, “Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information,” *Proceedings of the*

Royal Society of London B: Biological Sciences, vol. 259, no.1355, pp. 137–143, 1995.

10. A. J. O’toole, T. Vetter, H. Volz, E. M. Salter et al., “Threedimensional caricatures of human heads: distinctiveness and the perception of facial age,” *Perception*, vol. 26, pp. 719–732, 1997.

11. N. Harrati, I. Bouchrika, A. Tari, and A. Ladjailia, “Automating the evaluation of usability remotely for web applications via a model-based approach,” in *International Conference on New Technologies of Information and Communication (NTIC)*, 2015.

12. N. Ramanathan and R. Chellappa, “Face verification across age progression,” *Image Processing, IEEE Transactions on*, vol. 15, no. 11, pp. 3349–3361, 2006.

13. I. Bouchrika, J. N. Carter, and M. S. Nixon, “Towards automated visual surveillance using gait for identity recognition and tracking across multiple non-intersecting cameras,” *Multimedia Tools and Applications*, pp. 1–21, 2014.

14. Y. H. Kwon and N. da Vitoria Lobo, “Age classification from facial images,” in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR ’94., 1994 IEEE Computer Society Conference on*, Jun 1994, pp. 762–767.

15. J. Ylioinas, A. Hadid, X. Hong, and M. Pietikäinen, “Age estimation using local binary pattern kernel density estimate,” in *Image Analysis and Processing–ICIAP 2013*. Springer, 2013, pp. 141–150.

16. I. Bouchrika, J. N. Carter, M. S. Nixon, R. Morzinger, and G. Thallinger, “Using gait features for improving walking people detection,” in *20th International Conference on Pattern Recognition (ICPR)*. IEEE, 2010, pp. 3097–3100.

17. T. Ojala, M. Pietikäinen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions,” *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

18. C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on local binary patterns: A comprehensive study,” *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009.

19. I. Bouchrika, “Gait analysis and recognition for automated visual surveillance,” Ph.D. dissertation, University of Southampton, 2008.

20. K. M. Kennedy, K. Hope, and N. Raz, "Life span adult faces: Norms for age, familiarity, memorability, mood, and picture quality," *Experimental aging research*, vol. 35, no. 2, pp. 268–275, 2009.
21. D Gabor, *Theory of communication*. *J. Inst. Electr. Eng.* 93, 429–457 (1946)
22. SE Choi, YJ Lee, JL S, RP K, J Kim, Age estimation using hierarchical classifier based on global and local features. *Pattern Recogn.* 44, 1262–1281 (2015)
23. BS Manjunathi, WY Ma, Texture features for browsing, retrieval of image data. *IEEE Trans. Pattern Anal. Mach. Intell.* 18, 837–842 (1996)
24. F Gao, H Ai, in *Proceedings of 3rd International Conference on Advances in Biometrics*. Face age classification on consumer images with gabor feature and fuzzy lda method: lecture notes in computer science (Springer, Alghero, 2009), pp. 132–141
25. WB Horng, CP Lee, CW Chen, Classification of age groups based on facial features. *Tamkang J. Sci. Eng.* 4, 183–192 (2001)
26. H Takimoto, Y Mitsukura, M Fukumi, N Akamatsu, Robust gender and age estimation under varying facial poses. *Electron. Commun. Jpn.* 91, 32–40 (2008)
27. J Hayashi, M Yasumoto, H Ito, H Koshimizu, in *Proceedings of the 7th International Conference on Virtual Systems and Multimedia*. A method for estimating and modeling age and gender using facial image processing (IEEE, Berkeley, 2001)
28. Y Kwon, N Lobo, Age classification from facial images. *Comp. Vision Image Underst.* 74, 1–21 (1999)
29. SE Choi, YJ Lee, JL S, RP K, J Kim, Age estimation using hierarchical classifier based on global and local features. *Pattern Recognition.* 44, 1262–1281 (2011)
30. RA Fisher, The statistical utilization of multiple measurements. *Ann. Eugenics.* 8, 376–386 (1938)
31. K Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd edn. (Academic Press, San Diego, 1990)
32. PN Belhumeur, JP Hespanha, DJ Kriegman, Eigenfaces vs fisherfaces:

- recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 711–720 (1997)
33. DL Swets, JJ Weng, Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* 18, 71–86 (1996)
34. AM Martinez, AC Kak, PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 228–233 (2001)
35. G Panis, A Lanitis, N Tsapatsoulis, TF Cootes, Overview of research on facial ageing using the FG-NET ageing database. *IET Biometrics.* 5, 37–46. (2016)
36. T Ojala, M Pietikainen, D Harwood, A comparative study of texture measures with classification based on featured distribution. *Pattern Recogn.* 29, 51–59 (1996)
37. A Gunay, VV Nabiyeu, in *Proceedings of 23rd International Symposium of Computer and Information Sciences. Automatic age classification with LBP* (IEEE, Istanbul, 2008), pp. 1–4
38. JP P, H Moon, SA Rizvi, PJ Rauss, The feret evaluation methodology for face recognition algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 1090–1104 (2000)
39. Z Yang, H Ai, in *Proceedings of International Conference on Biometrics. Demographic classification with local binary patterns* (Springer, Seoul, 2007), pp. 464–473
40. RA Fisher, The statistical utilization of multiple measurements. *Ann. Eugenics.* 8, 376–386 (1938)
41. T Ojala, M Pietikainen, T Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 971–987 (2002)
42. T Ahonen, A Hadid, M Pietikainen, in *Proceedings of European Conference on Computer Vision. Face recognition with local binary patterns* (Springer, Prague, 2004), pp. 469–481
43. A Lanitis, C Draganova, C Christodoulou, Comparing different classifiers for automatic age estimation. *IEEE Trans. Man Syst. Cybern.* 34, 621–628 (2004)

44. K Ueki, T Hayashida, T Kobayashi, in Proceedings of IEEE Conference on Automatic Face and Gesture Recognition. Subspace-based age group classification using facial images under various lighting conditions (IEEE, Southampton, 2006), pp. 43–48
45. I Huerta, C Fernandez, C Segura, J Hernando, A Prati, A deep analysis on age estimation. *Pattern Recognit. Lett.* 68, 239–249 (2015)
46. H Bay, T Tuytelaars, LV Gool, Surf: Speeded up robust features. *Comput. Vis.-ECCV*. 3951, 404–417 (2006)
47. B Triggs, N Dalal, in Proceedings of IEEE on Computer Vision and Pattern Recognition. Histograms of oriented gradients for human detection (IEEE, San Diego, 2005), pp. 886–893
48. Z Hu, Y Wen, J Wang, M Wang, R Hong, S Yan, Facial age estimation with age difference. *IEEE Trans. Image Process.* 1–13 (2016)
49. A Lanitis, J Taylor, TF Cootes, Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 442–455 (2002)
50. S Yan, H Wang, X Tang, TS Huang, in Proceedings of IEEE Conference on Computer Vision. Learning auto-structured regressor from uncertain non-negative labels, (2007)
51. DT Nguyen, SR Cho, KR Park, Age estimation-based soft biometrics considering optical blurring based on symmetrical sub-blocks for mlbp. *Symmetry*, 1882–1913 (2015)
52. G Guo, G Mu, in Proceedings of IEEE Conference on Face and Gesture Recognition. Joint estimation of age, gender and ethnicity:cca vs pls (IEEE, Shanghai, 2013), pp. 1–6
53. J Lu, Y Tan, Ordinary preserving manifold analysis for human age and head pose estimation. *IEEE Trans. Hum. Mach. Syst.* 43, 249–258 (2013)
54. OFW Onifade, DJ Akinyemi, A groupwise age ranking framework for human age estimation. *Int. J. Image Graphics Signal Process.* 5, 1–12 (2015)

55. JD Akinyemi, OFW Onifade, in Proceedings of IEEE Symposium on Technologies for Homeland Security. An ethnic-specific age group ranking approach to facial age estimation using raw pixel features (IEEE, Waltham, 2016), pp. 1–6
56. G Guo, Y Fu, TS Huang, C Dyer, in Proceedings of IEEE Workshop on Applications of Computer Vision. Locally adjusted robust regression for human age estimation (IEEE, Copper Mountain, 2008)
57. G Guo, Y Fu, C Dyer, T Huang, Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Trans. Image Process.* 17, 1178–1188 (2008)
58. G Guo, Y Fu, TS Huang, C Dyer, in Proceedings of IEEE Workshop on Applications of Computer Vision. Locally adjusted robust regression for human age estimation (IEEE, Copper Mountain, 2008)
59. A Gunay, VV Nabiyeu, Facial age estimation based on decision level fusion of amm, lbp and gabor features. *Int. J. Adv. Comput. Sci. Appl.* 6, 19–26 (2015)
60. H Han, XO CLiu, AK Jain, Demographic estimation from face images: human vs machine performance. *IEEE Trans. Pattern Anal. Mach. Intell.* 37, 1148–1161 (2015)