

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ

**НАВЧАЛЬНО–НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

Кафедра інженерії програмного забезпечення

**Пояснювальна записка
до магістерської роботи
на ступінь вищої освіти магістр
на тему: «РЕАЛІЗАЦІЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ
РОЗПІЗНАВАННЯ ЕМОЦІЙ ІЗ ЗАСТОСУВАННЯМ НЕЙРОННИХ
МЕРЕЖ»**

Виконав: студент 6 курсу, групи ПДМ-61

спеціальності:

121 Інженерія програмного забезпечення

(шифр і назва спеціальності)

Кутц А. В.

(прізвище та ініціали)

Керівник Дібрівний О. А.

(прізвище та ініціали)

Рецензент _____

(прізвище та ініціали)

Нормоконтроль Трінтіна Н.А.

(прізвище та ініціали)

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
Кафедра Інженерії програмного забезпечення
Ступінь вищої освіти - «Магістр»
Спеціальність – 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ
Завідувач кафедри
інженерії програмного забезпечення

“ ____ ” _____ 2022 року

ЗАВДАННЯ
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

Кутц Андрій Валерійович

(прізвище, ім'я, по батькові)

1. Тема роботи: «Реалізація інтелектуальної системи розпізнавання об'єктів із застосуванням нейронних мереж»

Керівник роботи Дібрівний Олесь Андрійович

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом вищого навчального закладу від “ ____ ” __ 2021 року № ____.

2. Строк подання студентом роботи _____

3. Вихідні дані до роботи:

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) _____

4.1 Аналіз існуючих розробок та визначення вимог

4.2 Аналіз та архітектура програмного забезпечення

4.3 Концепція додатку

4.4 Розробка архітектури

5. Перелік графічного матеріалу.

1. Титульний слайд

2. Мета, об'єкт дослідження, предмет дослідження

3. Аналіз аналогів

4. Технічне завдання

5. Програмне забезпечення

6. Діаграма діяльності

7. Зовнішній вигляд додатку

8. Висновки

6. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів магістерської роботи	Строк виконання етапів роботи	Примітка
1	Вибір теми проекту	01.09.2020 - 31.10.2020	
2	Постановка задачі	01.11.2020 – 20.11.2020	
3	Вибір засобу реалізації	21.11.2020 – 10.06.2021	
4	Розробка архітектури	11.06.2021- 22.07.2021	
5	Розробка візуальної частини	23.07.2021- 29.08.2021	
6	Розробка функціональної частини	30.08.2021- 30.11.2021	
7	Здача диплому в деканат		

Студент _____ Кутц А.В.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Дібрівний О. А.
(підпис) (прізвище та ініціали)

РЕФЕРАТ

Текстова частина дипломної роботи 70 с., 16 рис., 94 джерел.

«Реалізація інтелектуальної системи розпізнавання емоцій із застосуванням нейронних мереж»

Актуальність дипломної роботи полягає в використанні нейронних мереж для розв'язання формалізованих задач в інтелектуальному аналізі даних. Зростання обсягів інформації, а також розширення кола технічно складних задач прийняття рішень вимагають систематизації наявних методів і розробки нових методик і алгоритмів рішення. У магістерській роботі розглядається можливість застосування нейронної мережі при розв'язанні задачі розпізнавання емоцій на фотографіях і відеозаписах.

У роботі проведено аналіз наявних проблем при створенні нейронних мереж для розпізнавання емоцій та запропоновано рішення у вигляді розробленої нейронної мережі.

Об'єкт роботи: створення інтелектуальної системи для розпізнавання емоцій за допомогою нейронних мереж.

Предмет роботи: нейронна мережа для розпізнавання емоцій.

Мета роботи: створення варіанту розв'язання проблеми розпізнавання емоцій обчислювальними машинами та інтелектуальними системами за допомогою нейронних мереж.

Ключові слова: інтелектуальна система, нейронна мережа, розпізнавання емоцій.

ЗМІСТ

Вступ	7
1. ОСНОВНІ ДОСЛІДЖЕННЯ В ОБЛАСТІ РОЗПІЗНАВАННЯ ЕМОЦІЙ ..	10
1.1 Ідеальна система для розпізнавання виразу обличчя	12
1.2 Розпізнавання обличчя та виділення ознак.....	12
2. ФОН НЕЙРОННОЇ МЕРЕЖІ	18
2.1 Конформація та функція нейрона.....	18
2.2 Коротка історія штучних нейронних мереж	19
2.3 Вибір моделі нейронної мережі та теорія зворотного поширення	20
3. ОПТИМАЛЬНІ МЕТОДИ: POWELL AND DOWNHILL SIMPLEX	26
3.1 Метод встановлення напрямків Пауелла	26
3.2 Симплексний алгоритм спуску	27
4. ПРОЕКТУВАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ЕМОЦІЙ	28
4.1 Загальний опис	28
4.2 Вага	28
4.3 Основні параметри та їх комбінація	30
4.4 Кількість прихованих шарів і прихованих вузлів	30
4.5 Об'єднання наборів напрямків Пауелла	31
4.6 Емпіричні методи	33
4.6.1 Нормалізація виводу	33
4.6.2 Використання пікових кадрів даних	33
4.6.3 Медіанний фільтр	33
4.6.4 Ігнорування всіх нейтральних значень	34
5. ЕКСПЕРИМЕНТИ	35
5.1 Налаштування експерименту	35
5.2 Результати тесту на наборі даних Кона-Канада — Емпіричний метод .	41
5.2.1 One_hidden_layer нейронні мережі для тестування емпіричних методів	41
5.2.2 Нейронні мережі для спеціальних категорій емоцій	49
5.3 Результати тестування бази даних Кон-Канаде	52

5.4 Результати тесту на наборі даних Кона-Канада (різні початкові ваги) ..	57
ВИСНОВОК	58
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	62

ВСТУП

У світі штучні інтелектуальні системи набули великої популярності завдяки своїм багатим можливостям і ефективністю використання.

Існує безліч сфер, де існує ряд завдань вирішення яких можливе з використанням інтелектуальних систем: економіка та бізнес, робототехніка, геологорозвідка, математика, біофізика, авіоніка, системи охорони та безпеки, медицина та інші [32, с. 34].

Така велика кількість областей застосування вказує на те, що Інтелектуальні системи – це унікальний набір для вирішення завдань аналізу та обробки великих обсягів даних, вирішення завдань різного рівня складності.

На даний час системи штучного інтелекту активно використовуються штучні нейронні мережі.

Нейронна мережа – це обчислювальна структура, що складається з набору однотипних елементів. Ці елементи виконують прості функції, і всі процеси, що відбуваються в штучній нейронній мережі можуть бути пов'язані з процесами, що відбуваються в нервовій системі живих організмів.

Нейронні мережі нелінійні за своєю природою, у них немає явної залежності, що дозволяє відразу використовувати розроблену технологію (Інформаційна модель нейронної мережі). Протягом багатьох років лінійне моделювання було основним методом моделювання, тому що для нього були добре розроблені процедури оптимізації.

Актуальність магістерської роботи полягає у використанні нейронних мереж для вирішення слабо формалізованих завдань у видобутку даних. Збільшення обсягу інформації, а також розширення кола технічно складних завдань прийняття рішень вимагає систематизації наявних методів та розробки нових методик та алгоритмів ухвалення рішень. У цій магістерській роботі розглядається можливість використання нейронної мережі для розв'язання задачі розпізнавання людських емоцій.

Штучні нейронні мережі мають великі перспективи в розвитку, а програмне

забезпечення має величезні переваги від їх використання. Крім того, кожне реалізоване завдання має необмежений та нестандартний набір методів рішення.

Темою магістерської роботи є реалізація інтелектуальної системи розпізнавання емоцій із застосуванням нейронних мереж.

Основною метою роботи є вибір інформаційної моделі нейронної мережі та реалізація алгоритму розпізнавання рухової активності особи. Нейронна мережа має бути оптимальною за своєю внутрішньою структурою та способом керування інформаційними потоками між нейронами.

Вибрана інформаційна модель буде використана для вирішення практичного завдання.

Основними завданнями дипломної роботи є:

1) Вивчення наявних типів штучних інтелектуальних систем і методів їх функціонування;

2) Вивчення основних типів інформаційних моделей штучних нейронних мереж. Вибір оптимальної інформаційної моделі нейронної мережі для розв'язання задачі розпізнавання емоцій;

3) Вивчення наявних методів розпізнавання виразів обличчя та вибір універсального методу;

4) Реалізація та опис алгоритму розпізнавання рухової активності особи для інтелектуальної системи та вирішення практичного завдання.

Об'єктом цього дослідження є підходи, методи розпізнавання міміки. Предметом дослідження є інформаційні моделі штучних нейронних мереж, а також реалізація та опис алгоритму розпізнавання рухової активності особи. Наукова новизна магістерської роботи полягає в використанні нейромережевих технологій (інформаційних моделей), а також системи рухової активності особи для реалізації алгоритму розпізнавання емоцій людини. Для реалізації нейронної мережі, яка розпізнаватиме емоції, необхідно визначити ключові риси особи та основні лицьові прояви, які відрізняють конкретну людину від багатьох інших. Складність реалізації полягає у навчанні нейронної мережі. Вибір вихідних даних для вирішення поставленого завдання - це, перш за все, математичний та

геометричний опис виразів обличчя.

У зв'язку зі наростаючою складністю розв'язуваних науково-технічних завдань, автоматична обробка та аналіз візуальної інформації стають дедалі більш актуальними питаннями. На даний час вже існують цифрові та інтелектуальні камери, а також програмне забезпечення, які обробляють зображення.

Сучасні технології швидко рухаються до того моменту, коли комп'ютери стануть самостійними системами, здатними до самонавчання та зростання "особистості", як це відбувається з людьми з самого їх народження. Вже зараз багато розробників розв'язують складні задачі за допомогою машин комп'ютерного зору, використовуючи системи за допомогою різних методів, у тому числі нейронних мереж.

У третьому розділі представлений алгоритм розпізнавання емоцій з застосуванням нейронної мережі.

1. ОСНОВНІ ДОСЛІДЖЕННЯ В ОБЛАСТІ РОЗПІЗНАВАННЯ ЕМОЦІЙ

На початку 1970-х років Пол Екман та його колеги провели широкі дослідження, що стосувалися виразу обличчя людини [20]. Вони знайшли докази, що підтверджують універсальність у виразах обличчя.

Ці «універсальні вирази обличчя» представляють щастя, смуток, гнів, страх, здивування та огиду. Вони вивчали вирази обличчя в різних культурах, включаючи дописьменні культури, і знайшли багато спільного у вираженні та розпізнаванні емоцій на обличчі. Однак вони також помітили відмінності у виразах і запропонували модель, у якій вираз обличчя регулюється «правилами відображення» в різних соціальних контекстах.

Наприклад, японські та американські дослідження зафіксували схожі вирази обличчя під час перегляду одного і того ж фільму. Проте в присутності влади японські глядачі неохоче демонстрували свої справжні вирази обличчя. З іншого боку, немовлята демонструють широкий спектр виразів обличчя з самого народження, що свідчить про те, що ці вирази є вродженими.

Екман і Фрізен [19] розробили систему кодування дій на обличчі (FACS) для кодування виразів обличчя, де рухи на обличчі описуються набором одиниць дії (AU). Кожна АС має певну пов'язану м'язову основу. Ця система кодування виразів обличчя виконується вручну, дотримуючись набору встановлених правил. Вхідними є нерухомі зображення виразів обличчя, часто на піку виразу. Цей процес дуже трудомісткий.

Роботи Екмана надихнули багатьох дослідників аналізувати вирази обличчя за допомогою обробки зображень і відео. Відстежуючи риси обличчя та вимірюючи кількість рухів обличчя, вони намагаються класифікувати різні вирази обличчя. Нещодавні роботи з аналізу та розпізнавання виразів обличчя [13] використовували ці «базові вирази» або їх підмножину. У [4] Пантік і Роткранц надають поглиблений огляд багатьох досліджень, проведених в області автоматичного розпізнавання виразів обличчя за останні роки.

Робота з комп'ютерної кількісної оцінки виразів обличчя почалася лише в

1990-х роках. Массе [3] використовував оптичний потік для розпізнавання виразів обличчя. Він був одним із перших, хто застосував методи обробки зображень для розпізнавання виразів обличчя. Ланітіс та ін. [6] використовували гнучку форму та модель зовнішнього вигляду для кодування зображень, ідентифікації людини, відновлення пози, розпізнавання статі та розпізнавання виразу обличчя.

Блек і Якоб [16] використовували локальні параметризовані моделі руху зображення для відновлення нетвердого руху. Після відновлення ці параметри використовувалися як вхідні дані для класифікатора на основі правил для розпізнавання шести основних виразів обличчя.

Якуб і Девіс [17] обчислили оптичний потік і використали подібні правила для класифікації шести виразів обличчя. Розенблум, Якоб і Девіс [21] також обчислили оптичний потік областей на грані, а потім застосували мережу радіальних базисних функцій для класифікації виразів. Есса і Пентленд [22] використовували метод на основі області оптичного потоку для розпізнавання виразів. Донато та ін. [23] перевірили різні особливості для розпізнавання АС на обличчі та визначення виразу обличчя в кадрі.

Такахіро Оцуко та Джун Ойя [24] спочатку обчислили оптичний потік, а потім обчислили коефіцієнти двовимірного перетворення Фур'є, які були використані як вектори ознак для прихованої моделі Маркова (НММ) для класифікації виразів. Навчена система змогла розпізнати один із шести виразів майже в реальному часі (близько 10 Гц). Крім того, вони використовували відстежені рухи для контролю виразу обличчя анімованої системи Кабукі [25].

Мартінес [28] запровадив підхід до індексації, заснований на ідентифікації зображень лобового обличчя за різних умов освітлення, виразів обличчя та оклюзій. Байєсівський підхід був використаний для пошуку найкращої відповідності між місцевими спостереженнями та моделлю вивчених місцевих особливостей, а НММ використовувався для досягнення хорошого розпізнавання, навіть якщо нові умови не відповідали умовам, які були раніше на етапі навчання. Олівер та ін. використовував відстеження нижньої частини обличчя для виділення рис форми рота та використовував їх як вхідні дані для системи розпізнавання

виразу обличчя на основі НММ (розпізнавання нейтрального, щасливого, сумного та відкритого рота).[29]

Ці методи подібні тим, що спочатку виділяють деякі ознаки із зображень, потім ці ознаки використовуються як вхідні дані в систему класифікації, а результатом є одна з попередньо вибраних категорій емоцій. Вони відрізняються головним чином ознаками, витягнутими з відеозображень, і класифікаторами, які використовуються для розрізнення різних емоцій.

1.1 Ідеальна система для розпізнавання виразу обличчя

Є три основні фактори для створення системи розпізнавання виразів обличчя, а саме розпізнавання обличчя, виділення рис обличчя та класифікація емоцій.

Ідеальний аналізатор емоцій повинен розпізнавати суб'єктів незалежно від статі, віку та будь-якої етнічної приналежності. Система повинна бути інваріантною до різних умов освітлення та відволікання, як-от окуляри, зміни зачіски, волосся на обличчі, вуса, борода тощо, а також повинна мати можливість «заповнити» відсутні частини обличчя та побудувати ціле обличчя. Він також повинен виконувати надійний аналіз виразу обличчя, незважаючи на значні зміни умов перегляду, жорсткі рухи тощо. Хорошою системою відліку є зорова система людини [4]. Нинішні системи далекі від ідеалу, і вони мають довгий шлях для досягнення цих цілей.

1.2 Розпізнавання обличчя та виділення ознак

Більшість систем розпізнають обличчя в контрольованих умовах, наприклад, без волосся на обличчі/окулярів, будь-яких жорстких рухів голови, перший кадр має бути нейтральною емоцією тощо, і, отже, сьогодні довільне виявлення обличчя повертає великі наміри [4].

Зазвичай розпізнавання обличчя виконується двома способами. При

цілісному підході обличчя визначається як єдине ціле, тоді як при аналітичному підході виявляються лише деякі важливі риси обличчя.

Після виявлення обличчя є 2 способи виділити риси. У цілісній моделі обличчя використовується метод на основі шаблону. В аналітичній моделі обличчя будуть використовуватися методи на основі ознак, щоб відстежувати риси обличчя, поки люди демонструють вираз обличчя.

У нашій системі ми в основному зосереджуємось на частині класифікації емоцій, а не на розпізнаванні обличчя чи виділенні рис обличчя. Для виділення рис обличчя ми використовуємо систему розпізнавання виразу обличчя в реальному часі, розроблену Себе і Коеном [1] (див. рисунок 2.1).

Ця система складається з частини відстеження обличчя, яка виводить вектор ознак руху певних областей обличчя. Ознаки використовуються як вхідні дані для класифікатора.

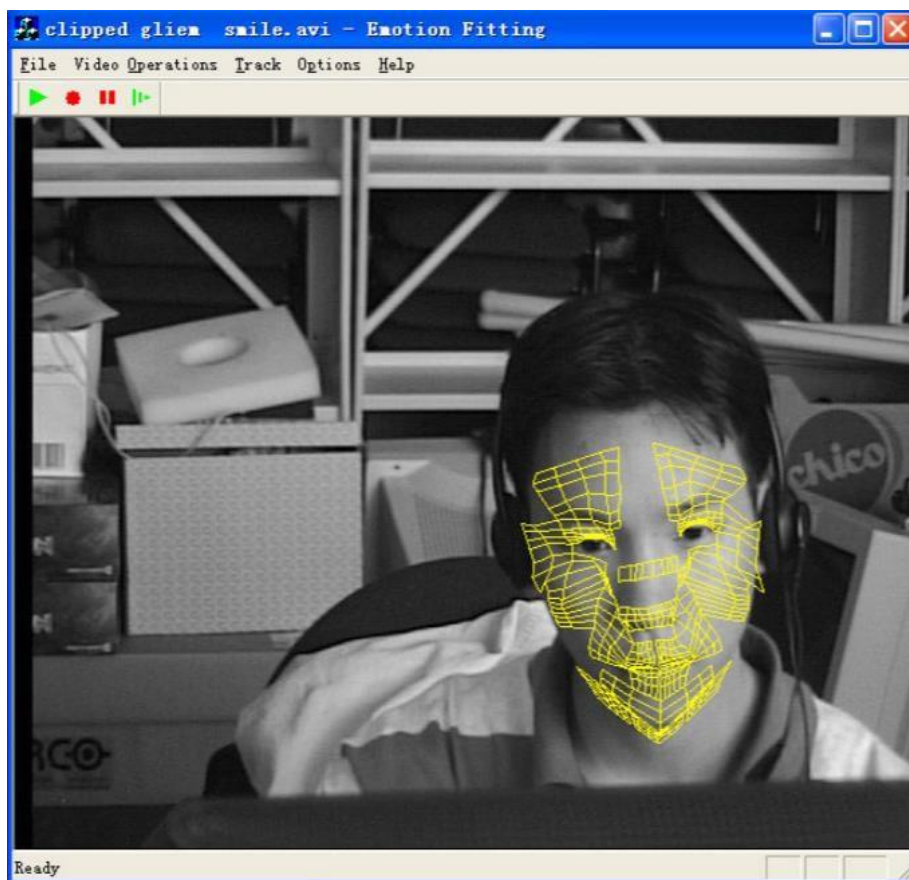


Рисунок 1.2.1 - Знімок трекера обличчя в реальному часі

Цей трекер обличчя використовує підхід на основі моделі, коли

створюється явна 3D-каркасна модель обличчя.

У першому кадрі послідовності зображень визначні риси обличчя, такі як куточки очей та куточки рота, вибираються інтерактивно. Загальна модель обличчя потім деформується, щоб відповідати вибраним рисам обличчя. Модель обличчя складається з 16 поверхневих латок, вбудованих в томи Без'є. Визначені таким чином ділянки поверхні гарантовано будуть безперервними та гладкими. Зміна розташування контрольних точок в об'ємі Без'є може змінити форму сітки. Перш ніж описувати об'єм Без'є, почнемо з кривої Без'є.

Враховуючи набір контрольних точок b_0, b_1, \dots, b_n , відповідна крива Без'є (або крива Бернштейна-Без'є) визначається як (1.1)

$$x(u) = \sum_{i=0}^n b_i B_i^n(u) = \sum_{i=0}^n b_i \binom{n}{i} u^i (1-u)^{n-i} \quad (1.1)$$

де форма кривої контролюється контрольними точками b_i та u в діапазоні між $[0,1]$.

У міру переміщення контрольних точок виходить нова форма відповідно до поліномів Бернштейна $B_n(u)$ у наведеному вище рівнянні. Зміщення точки на кривій можна описати в термінах лінійних комбінацій переміщень контрольних точок.

Об'єм Без'є є прямим розширенням кривої Без'є і визначається як $V=BD$, записаним у матричній формі. У цьому рівнянні V — зміщення вузлів сітки, D — матриця, стовпці якої — вектори зміщення контрольної точки об'єму Без'є, а B — відображення в термінах поліномів Бернштейна. Іншими словами, зміну форми моделі обличчя можна описати через деформації в D .

Після того, як модель сконструйована та встановлена, можна відстежити рух голови та локальні деформації таких рис обличчя, як брови, повіки та рот.

Спочатку вимірюються рухи 2D-зображення за допомогою шаблонного узгодження міжкадрами з різною роздільною здатністю. Шаблони зображень із попереднього кадру та з самого першого кадру використовуються для більш надійного відстеження.

Виміряні рухи 2D-зображення моделюються як проекції справжніх 3D-рухів на площину зображення. З двовимірних рухів багатьох точок на сітці тривимірний рух можна оцінити, розв'язавши надвизначену систему рівнянь проєктивних рухів у сенсі найменшого квадрата.

Відновлені рухи представлені в термінах величин певного попередньо визначеного руху різних рис обличчя. Кожен рух об'єкта відповідає простій деформації на грані, визначеній параметрами регулювання гучності Без'є. Ми називаємо ці вектори руху одиницями руху (MU). Зауважимо, що вони подібні, але не еквівалентні AC Екмана, і є числовими за своєю природою, що представляють не тільки активацію ділянки обличчя, але також напрямок та інтенсивність руху.

MU, що використовується в трекері обличчя, показано на рисунку 1.2.2 і описано в таблиці 1.2.1.

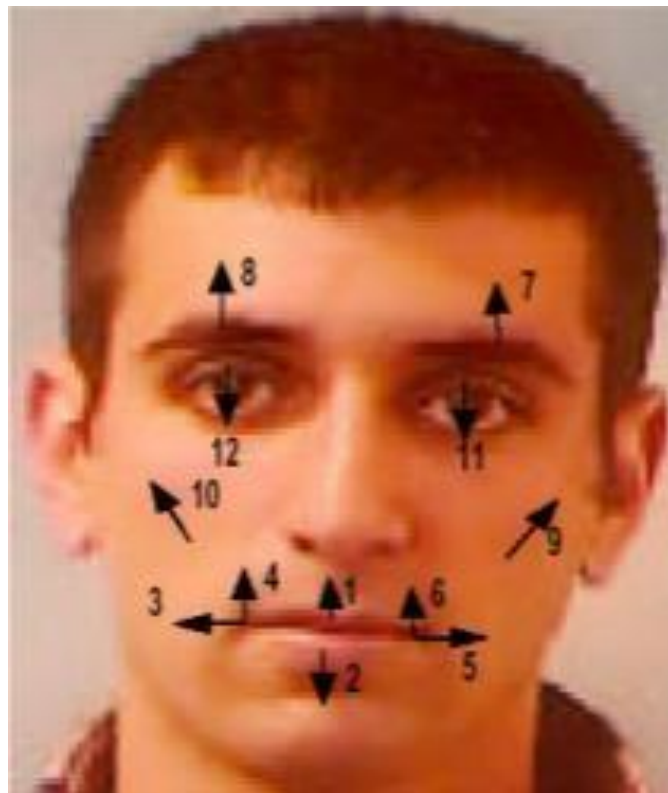


Рисунок 1.2.2 12 вимірювань рухів обличчя

Таблиця 1.2.1 Одиниці руху, що використовуються в трекері обличчя

MU	Опис
1	Вертикальний рух центру верхньої губи
2	Вертикальний рух центру нижньої губи
3	Горизонтальний рух лівого кута рота
4	Вертикальний рух лівого кута рота
5	Горизонтальний рух правого кута рота
6	Вертикальний рух правого кута рота
7	Вертикальний рух правої брів
8	Вертикальний рух лівої брови
9	Піднімання правої щоки
10	Піднімання лівої щоки
11	Моргання правим оком
12	Моргання лівим оком

Кожен вираз обличчя моделюється як лінійна комбінація МУ (1.2):

$$V = B[D_0 D_1 \dots D_m] \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_m \end{bmatrix} = BDP \quad (1.2)$$

де кожен з D відповідає МУ, а P_i — відповідні величини (або коефіцієнти) кожної деформації. Загальний рух голови та обличчя (1.3):

$$R(V_0 + BDP) + T \quad (1.3)$$

де R — матриця тривимірного обертання, T — матриця тривимірного трансляції, а V_0 — початкова модель грані. МУ використовуються як основні ознаки для схеми класифікації [1].

2. ФОН НЕЙРОННОЇ МЕРЕЖІ

Одна з головних цілей штучного інтелекту — імітувати людський мозок. Нині можливості пам'яті, швидкість обчислень та інші логічні функції числового комп'ютера вже перевищують людський мозок, але все ще є багато функцій людського мозку, які не можуть бути змодельовані комп'ютером, тобто здатність людини розпізнавати й приймати рішення. Теорія нейронних мереж базується на дослідженні моделювання методів обробки інформації людиною.

2.1 Конформація та функція нейрона

Штучні нейронні мережі спочатку були розроблені, щоб певним чином моделювати функціональність біологічних нейронних мереж, які є частиною людського мозку. Наш мозок містить близько 1011 нейронів. Кожен біологічний нейрон складається з тіла клітини, сукупності дендритів, які переносять електрохімічну інформацію в клітину, і аксона, який передає електрохімічну інформацію з клітини.

Нейрон виробляє вихід уздовж свого аксона, тобто він спрацьовує, коли сукупний ефект його входів досягає певного порогу. Аксон одного нейрона може впливати на дендрити іншого нейрона через з'єднання, які називаються синапсами. Деякі синапси будуть генерувати позитивний ефект в дендриті, тобто той, який спонукає його нейрон спрацьовувати, а інші вироблять негативний ефект, тобто такий, який відбиває нейрон від спрацьовування. Один нейрон отримує вхідні дані приблизно від 105 синапсів, а загальна кількість синапсів у нашому мозку може бути порядку 10^{16} . Досі не зрозуміло, як саме наш мозок навчається і запам'ятовує, але, схоже, це пов'язано із взаємозв'язками між нейронами. (тобто в синапсах).

Штучні нейронні мережі намагаються змоделювати цю низькорівневу функціональність мозку. Це контрастує з символічним міркуванням високого рівня у штучному інтелекті, який намагається моделювати процеси міркування високого

рівня в мозку. Коли ми думаємо, що усвідомлюємо, що маніпулюємо поняттями, до яких ми додаємо імена (або символи), напр. для людей чи предметів ми не усвідомлюємо низького рівня електрохімічних процесів, що відбуваються під ними. Аргумент на користь підходу нейронної мережі до ШІ полягає в тому, що, якщо ми зможемо правильно моделювати низько рівневі дії, функціональність високого рівня може бути створена як властивість, що виникає.

2.2 Коротка історія штучних нейронних мереж

У 1943 році психолог Маккаллок і математик Піттс створили першу модель нейронної мережі під назвою МР-модель.

1943 – 1970 рр. було розроблено різноманітні моделі та алгоритми навчання: Хебб вийшов зі своїм правилом навчання; модель перцептрона була розроблена Розенблатом та ін.

1970-1980 рр. для цієї місцевості був період відливу; розроблена до цього часу теорія мала багато неповноти.

1980-1990 рр., після того, як американський вчений-фізик Дж. Хопфілд пропагував нейронну мережу Хопфілда (HNNS) і досяг прогресу в проблемі мандрівного продавця, почали з'являтися численні моделі та алгоритми.

Після 1990 року ця сфера постійно розвивається. Практично існує багато застосувань у багатьох областях, включаючи розпізнавання образів, стиснення даних, кодування, аналіз шифру, аналіз запасів, економічне управління та оптимальний контроль.

2.3 Вибір моделі нейронної мережі та теорія зворотного поширення

У досліджених нами статтях літератури зворотне поширення [4, 14] було широко використовуваною моделлю для класифікації емоцій. Чи є якась особлива причина, яка робить зворотне поширення кращим, ніж численні інші моделі нейронних мереж? Або нам варто спробувати деякі нові моделі, щоб побачити, чи

працюють вони краще, ніж «популярне» зворотне поширення? Деякі відповіді на ці запитання можна знайти в [15] (див. таблицю 3.3.1). SOM, також відома як самоорганізуюча карта Кохонена, яка є методом кластеризації, який можна використовувати для автоматичного розуміння природи даних без нагляду. Ми можемо перетворити цю неконтрольовану нейронну мережу в контрольовану нейронну мережу LVQ (Learning Vector Quantum), що є ще одним методом кластеризації. У таблиці ми можемо зробити висновок, що при обробці шуму та багаторазового введення даних зворотне поширення працює краще, ніж SOM. LVQ відмінно підходить для класифікації, але при обробці шуму трохи гірше, ніж зворотне поширення. Беручи до уваги всі ці факти, ми вирішили використовувати зворотне поширення як наш основний метод.

Таблиця 2.2.3 Порівняння різних моделей нейронних мереж.

Тип мережі	Тип проблеми		Обмеження даних	
	Класифікація	Прогноз	Шумний: Здатність досправляються з дуже шумними даними (наприклад, на складі ринкові моделі, маркетинг тощо)	Багато вхідних даних: Ефективність з великим числом полів введення
Зворотне поширення	Достатній	Відмінний	Відмінний	Відмінний
SOM	Достатній	Середній	Достатній	Достатній

LVQ	Відмінний	Дужепоганий	Достатній	Відмінний
-----	-----------	-------------	-----------	-----------

SOM посилається на самоорганізуючу карту Кохонена, а LVQ відноситься до нейронної мережі Learning Vector Quantum.

На рисунку 2.3.1 показано простий нейрон (персептрон) у нейронній мережі з $n-1$ входами, 1 зміщенням і 1 виходом. Загальна кількість вхідних подразників до цього нейрона у вихідному шарі становить (1.4):

$$z_{in} = \sum_{i=0}^n x_i w_i = x_0 w_0 + x_1 w_1 + x_2 w_2 + \dots + x_n w_n \quad (1.4)$$

Виходом цього нейрона є функція активації $f(z_{in})$, де f — функція активації.

Під час експериментів функція активації визначається як (1.5):

$$f(x) = \frac{1}{1 + e^{-\alpha x}} \quad (1.5)$$

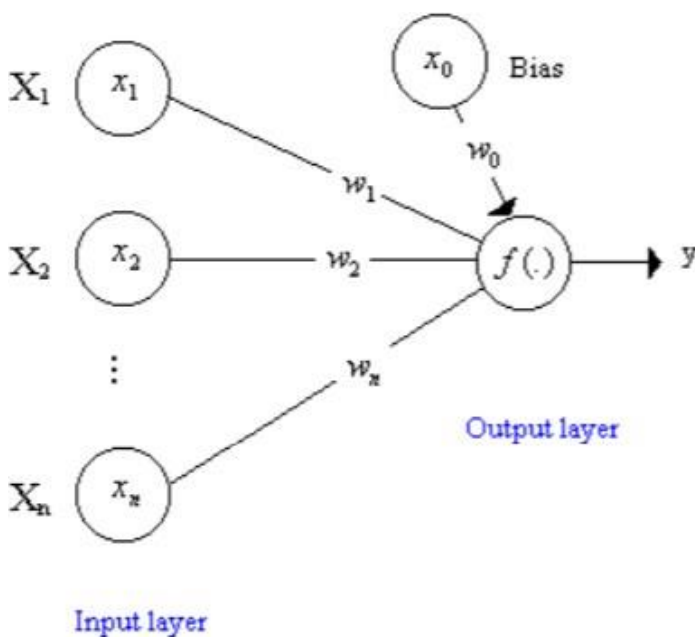


Рисунок 2.3.1 - Один персептрон в нейронній мережі

Після того як нейрон виводить своє значення, зазвичай існує різниця між цим значенням і його правильним значенням, цільовим значенням. Різниця визначається як функція помилки (1.6):

$$E = \frac{1}{2} \sum_{k=1}^m (t_k - y_k)^2 \quad (1.6)$$

Ми використовуємо частковий диференціал цієї функції помилки, щоб змінити вагові коефіцієнти назад, щоб зменшити помилку. Ця процедура називається зворотним поширенням. Рівняння модифікації ваг з використанням часткового диференціала функції помилки задається так (1.7):

$$\Delta w_{jk} = -\alpha \frac{\partial E}{\partial w_{jk}} \quad (1.7)$$

де α - швидкість навчання, а w_{jk} - вага між нейроном j і нейроном k .

Після ряду модифікацій ваг частковим диференціалом функції помилки похибка зменшиться до певної міри, що можна прийняти. На цьому етапі нейронну мережу можна використовувати для перевірки даних. На рисунку 2.3.2 показано повне зображення мережі одного прихованого шару з кількома нейронами, де x — вхідні дані, z — приховані вузли, y — вихід, t — цільове значення, w і v — ваги, зміщення зазвичай дорівнює 1.

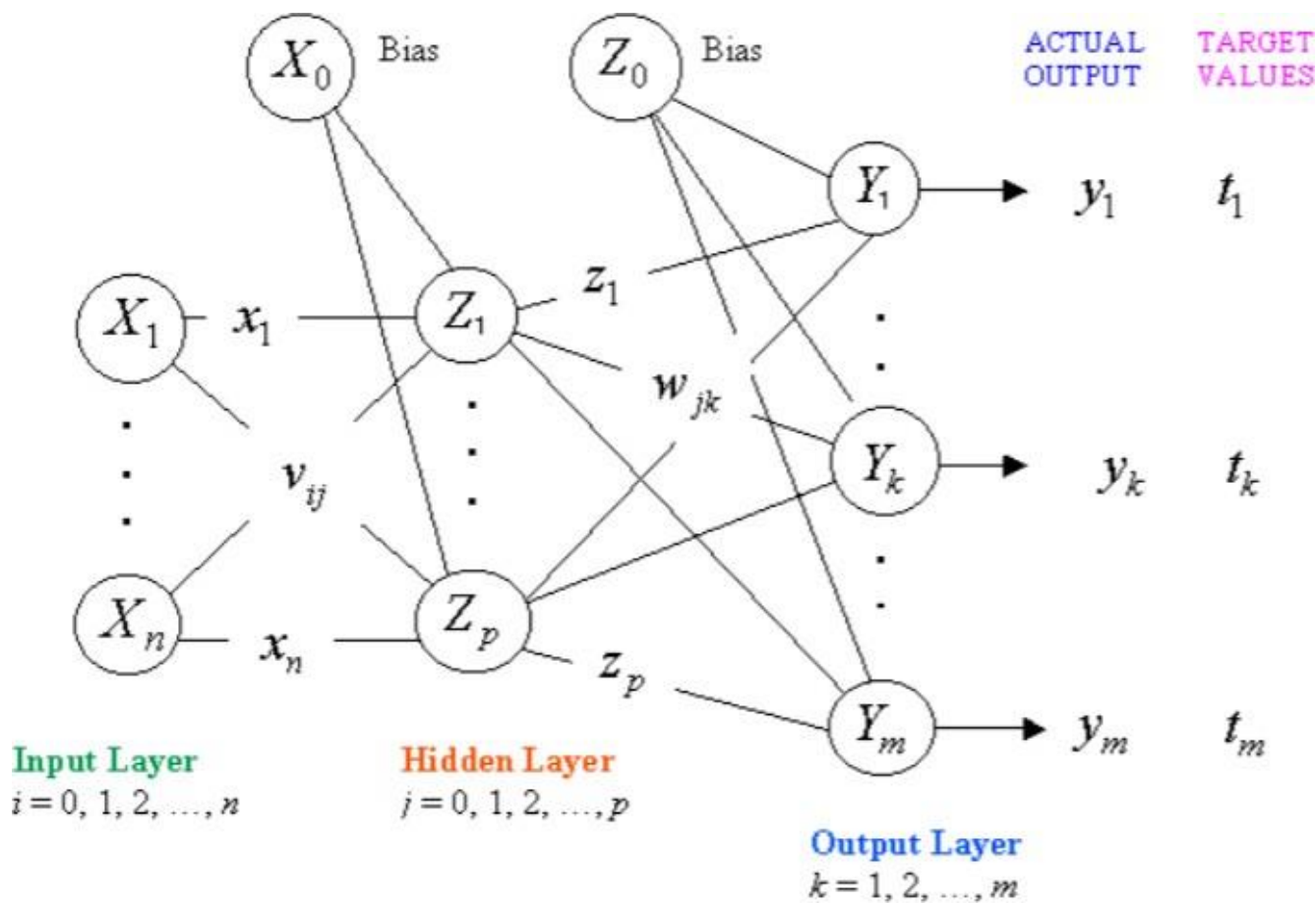


Рисунок 2.3.2 Повна фігура мережі одного прихованого шару з кількома нейронами

Алгоритм для одного прихованого шару наведено на наступній сторінці. У цьому алгоритмі i представляє кількість вхідних вузлів, j — кількість прихованих вузлів, k — кількість вихідних вузлів, w_{jk} являє ваги між нейроном k у вихідному шарі та нейроном j у прихованому шарі, а v_{ij} — ваги між нейронами j у прихованому шарі та нейроном i у вхідному шарі.

Теоретично, нейронна мережа зворотного поширення двох прихованих шарів може моделювати будь-яку довільну функцію, тому ми також спробували двошарову мережу, щоб перевірити, чи вона працює краще, ніж одношарова мережа. Для двошарової мережі крок дії подібний до дії одного прихованого шару, за винятком додавання та зміни ваги одного додаткового шару.

Поетапна дія одношарового зворотного поширення (1.8).

Ініціалізуйте ваги до малих випадкових значень під час (умова зупинки
хибна)

{

для (кожен шаблон навчання)

{

Встановіть всі x_i для $i = 1, 2, \dots, n$ для $j = 1, 2, \dots, p$

{

$$z_{in_j} = \sum_{i=1}^n x_i v_{ij}$$

$$z_j = f(z_{in_j})$$

(1.8)

}

Для $k = 1, 2, \dots, m$

{

$$y_{in_k} = \sum_{j=1}^p z_j w_{jk}$$

$$y_k = f(y_{in_k})$$

}

$$\delta_k = (t_k - k_k) f'(y_{in_k})$$

$$\Delta w_{kj} = \alpha \delta_k z_j$$

$$m$$

$$\delta_j = f'(z_{in_j}) \sum_{k=1} \delta_k w_{jk}$$

$$\Delta v_{ij} = \alpha \delta_j x_j$$

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk}$$

$$v_{jk}(new) = v_{jk}(old) + \Delta v_{ij}$$

}/* end of for loop */

}/* end of while loop */

3. ОПТИМАЛЬНІ МЕТОДИ: POWELL AND DOWNHILL SIMPLEX

Суть процедури навчання нейронної мережі полягає у знаходженні набору ваг, які можуть мінімізувати функцію помилки. Типовим способом нейронної мережі є поширення назад часткового диференціала функції помилки. Нам цікаво дізнатися, чи працює традиційний оптимальний алгоритм для нейронних мереж. Серед розглянутих методів мінімізації та максимізації, представлених у [8], є три методи багатовимірного введення, які не вимагають, щоб функція мала першу похідну. Такими методами є імітований відпал, симплекс на схилі та метод Пауелла. Оскільки вже є повідомлення про поєднання імітованого відпалу з нейронними мережами, ми вирішили спробувати два останні методи в наших експериментах.

3.1 Метод встановлення напрямків Пауелла

Припустимо, що ми починаємо з точки P в N -вимірному просторі і рухаємося звідти в напрямку, визначеному вектором E . Тоді будь-яку функцію з N змінних $f(P)$ можна мінімізувати таким чином: Візьмемо одиничні вектори e_1, e_2, \dots, e_N як набір напрямків, рухайтесь по першому напрямку до мінімуму, потім звідти по другому напрямку до мінімуму і так далі, перебираючи весь набір напрямків стільки разів, скільки необхідно, поки функція не перестане зменшуватися. [8] Ініціалізуйте набір напрямків u_i до базових векторів (уже визначених векторів або дуже малого випадкового значення), $u_i = e_i$ з $i = 1, \dots, N$. Тепер повторюйте таку послідовність кроків, поки цільова функція не перестане зменшуватися (1.9):

Збережіть вихідну позицію як P_0 .

Для $i = 1, \dots, N$ перемістіть P_{i-1} до мінімуму вздовж напрямку u_i і назвіть цеточка P_i .

Для $i = 1, \dots, N-1$ встановіть $u_i = u_{i+1}$. Установіть $u_{n=P_n} = P_0$. (1.9)

Перемістіть P_N до мінімуму вздовж напрямку u_N і назовемо цю точку P_0 [8].

3.2 Симплексний алгоритм спуску

Симплекс — це геометрична фігура в N вимірах з $N + 1$ точок (або вершин).

У двох вимірах симплекс є трикутником. У трьох вимірах це тетраедр [8].

Симплекс-метод вниз по схилу починається не просто з однієї точки, а з $N + 1$ точок, визначаючи початковий симплекс.

Припустимо, що будь-яка з цих точок є початковою точкою P_0 , а потім візьміть інші N точок як $P_i = P_0 + e_i$, де $e_i \in N$ одиничних векторів [8].

Симплекс-метод спуску тепер виконує ряд кроків, більшість кроків просто переміщує точку симплексу, де функція найбільша («найвища точка») через протилежну грань симплексу до нижчої точки. Ці кроки називаються відображеннями, і вони побудовані для збереження об'єму симплексу (отже, збереження його невивроженості). Коли він може це зробити, метод розширює симплекс в тому чи іншому напрямку, щоб зробити більші кроки. Коли він досягає «дна долини», метод стискається в поперечному напрямку і намагається сочитися вниз по долині. Якщо є ситуація, коли симплекс намагається «пройти крізь вушко голки», він стискається в усіх напрямках, втягуючи себе навколо своєї найнижчої (найкращої) точки. Тоді можна завершити, коли векторна відстань, переміщена на цьому кроці, буде вдвічі меншою за величиною, ніж деякий допуск [8].

У наступному розділі є докладний опис того, як об'єднати представлені алгоритми в нейронну мережу.

4. ПРОЕКТУВАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ЕМОЦІЙ

4.1 Загальний опис

Тепер, маючи всі базові знання, ми можемо приступити до проектування нейронних мереж для розпізнавання емоцій. Дані про 12 ознак, отримані від трека обличчя, використовуються як вхідні дані для 12 вхідних вузлів нейронної мережі. Вихідний шар містить 2-7 вузлів, які представляють категорії емоцій, залежно від різних мереж. Існує 1 або 2 прихованих шари, а кількість прихованих вузлів коливається від 1 до 29x29. Швидкість навчання, число імпульсу та параметр функції активації сигмовидної кривої автоматично налаштовуються під час процедури навчання. В одних мережах розглядається метод Пауелла, а в інших комбінується набір емпіричних способів, тобто беруть пікові кадри послідовності даних емоцій, сортують навчальний набір, видаляють частину емоцій, нормалізують вихід, встановлюють поріг. до ваг тощо.

Результати тестування базуються на базі даних Кон-Канаде та окремо на автентичній базі даних. Функція активації, яку ми використовували, - це сигмовидна функція. У наступних розділахми представляємо опис усіх параметрів та результатів їх комбінацій в експериментах.

4.2 Вага

Зворотне поширення є пошуком градієнтного спуску, тому його легко зупинити на локальному мінімумі, тоді як випадково вибрані ваги допомагають уникнути цього [14]. Якщо ваги занадто великі, мережі мають тенденцію до перенасичення. Рішення полягає в тому, щоб гарантувати, що після ініціалізації ваги та до навчання вихідні дані всіх нейронів мали невелике значення між $[-0,5, 0,5]$. Ми ініціалізуємо вагові коефіцієнти випадковою функцією та ігноруємо ті ваги, які перевищують певний поріг, який також можна налаштувати як один із параметрів мережі.

Одне питання: чи слід під час тренування обмежувати ваги? Це частково залежить від того, наскільки великий вхід, оскільки сигмовидна функція дуже близька до одиниці, коли вхід більше 10 [14]. Оскільки наші дані про функції для введення дуже малі, зазвичай менші за 2, ми встановлюємо поріг ваги, ігноруючи ті ваги, які перевищують це обмеження під час навчання. Виявилось, що це не має великої різниці в покращенні показника попадання.

З іншого боку, коли ми намагалися встановити дуже суворий поріг у 2-шаровій нейронній мережі під час процедури навчання, іноді це призводило до жахливої продуктивності 2-шарової мережі. Це пов'язано з тим, що параметри активації, які ми встановили в мережі з 2 прихованими шарами, не відповідали цьому порогу і викликали насичення нейрона. Отже, ми надали дуже великий поріг після ініціалізації ваг, щоб уникнути подібних проблем.

Інша річ, про яку ми повинні подбати, це початкова точка, яка також може вплинути на напрямок пошуку, щоб знайти хороший локальний мінімум (див. Рисунок 4.2.1). Якщо ми почнемо з точки А, ми отримаємо глобальний мінімум, а з С — локальний мінімум. Тому ми повинні спробувати різні вихідні точки, ініціалізувавши ваги різними випадковими значеннями. Ми перевірили це в деяких мережах і виявили, що крива точності коливається, але не дуже сильно.

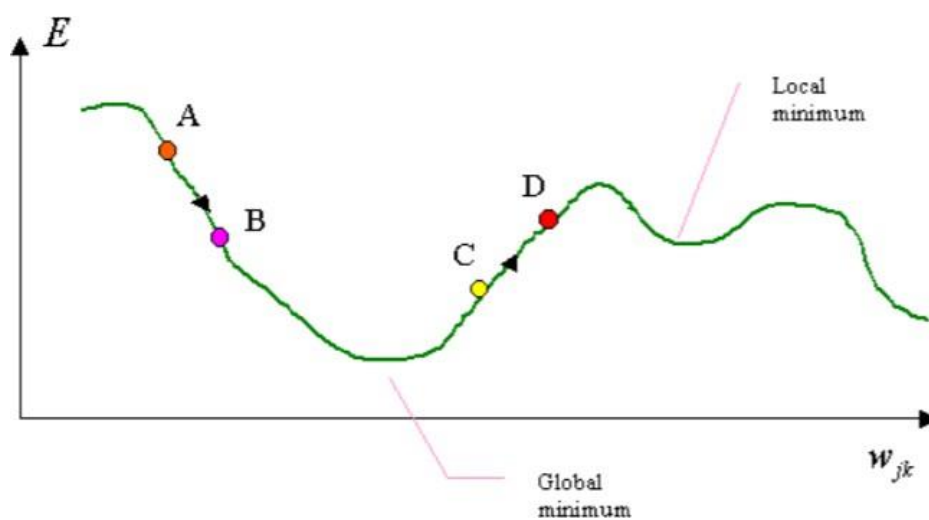


Рисунок 4.2 - Локальний мінімум і глобальний мінімум

4.3 Основні параметри та їх комбінація

При розробці нейронних мереж необхідно встановити деякі критичні параметри, тобто швидкість навчання α , число імпульсу λ і параметр функції активації σ .

Швидкість навчання визначається швидкістю навчання α . Число імпульсу переносить зміну ваги, таким чином, воно має тенденцію згладжувати простір ваги помилок. Неправильне значення σ призведе до насичення нейрона.

Загалом, продуктивність нейронних мереж буде дуже жахливою, якщо ці значення не підбрані правильно. На жаль, немає точних правил чи математичного визначення того, коли слід вибирати, яке саме значення цих параметрів. Зазвичай налаштування параметрів здійснюється емпірично. Чи допоможе 30 знайти кращі комбінації, якщо ми дозволимо комп'ютеру виконувати частину роботи? Ми спробували це в такий спосіб.

Спочатку ми визначили три різні категорії. Збільшення або зменшення розміру кроку α , λ і σ задається за допомогою введення або макровизначення. Це залежить від того, як часто ці категорії слід змінювати під час процедури навчання, наприклад, для тих категорій, які потребують невеликого втручання, ми даємо макровизначення для ефективності навчання. Коли досягається краща точність, показники разом із параметрами, які дають змогу досягти такої точності, записуються у файл. Повторюємо тренування до тих пір, поки точність не перестане покращуватися на кілька ходів. Під час тестування ми будуємо нейронну мережу, зчитуючи ці параметри з файлу.

Оскільки навчання з повними комбінаціями цих параметрів займає велику кількість часу, ми спробували лише частину з цих трьох комбінацій. Тому можна пропустити деякі кращі їх комбінації.

4.4 Кількість прихованих шарів і прихованих вузлів

Кількість прихованих шарів і прихованих вузлів також вплине на

продуктивність. Мережа не створить гарної моделі проблеми, якщо використовується занадто мало прихованих нейронів, тоді як занадто багато прихованих нейронів призведе до поганого узагальнення. Тільки з належною кількістю прихованих нейронів, правильною швидкістю навчання, числом імпульсу та параметром функції активації нейронна мережа може добре працювати. Знову ж таки, це забирає дуже багато часу, щоб автоматично спробувати різну кількість прихованих вузлів з різними комбінаціями інших ключових параметрів.

Таким чином, хоча код був розроблений для автоматичного вибору комбінації, під час практичного навчання ми все одно емпірично вибираємо деякі типові розміри (кількість прихованих вузлів) мережі як відправну точку.

4.5 Об'єднання наборів напрямків Пауелла

Ми спробували багато комбінацій параметрів мережі зворотного поширення, але, здається, рівень розпізнавання залишається на рівні приблизно 60%-76,8% (результати в базі даних Кон-Канада). Цей факт спонукав нас спробувати деякі покращення. Один очевидний вибір — спробувати деякі з існуючих оптимальних алгоритмів. Вони розглядають нейронну мережу як функцію, яка повертає помилку, яку можна мінімізувати. Ми розглянули декілька алгоритмів мінімізації [8]. Лише імітований відпал, метод Пауелла та симплекс на схилі можуть обробляти багатовимірні вхідні дані і не потребують першої похідної функції. Спочатку ми спробували спускний симплекс, але коли спробували цей метод на звичайній функції, наприклад (1.10):

$$f(x, z) = (x-20) * (x-20) + (z-400) * (z-400), \quad (1.10)$$

він не міг обробляти більше, ніж двовимірне введення, тому ми зупинилися на цьому.

Що стосується методу Пауелла, ми спробували використовувати до 6-вимірної функції введення (1.11):

$$\begin{aligned}
Y(x_1, x_2, x_3, x_4, x_5, x_6) &= (x_1 - 20, 0) * \\
&* (x_1 - 20, 0) + (x_2 - 40, 0) * (x_2 - 40, 0) + (x_3 - 3, 0) * \\
&* (x_3 - 3, 0) + (x_4 - 40, 0) * (x_4 - 40, 0) + (x_5 - 30) * \\
&* (x_5 - 30) + (x_6 - 95) * (x_6 - 95)
\end{aligned}
\tag{1.11}$$

Він може успішно знайти всі значення, які можуть мінімізувати y : 20, 40, 3, 40, 30, 95. Тому ми припускаємо, що він може працювати з довільним розміром.

Ми встановлюємо ваги мережі як вхідні дані функції X , розмірність — це кількість усіх ваг. Якщо це мережа `2_hidden_layer`, просто розрахуйте кількість усіх ваг незалежно від того, на якому рівні знаходиться вага. Виберіть випадковий набір напрямків з дуже малим значенням, потім почніть мінімізацію помилок і змініть ваги за процедурою Пауелла.

На жаль, цього разу нам не вдалося отримати хороший показник розпізнавання. Порівняно зі звичайною нейронною мережею зворотного поширення, цьому методу знадобилося кілька хвилин, щоб досягти показника 30%, тоді як нейронній мережі зворотного поширення знадобилося не більше півхвилини, щоб досягти такого ж успіху. Тому ми не продовжували в цьому напрямку.

Виникає закономірне запитання: чи допоможе, якщо ми почнемо з навченої ваги, встановленої зі зворотного шляху поширення?

Ми спробували цю процедуру, але результати коливалися, тому ми не змогли зробити висновок про те, чи завжди цей метод буде покращуватися з наявними навченими вагами.

Виходячи з продуктивності та швидкості навчання для досягнення однакової швидкості розпізнавання, ми можемо зробити висновок, що нейронні мережі з вагами, оптимізованими шляхом зворотного поширення диференціала помилок, перевершують інші методи. Причина може полягати в тому, що під час модифікації ваг за допомогою зворотного поширення він вважає частковий диференціал функції помилки своїм орієнтиром, тоді як у методі Пауелла напрямки вибираються випадковим чином. Тому більше немає гідної інформації, і

легше потрапити в локальний мінімум.

4.6 Емпіричні методи

4.6.1 Нормалізація виводу

Зворотне поширення – це контрольований спосіб навчання та тестування. У нашому випадку як дані тесту, так і навчальні дані позначаються цілим числом від 1 до 7, що представляє 7 можливих виразів: нейтральний, радість, здивування, злість, огида, страх і сум. З іншого боку, вихід нейронної мережі в кожному шарі є float, тому ми нормалізували значення кінцевого вихідного шару та встановили максимальне значення на 1, інші на 0. У деяких випадках наших експериментів, особливо в 1_hidden_layer нейронні мережі, ця процедура допомогла покращити швидкість розпізнавання.

4.6.2 Використання пікових кадрів даних

Кадри даних безперервні, нейтральні емоції завжди на початку кожної послідовності емоцій, після чого слідує серія кадрів «справжньої» емоції дані. Таким чином, у деяких наших мережах ми зібрали лише перші 3 кадри нейтральні дані та 3 кадри «справжніх» емоційних даних у середині вираз, і ігнорував усі інші кадри, які в основному складаються з перехідних емоцій дані. Ця процедура повинна певною мірою підвищити точність, але є недоліком полягає в тому, що після фільтрації всіх проміжних кадрів, кількість кадрів для навчання різко зменшується. Це, мабуть, і є причиною зниження рівня класифікації.

4.6.3 Медіанний фільтр

У наборі даних Кон-Канаде та створеному нами наборі автентичних даних, зокрема емоція представлена набором послідовних кадрів. Цей факт змусив нас спробувати а Процедура медіанної фільтрації на виходах нейронних мереж з інтенцією покращення рівня класифікації. Ідея полягає в тому, що коли визнається значення емоції (значення, що надається нейронною мережею вихідний шар)

відрізняється від свого наступного сусіда і його попереднього сусіда, він повинен бути таким неправильне значення. Середній фільтр скине його до значення його сусідів. Розмір околиці налаштує покращення, надане цим методом. В даний час, Середній фільтр сприяє підвищенню точності приблизно від 1% до 5%, але нормально.

4.6.4 Ігнорування всіх нейтральних значень

У підмножині Кон-Канаде, яку ми використовуємо, нейтральні дані становлять приблизно $1/3$ весь навчальний набір, що може призвести до більшої настройки параметрів нейронної мережі за нейтральними даними. Хоча в практичному застосуванні нейтральних даних може бути найменше Цікавий вираз, у деяких експериментах ми вирішили ігнорувати всі нейтральні дані для підвищення точності. Ця процедура дещо покращила показник.

Сортувати та вибрати навчальний набір. У деяких експериментах ми сортували навчальний набір даних і використовували точно так само кількість даних кожної емоції в деяких наших одношарових мережах, щоб збалансувати вплив кожної емоції на всю мережу. Реорганізація (сортування і вибір) даних допомагає підвищити точність від 1% до 10%. Але як це впливає на 2-приховані мережі, ми не пробували.

Про перенавчання. Перенавчання може виникнути, коли кількість навчальних даних перевищує певну кількість значення. Це значення буде відрізнятися при зміні параметрів. Це досить трудомістко щоб навчити кожну мережу бачити, скільки даних має бути в кожній групі параметри, особливо коли ми все ще шукаємо найкращий набір параметрів. Ми перевірили перетренованість у кількох випадках, але ми не дійшли висновку щодо оптимальна кількість навчальних даних, яка не призведе до перетренування нейронної мережі.

5. ЕКСПЕРИМЕНТИ

5.1 Налаштування експерименту та результати для автентичних даних

У досліджених нами роботах усі експериментальні результати базуються на базах даних, які не фіксували природні емоції, напр. суб'єкту було вказано показати певну емоцію перед камерою.

Ідеальним способом отримати дані про природні емоції є використання секретної камери і, таким чином, не повідомляти людині заздалегідь, що її знімають. Однак такий тип експериментів може спричинити деякі етичні проблеми, тому ми тестували лише з нашими друзями та колегами. Їм сказали, що ми проводимо психологічний тест або тест на якість зображення після декомпресії. Ми дозволяємо їм переглянути короткий період (9 хвилин) відеороликів, витягнутих з фільму жахів і комедії, огидних, дивовижних відеороликів. Деяким з них ми також показали цілий фільм жахів. Згодом, коли ми пояснили їм наш намір, більшість з них були розважені та трохи здивовані.

Знайти достатню кількість людей, які б витратили 2 години на перегляд цілого фільму в особливих ситуаціях і наказали їм не говорити і не їсти, сидіти на належній відстані і правильно позувати, щоб камера могла зловити їхні обличчя, — це дуже виснажлива і складна задача. Тому ми вирішили використовувати короткі відеоролики в більшості випадків, навіть якщо вони недостатньо довгі, щоб нарощувати більшість емоцій, таких як сум, страх, гнів або навіть здивування та відраза.

Було також складно та багато часу знайти відповідні відео та фільми, які могли б відразу викликати емоції людей, тобто важко викликати у людей огиду, щоб показати вираз обличчя, не дратуючи їх. Перед записом справжніх експериментів ми перепробували багато відеороликів на деяких викладачах у LIACS і кілька разів змінювали відеоролики-кандидати. Інша проблема полягала в тому, куди сховати камеру нормального розміру для таємних зйомок. Іноколи ми використовували штору, іноді камера стояла за отвором у книжковій папці, а іноді

— у великій мавпі-іграшці.

Незважаючи на всі ці труднощі, загалом ми протестували 28 осіб. Проте деякі люди не завжди сиділи в належній позі, або розмовляли та рухалися. Інші навіть не виявляли жодних емоцій під час перегляду коротких кліпів, оскільки вони лише дуже серйозно дивилися на якість зображень, як їм було сказано! Зрештою, ми відібрали дані з 15 осіб, у тому числі 4 китайки, 1 китайський чоловік, 2 голландки та 8 голландців. Ми не ставили жодних обмежень щодо їх віку.

Ми проводили цей тест переважно з короткими кліпами, тому швидше за все було отримано лише окремі вирази, такі як нейтральність, радість, а іноді й огида. Китайські дівчата, які брали участь у всьому 2-годинному експерименті, показали кілька емоцій, таких як радість, страх, огида та здивування. Деякі з них зізналися, що були дуже напружені та налякані, але не показали очевидних емоцій, як ми очікували. Вони виглядали дуже серйозно, а інші були здивовані, але змішані зі страхом. Враховуючи все це, ми вирішили відібрати дані нейтральні, радості та огиди.

На рисунку 4.1.1 показано три кадри однієї людини з нашої автентичної бази даних, які показують різні вирази: нейтральне, радість та огиду. На рисунку 4.1.2 показано приклади двох людей із бази даних Кона-Канада, які показують 6 різних виразів.



Рисунок 5.1.1 Зображення з автентичної бази даних



Рисунок 5.1.2 Зображення з бази даних Cohn-Kanade

Після запису емоції ми використали Adobe Premiere та Virtual Duba нотувати дані, а також засіб відстеження обличчя, згаданий у главі 2, для вилучення ознак. В загалом, 11 осіб використовуються для навчального набору, а інші 4 як тестові особи. Ми також провів експеримент, використовуючи навчальний набір, що складається з 70% даних з 11 людей, тоді як тестовий набір склав решту 30% даних. Результати наведені в таблиці 5.1.1.

Таблиця 5.1.1 Результати на основі автентичних даних.

	Спокій	Радість	Відраза
Спокій	96,08%	3,92%	0,0%
Радість	4,67%	95,33%	0,0%
Відраза	1,59%	0,0%	98,41%

Середній показник класифікації становить 96,6%. Тренувальний набір складається з 70% даних від 11 осіб і тестового набору – це решта 30% даних. Параметри мережі: 1 прихований шар, 10 прихованих вузлів, $\lambda = 0,55$, $\alpha = 0,42$, $\sigma = 0,3$, кількість тестових даних дорівнює 264, кількість навчальних даних — 677, із медіанним фільтром і нормованим виводом.

Для порівняння ми також спробували класифікувати лише нейтральні, огидні та радість з Кон-Канаде.

Результати наведені в таблиці 5.1.2.

Таблиця 5.1.2 Результати набору даних Кона-Канада.

	Спокій	Радість	Відраза
Спокій	82.78	8.71	8.51
Радість	19.67	77.05	3.28
Відраза	25.00	6.82	68.18

Параметри мережі: 1 прихований шар, 10 прихованих вузлів, $\lambda = 0,5$, $\alpha = 0,3$, $\sigma = 0,3$, тестовий вектор номер 687, навчальний вектор число 2300, використовувати медіанний фільтр і нормований вихід. У середньому тут 76% нижче, ніж з автентичної бази даних.

Хоча в базі даних Кон-Канаде 7 емоцій, а в нашій лише 3, але при обробці лише 3 емоцій в базі даних Cohn_Kanade, середнє значення швидкості розпізнавання також значно нижча, ніж отримана з автентичної бази даних. Однак поки ще рано робити висновок про різницю між цими двома результатами наведені в таблиці 5.1.1 і таблиці 5.1.2.

Оскільки під час вилучення даних для автентичній базі даних, ми підібрали пікові кадри вручну, таким чином проігнорувавши перехідна емоція. З іншого боку, взявши пікові кадри від Cohn_Kanade База даних була зроблена за кодом, із фіксованою піковою позицією кадру, про що було важко сказати був справжнім піком цієї емоції.

Причина, по якій ми вибрали найбільш очевидні дані про емоції з наших відео, полягає в тому, що ми хочемо знати, наскільки добре нейронні мережі будуть працювати з емоціями, які можуть легко відрізнити людину. Таблиця 5.1.1 показує, що це може добре працювати в цьому випадку, але як це працює з

даними переходу емоцій, невідомо.

Таблиця 5.1.3 Результати на основі автентичних даних.

	Спокій	Радість	Відраза
Спокій	100%	0,00%	0,00%
Радість	-	-	-
Відраза	0,00%	0,00%	100%

Тест з предмета 1 (не входить до навчального комплекту). Параметри мережі:

1 прихований шар з 10 прихованими вузлами, $\lambda = 0,55$, $\alpha = 0,36$, $\sigma = 0,3$, кількість тестові дані 22, кількість навчальних даних 677, із медіанним фільтром і нормованим виходом. «-» означає, що відповідних вимірювань не було.

Таблиця 5.1.4. Результати на основі автентичних даних.

	Спокій	Радість	Відраза
Спокій	100%	0,00%	0,00%
Радість	33,33%	66,67%	0,00%
Відраза	0,00%	0,00%	100%

Тест з предмета 2 (не входить до навчального комплекту). Параметри мережі:

1 прихований шар з 10 прихованими вузлами, $\lambda = 0,55$, $\alpha = 0,3$, $\sigma = 0,3$, кількість тестові дані 100, кількість навчальних даних 677, із медіанним фільтром і нормованим виходом.

Таблиця 5.1.5 Результати за автентичним набором даних, тест за темою 3 (не входить до навчального набору)

	Спокій	Радість	Відраза
Спокій	100%	0,00%	0,00%
Радість	0,00%	100%	0,00%
Відраза	—	—	—

Параметри мережі: 1 прихований шар з 10 прихованими вузлами, $\lambda = 0,55$, $\alpha = 0,36$, $\sigma = 0,3$, кількість тестові дані 83, кількість навчальних даних 677, із медіанним фільтром і нормованим виходом.

Таблиця 5.1.6. Результати на основі автентичних даних. Тест за темою 4 (не входить до навчального комплекту).

	Спокій	Радість	Відраза
Спокій	100%	0,00%	0,00%
Радість	0,00%	93,75%	6,25%
Відраза	—	—	—

Параметри мережі: 1 прихований шар з 10 прихованими вузлами, $\lambda = 0,55$, $\alpha = 0,3$, $\sigma = 0,3$, кількість тестових даних 17, кількість навчальних даних 677, із медіанним фільтром та нормованим виводом.

З таблиці 5.1.3, таблиці 5.1.5 і таблиці 5.1.6 можна зробити висновок, що навіть якщо випробувані не входили до комплекту трейнів, рівень розпізнавання був відносним високий. Причина може полягати в тому, що ми підбирали кадри піків вручну, коли коментуючи відеокліпи, таким чином ігноруючи емоцію переходу; Або легше розпізнати справжні емоції, ніж не природні. У таблиці 5.1.4 радість 33% плутає з нейтральною. Ми помітили, що ця тема була дуже веселою

весь час в експерименті, навіть нейтральні емоції виглядали трохи радісно, але не можливо точно сказати, чи є це причиною значно нижчого коефіцієнту розпізнавання. Якщо ми можемо перевірити більше людей, включіть дані про емоції переходу та отримайте більше емоцій, стає зрозуміліше, як нейронна мережа працюватиме з автентичними даними.

5.2 Результати тесту на наборі даних Кона-Канада — Емпіричний метод

5.2.1 One_hidden_layer нейронні мережі для тестування емпіричних методів

Наступна група є результатом нейронних мереж one_hidden_layer з 40 прихованих вузлів, виключити Спокій, не приймати пік , без числа імпульсів і комбінувати з використанням медіанного фільтра та сортуванням даних трейна .

Таблиця 5.2.1.1.

Емоція	Радість	Здивування	Гнів	Відраза	Страх	Сум
Радість	77.87	0.00	10.66	1.64	9.84	0.00
Здивування	0.00	92.93	0.00	0.00	0.00	7.07
Гнів	0.00	0.00	70.75	27.89	0.00	1.36
Відраза	4.55	0.00	3.41	67.62	1.14	9.09
Страх	25.67	0.00	4.76	0.95	67.62	0.00
Сум	12.70	3.97	8.73	8.73	9.52	55.35
Середній	74.55					

Результат отримано нейронною мережею з 1 прихованим шаром, 40

прихованими вузлами, немає числа імпульсу, $\sigma = 0,91$, $\alpha = 0,81$ і проігнорував нейтральну емоцію, відсортнував дані трейна набір із середнім фільтром розміром

Тестовий набір мав 687 векторів і кількість вихідних навчальний набір становив 2300, після сортування даних трейна кількість даних про трейни зменшилася до 1445.

Таблиця 5.2.1.2.

Емоція	Радість	Здивування	Гнів	Відраза	Страх	Сум
Радість	77.05	0.00	10.66	2.46	9.84	0.00
Здивування	0.00	90.91	0.00	0.00	2.02	7.07
Гнів	0.68	0.00	70.75	27.21	0.00	1.36
Відраза	5.68	0.00	5.68	78.41	1.14	9.09
Страх	25.67	0.00	3.81	0.95	67.62	0.95
Сум	12.70	4.76	8.73	8.73	9.52	55.56
Середній	73.38					

Результат отримано нейронною мережею з 1 прихованим шаром, 40 прихованими вузлами, немає числа імпульсу, $\sigma = 0,91$, $\alpha = 0,81$, якщо ігнорувати нейтральну емоцію та відсортнувати набір даних трейна без медіанного фільтра .

У тесті було 687 векторів, а набір трейнів – 2300, після сортування трейна, набір векторів трейнів зменшено до 1445.

Єдина відмінність між таблицею 5.2.1.1.2 і таблицею 5.2.1.1.1 полягає в тому, що пізніший один використовував медіанний фільтр, що спричинило середній показник на 1,17% вище, ніж попередній. Порівняння між таблицею 5.2.1.3.1 і таблицею 5.2.1.3.2 також показало, що середній фільтр покращить показник на 0,5% -2%.

Таблиця 5.2.1.3

Емоція	Радість	Здивування	Гнів	Відраза	Страх	Сум
Радість	72.95	0.00	10.66	4.92	11.48	0.00
Здивування	0.00	92.93	0.00	0.00	0.00	7.07
Гнів	0.00	0.00	65.67	31.97	0.00	1.36
Відраза	4.55	0.00	2.27	84.09	1.14	7.95
Страх	25.67	0.00	4.76	0.95	67.62	2.86
Сум	9.52	8.73	8.73	11.90	5.35	54.76
Середній	73.17					

Результат отримано нейронною мережею з 1 прихованим шаром, 40 прихованими вузлами, немає числа імпульсу, $\sigma = 0,81$, $\alpha = 0,71$ при ігноруванні нейтральної емоції, сортування трейна набір даних за допомогою медіанного фільтра з розміром 3.

Кількість тестових векторів становила 687, кількість вихідного навчального набору 2500, після сортування, кількість даних набору трейнів становило 1475. Різниця між таблицею 4.2.1.3 і таблицею 4.2.1.1 полягає в тому, що наступна мав 2300 векторів трейну, тоді як попередній мав 2500, швидкість знизилася під час нейронної мережі мав ще 200 векторів.

Це можна пояснити надмірним тренуванням. Слід зауважити що для різних нейронних мереж у цій роботі кількість набору трейнів, що викликає над тренуванням може відрізнятися в залежності від різних комбінацій ключових параметрів.

Таблиця 5.2.1.4

Емоція	Радість	Здивування	Гнів	Відраза	Страх	Сум

Радість	77.05	0.00	2.46	4.10	15.57	0.82
Здивування	0.00	79.80	0.00	0.00	9.09	11.11
Гнів	5.44	0.00	53.06	38.78	0.00	2.72
Відраза	5.68	0.00	3.41	72.73	4.55	13.64
Страх	37.14	0.00	0.00	3.81	55.19	2.86
Сум	8.73	4.73	3.97	5.35	7.94	68.25

Результат отримано нейронною мережею з 1 прихованим шаром, 40 прихованими вузлами, $\sigma = 0,91$, $\alpha = 0,71$, якщо ігнорувати нейтральну емоцію та не сортувати набір даних трейна. Набір трейнів мав 2300 векторів, тестовий набір – 687. Середнього фільтра не було. Різниця між 4.2.1.2 і таблицею 4.2.1.4 полягає в тому, що останній не використовує сортування трейнів, що призвело до зниження показника на 5,53%.

Наступна група — результати тестування нейронних мереж `one_hidden_layer` з 40 прихованими вузлами, включи Спокій, без числа імпульсу, комбінується із сортуванням даних трейна.

Таблиця 5.2.1.5

Емоція	Радість	Здивування	Гнів	Відраза	Страх	Сум
Радість	77.69	0.21	2.48	2.27	9.92	2.27
Здивування	0.00	92.93	0.00	0.00	0.00	7.07
Гнів	0.00	0.00	65.67	31.97	0.00	1.36
Відраза	4.55	0.00	2.27	84.09	1.14	7.95
Страх	25.67	0.00	4.76	0.95	67.62	2.86

Сум	9. 52	8. 73	8. 73	11. 90	5. 35	54. 76
Середній	72.45					

Результат отримано нейронною мережею з 1 прихованим шаром, 40 прихованими вузлами, $\sigma = 0,91$, $\alpha = 0,7$ і мав середній фільтр розміром 3 .

Набір даних трейна включав нейтральну емоцію і був відсортований, число якого було 1525. Кількість тестових векторів 1171. До сортування та відбору набір трейнів мав 3200 векторів.

Таблиця 5.2.1.6

	Нейтральна	Радість	Здивування	Гнів	Відраза	Страх	Сум
Спокій	72. 52	0. 00	2. 89	2. 89	13. 43	2. 69	5. 58
Радість	1. 64	55. 74	0. 00	8. 20	3. 28	27. 87	3. 28
Здивування	0. 00	0. 00	90. 91	0. 00	0. 00	1. 01	8. 08
Гнів	5. 44	0. 00	0. 00	51. 70	35. 05	2. 04	4. 76
Відраза	3. 41	3. 41	0. 00	14. 77	57. 95	2. 27	18. 18
Страх	11. 43	12. 38	1. 90	0. 95	12. 38	57. 14	3. 81
Сум	25. 40	0. 79	15. 67	5. 35	12. 70	2. 38	35. 71
Середній	64. 52						

Результат отримано нейронною мережею з 1 прихованим шаром, 40 прихованими вузлами, середній розмір фільтра = 3 , $\sigma = 0,91$, $\alpha = 0,61$. Тренувальний набір мав 3200 векторів, не відсортованих, не виділених і включав нейтральну емоцію. Тестовий набір мав 1171 вектор.

Порівнюючи таблицю 5.2.1.2.1 і таблицю 5.2.1.2.2, усі інші параметри майже те ж саме за винятком попереднього відсортував і відібрав комплект трейнів, який порівнюється до 7,93% вище, ніж у останньому. У цьому випадку важливу роль

відіграє сортування ролі для підвищення ставки.

Наступна група є результатом one_hidden_layer нейронних мереж з 40 прихованих вузлів, включають нейтральні, візьміть пікові кадри, виберіть розмір кадру 5, який означає, що ми зробили 5 кадрів на піку. Не було числа імпульсу. Це поєднано із використанням медіанного фільтра .

Таблиця 5.2.1.7

	Нейтральна	Радість	Здивування	Гнів	Відраз	Страх	Сум
Спокій	72.52	0.00	2.89	2.89	13.43	2.69	5.58
Радість	1.64	55.74	0.00	8.20	3.28	27.87	3.28
Здивування	0.00	0.00	90.91	0.00	0.00	1.01	8.08
Гнів	5.44	0.00	0.00	51.70	35.05	2.04	4.76
Відраза	3.41	3.41	0.00	14.77	57.95	2.27	18.18
Страх	11.43	12.38	1.90	0.95	12.38	57.14	3.81
Сум	25.40	0.79	15.67	5.35	12.70	2.38	35.71
Середній	60.24						

Результат отримано нейронною мережею без медіанного фільтра , 1 прихований шар, 40 прихованих вузлів, $\sigma = 0,81$, $\alpha = 0,61$. Тестовий набір мав 1171 вектор. У комплекті трейнів було 2300 включають нейтральні емоції, після досягнення піку кадрів кількість набору трейнів зменшилася до 1026 , потім після сортування вона зменшилась до 490.

Таблиця 5.2.1.8

	Нейтральна	Радість	Здивування	Гнів	Відраза	Страх	Сум
Спокій	74.38	0.00	2.69	2.48	13.22	2.27	4.96
Радість	1.64	58.20	0.00	8.20	1.64	27.05	3.28
Здивування	0.00	0.00	91.92	0.00	0.00	1.01	8.08
Гнів	5.12	0.00	0.00	53.74	35.37	0.68	4.08
Відраза	3.41	1.14	0.00	13.64	61.36	2.27	18.18
Страх	11.43	12.38	2.86	0.95	11.43	58.10	2.86
Сум	27.78	0.79	15.67	5.35	12.70	2.38	33.33
Середній	61.58						

Результат отримано нейронною мережею з 1 прихованим шаром, 40 прихованими вузлами, $\sigma = 0,81$, $\alpha = 0,61$, із середнім фільтром розміром 3. Кількість тестових векторів становила 1171. Набір трейнів мав 2300 векторів, включаючи нейтральні емоції, після отримання пікових кадрів кількість зменшилась до 1026, після сортування та вибору кількість кадрів дані трейнів були 490.

Порівнюючи таблицю 5.2.3.1 і таблицю 5.2.3.2, легко побачити, що всі інші параметри нейронної мережі аналогічні, за винятком використання медіанного фільтра в таблиці 5.2.1.9. Це оскільки той, що використовує середній фільтр, має на 1,34% вищий показник.

Наступна група є результатом `one_hidden_layer` нейронних мереж з 40 прихованих вузлів, виключити Спокій, візьміть кадри піків, без числа імпульсу, пік розміру = 5, поєднувати з використанням медіанного фільтра та сортуванням даних.

Таблиця 5.2.1.9

Train set	After take peak frame	After sort	Test set size	Rate (filter size 3)	Rate no filter
3200	No peak frame	1962	687	67.4%	67.3%
3200	1448	972	687	63.9%	63.2%
2300	1030	No sorting	687	63.9%	65.1
2300	No peak frame	1446	687	75.6%	73.4
3200	1448	No sorting	687	66.6%	64.1%

Результати тесту з 5 різних нейронних мереж у базі даних Cohn - Kanade. З цієї таблиці ми бачимо, що прийом пікових кадрів завжди призводить до нижчої частоти. І порівнюючи перший рядок і четвертий рядок, ми також бачимо, що швидкість зменшилася з 73,4% до 67,3%, тоді як кількість трейнів зросла з 2300 до 3200. Це можна пояснити як перетренування. Порівнюючи стовпці 5 і 6, ми можемо зробити висновок про медіану фільтр покращив швидкість у всіх випадках, крім випадку третього ряду.

Як правило, група таблиць у розділі 5.2.1.1 є результатом виключення Спокій, не бере пікових кадрів даних. Група таблиць у розділі 5.2.1.2 є результатом при включенні нейтрального, не взяття пікових кадрів даних. Група столів у розділі 4.2.1.3 є результат при включенні нейтралі, беремо пікові кадри даних. Група таблиць у розділі 5.2.1.4 є результатом без урахування нейтрального та пікового значення кадри даних.

У кожній групі ми також перевірили різні комбінації з використанням середнього фільтру, сортування набору трейнів. З різних порівнянь можна зробити висновок, що взяття пікових кадрів призведе до зниження швидкості, тоді як медіанний фільтр, сортування набору трейнів та виключення нейтральної волі покращить швидкість. З таблиці 5.2.1.4.1 ми можемо далі виявити, що сортування і взяття пікових кадрів може суттєво вплинути на результати, зазвичай внесок був більше 2%. З іншого боку, з порівняння між групою 5.2.1 (не взявши пік, без нейтрального) і групу 5.2.2 (не взявши пік, але з нейтральним), ми можемо зробити висновок, що виключення нейтральних емоцій, не має великої різниці. Ми можемо легко знайти те, що потрібно порівняти, вибравши різні пари таблиць,

таким чином знайдіть цікавий доказ наших дизайнерських ідей у розділі 5.

5.2.2 Нейронні мережі для спеціальних категорій емоцій

З попередніх результатів ми робимо висновок, що емоція здивування завжди має вищий рівень розпізнавання, ніж інші, можна запитати, чи правда ця емоція домінує над вагою нейронних мереж, таким чином обмежує розпізнавання інших емоцій ставка? Як щодо того, щоб спробувати різні категорії класифікації, щоб краще зрозуміти зв'язок між даними?

Ми намагалися ігнорувати всі дані про несподівані емоції, щоб отримати найвищий рівень розпізнавання серед іншого, також ігноруючи нейтральні дані про емоції через їх переважну кількість даних. Якщо рівень радості, який зазвичай має другий найвищий показник, може значно покращитися вище, то ми можемо зробити це рекурсивно: одна нейронна мережа для однієї емоції, ідеї і кроки:

- a) Спочатку визначте, чи є це здивуванням.
- b) Якщо є, то спробуйте другу нейронну мережу, щоб класифікувати, якщо це радість.
- c) Якщо є, то зупиніться і повторіть попередні кроки поки не буде збігу.

Якщо коефіцієнт розпізнавання всіх емоцій вище 90% так само ми здивуємося, тоді середній показник може бути більш задовільним, що означає припущення, що емоція з більш високою швидкістю стримує швидкість інших, є правильним. Після ігнорування здивування та нейтральності швидкість радості зростає, але не надто сильно. В у перевічених нами випадках показник радості покращився з приблизно 75% до 81,4%--84%, тому ми зупинилися на цьому методі. По-друге, ми спробували різні категорії емоцій, щоб зробити більше про їхнє співвідношення між різними категоріями емоцій. З тих таблиць, ми не знайшли більшого рівня класифікації, скинувши категорію цих емоцій, це означає, що якщо ми розробимо 7 нейронних мереж для кожної емоції, це не допоможе для підвищення точності доказування, що можна знайти в таблиці 5.2.2.2 до таблиці 5.2.2.5.

Таблиця 5.2.2.1

	Спокій	Позитивний	Здивування	Негативний
Спокій	79,05%	3,32%	0,83%	16,80%
Позитивний	13,11%	76,23%	0,00%	10,66%
Здивування	4,04%	0,00%	92,93%	3,03%
Негативний	11,16%	10,73%	1,29%	76,82%

Результат на наборі даних Кона-Канада. Параметр мережі: 1 прихований шар, 10 прихованих вузлів, $\lambda=0,5$, $\alpha=0,5$, $\sigma=0,3$. Номер тестового вектора був 687, номер вектора трейна був 3756, використовуйте медіанний фільтр і нормалізуйте вихід.

Таблиця 5.2.2.1 побудована, щоб побачити, наскільки може нейронна мережа розрізняти ці 4 категорії, де позитивні емоції означають радість, негативні емоції включають огиду, гнів і страх. Середній показник вищий за класифікаційний, їх поділяють на 7 категорій.

Таблиця 5.2.2.2

	Радість	Не радість
Радість	73. 77%	25. 33%
Не радість	4. 39%	95. 61%

Результат на наборі даних Кона-Канада.

Параметр мережі: 1 шар, 10 прихованих вузлів, $\lambda=0,55$, $\alpha=0,6$, $\sigma=0,3$. Перевірте вектор номер 687, трейн вектор номер 3756, використовуйте медіанний фільтр і нормалізуйте вихід.

Таблиця 5.2.2.3

	Здивування	Не дивно
Здивування	87. 82%	12. 12%

Не дивно	0,93%	99. 07%
----------	-------	---------

Результат по базі даних Кона-Канада. Параметр мережі: 1 прихований шар, 10 прихованих вузлів $\lambda= 0,55$, $\alpha=0,6$, $\sigma=0,3$. Перевірте вектор номер 687, трейн вектор номер 3756, використовуйте медіанний фільтр і нормалізуйте вихід.

Результат по базі даних Кона-Канада. Параметр мережі: 1 прихований шар, 10 прихованих вузлів, $\lambda= 0,55$, $\alpha=0,6$, $\sigma=0,3$. Перевірте вектор номер 687, трейн вектор номер 3756, використовуйте медіанний фільтр і нормалізуйте вихід.

Таблиця 5.2.2.4

	Відраза	Не відраза
Відраза	55. 68%	44. 32%
Не відраза	5. 08%	94. 92%

Параметр мережі: 1 прихований шар, 10 прихованих вузлів, $\lambda= 0,55$, $\alpha=0,6$, $\sigma=0,3$. Перевірте вектор номер 687, трейн вектор номер 3756, використовуйте медіанний фільтр і нормалізуйте вихід.

Таблиця 5.2.2.5

	Сум	Не Сум
Сум	32. 71%	65. 29%
Не Сум	2. 39%	97. 61%

Результат по базі даних Кона-Канада. Параметр мережі: 1 прихований шар, 10 прихованих вузлів, $\lambda= 0,55$, $\alpha=0,6$, $\sigma=0,3$. Перевірте вектор номер 687, трейн вектор номер 3756, використовуйте медіанний фільтр і нормалізуйте вихід.

5.3 Результати тестування бази даних Кон-Канаде

У наведених нижче таблицях представлені результати тесту на різних нейронних мережах з різними вузлами, шарами. Продуктивність мереж залежить від їх розміру. З метою порівняння ми намагаємося уникнути дуже великої кількості прихованих вузлів, які викликають меншу генерацію мережі. Одна мережа прихованого рівня з різними прихованими вузлами.

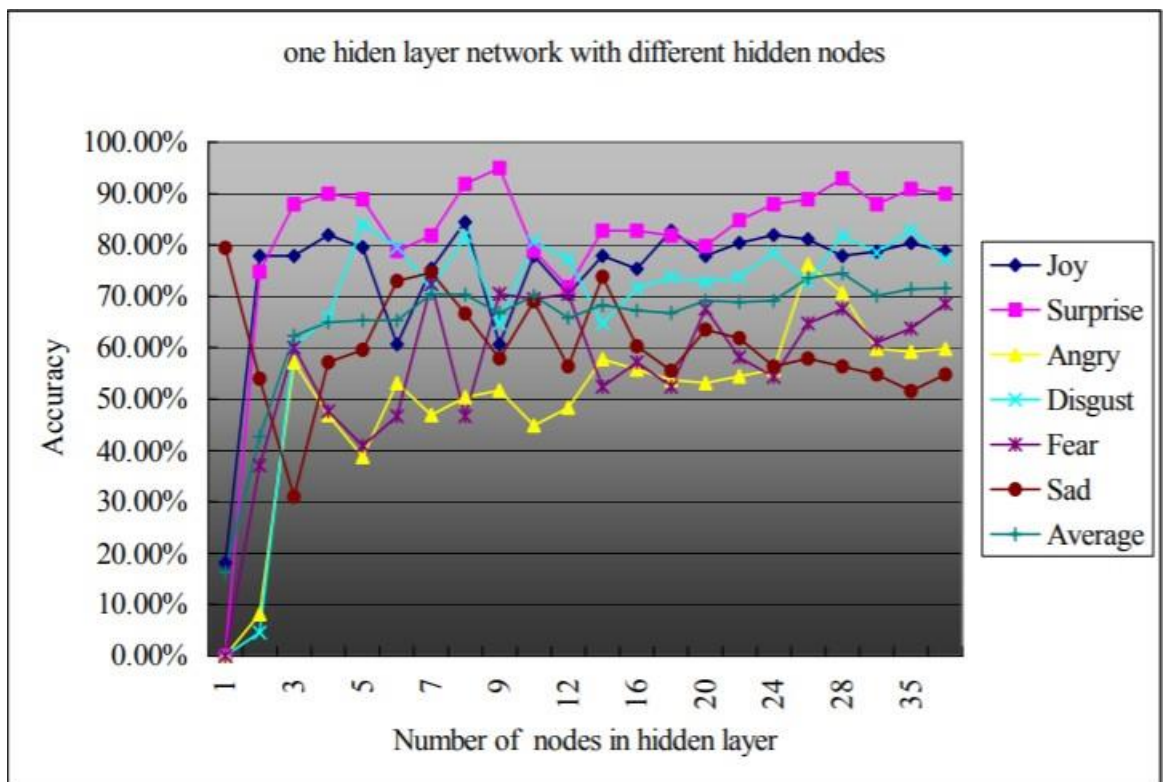


Рисунок 5.3.1 - Тестування в мережі one_hidden_layer з різними вузлами в прихованому шарі.

На цьому рисунку кількість тестових векторів становила 687; кількість векторів трейнів була 3755. Він використовував медіанний фільтр і нормований вихід, але без числа імпульсу. Кожна швидкість отримана за значеннями α і σ . Значення α і σ на цьому рисунку становило не записується кодом. На осі X інтервал прихованих вузлів перед 10 прихованими вузлами — 1, між 10 і 20 — 2 вузли на сітку на осі X і 5 вузлів після 20 прихованих вузлів.

З малюнка ми можемо зробити висновок, що середній показник класифікації

коливається в той час як кількість змін вузла. Це доводить, що кількість вузлів є ще одним важливим фактором нейронної мережі. Також можна помітити, що нейронна мережа з 40 вузлами зараз отримує найкращу точність на цьому рисунку, але це не завжди вірно змінюючи значення α і σ .

Одну цікаву річ, яку ми помітили, це хоча деякі документи повідомляли 2_hidden_layer нейронні мережі зазвичай краще, ніж 1_hidden_layer нейронні мережі [14], при виборі різної швидкості навчання, числа імпульсів, функції активації та його параметра, під час наших експериментів важко сказати, який з них перевершує інший.

Порівнюючи зі ставкою, отриманою з 1_hidden_layer, нам не вдалося повідомити про кращу швидкість нейронних мереж 2_hidden_layer. Причина в тому, що автоматичний вибір параметрів нейронної мережі працював краще в 1_hidden_layer, ніж у 2_hidden_layer, оскільки є ще один параметр – номер вузлів другого прихованого шару, що викликало більше комбінації параметрів складні, тому тренування більшої кількості нейронних мереж займає більше часу. Найімовірніше, що ми можемо знайти кращу модель, спробувавши більше комбінацій 2_hidden_layer.

З таблиці 5.3.1 можна зробити одне припущення, що дані є невідповідність між набором тренів і тестовим набором у базі даних Cohn_Kanade, оскільки трен швидкість може бути досягнута більш ніж на 98%, але показник тесту становить лише більше 60%. Це можна пояснити як переналаштування нейронних мереж, але справжня причина перепідгонки це невідповідність даних.

Чим краща узгодженість між треїном і тестовим набором, тим краще працює зворотне поширення. Якщо є ефективний спосіб вирішення невідповідності, то швидкість розпізнавання емоцій можна покращити набагато більше.

Інший висновок, який ми можемо зробити з таблиці 5.3.1, полягає в тому, що наш 2_hidden_layer мережі працюють належним чином, оскільки вони можуть досягати майже 100% швидкості тренів, тобто похибки зменшуються майже до нуля за рахунок правильного поширення їх диференціала.

Таблиця 5.3.1

Nodes (hidden1Xhidden2)	Train rate	Test rate	Train vector number	Test vector number	σ range	σ increase Step	α range	α increase Step	λ range	λ increase step
29x28	94.6	70.2	3756	687	2-2.5	0.5	0.2-1	0.3	0.5-1	0.3
29x29	96.9	68.5	3756	687	2-2.5	0.5	0.2-1	0.3	0.5-1	0.3
29x29	97.7	66.3	3756	687	1.8-2.5	0.5	0.2-1	0.3	0.5-1	0.3
4 x 4	81.83	67.32	3756	687	0.9-2	0.5	0.2-1	0.3	0.5-1	0.3
4 x 5	82.77	69.7	3756	687	0.9-2	0.5	0.2-1	0.3	0.5-1	0.3
29x29	98.7	66.5	3756	687	0.9-2	0.3	0.2-1	0.3	0.5-1	0.3

Результат тесту на 2_hidden_layer: різні вузли та різні категорії параметрів активації. Параметри нейронної мережі: тестовий вектор номер 687, вектор трейну номер 3755. Використовуйте середній фільтр, але не нормалізуйте вихід, кожна швидкість має різний набір λ , α та σ .

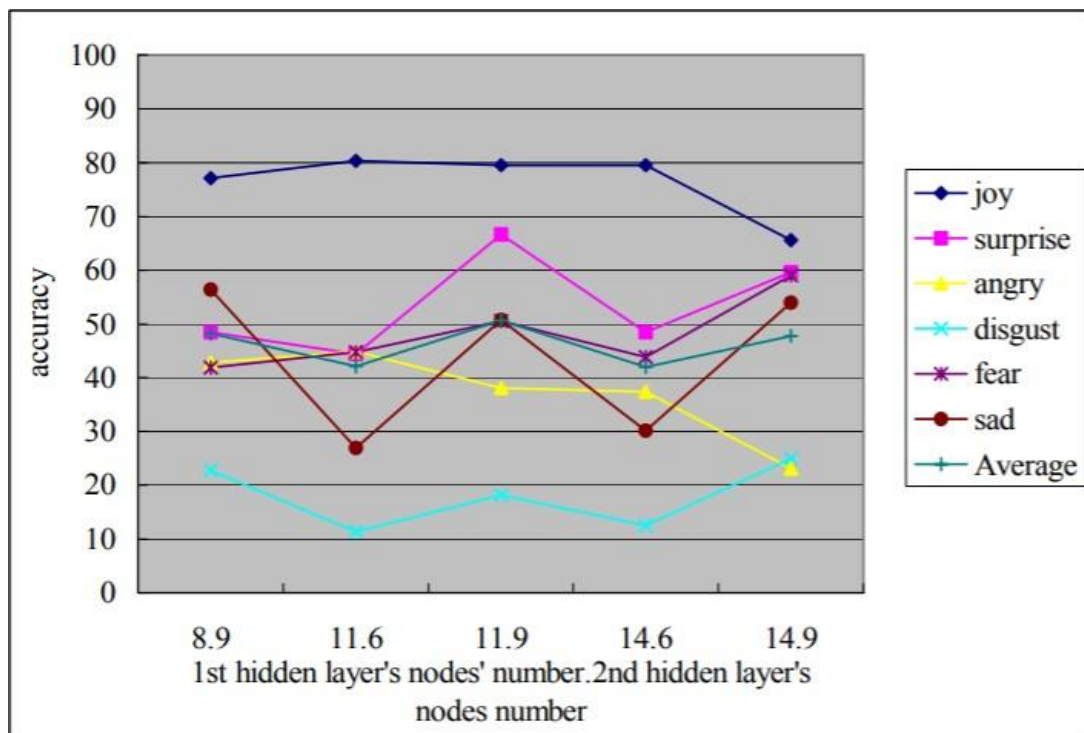


Рисунок 5.3.2

П'ять нейронних мереж 2_hidden_layer з неправильним порогом ваги [0,5, -0,5] не тільки для ініціалізації, але й утримання ваги до цього порогу протягом усього навчання процедури. Категорія параметра функції активації [0.1-1] не є

належною для цього випадку, тому викликаний нейронна насичення, що призвело до низького рівня розпізнавання. Оскільки це не вдалий результат, детальне поєднання параметри не зберігалися.

Таблиця 5.3.2

	Радість	Здивування	Сердитий	Відраза	Страх	Сум
Радість	63.11	5.74	7.38	4.10	19.67	0,00
Здивування	0,00	96,97	0,00	0,00	2.02	1.01
Гнів	2.04	0,00	79,59	15.33	0,00	2.04
Відраза	3.41	5.68	21.59	65,91	2.27	1.14
Страх	25,71	0,00	0,95	0,95	71,43	0,95
Сум	0,00	3.17	3,97	8.73	0,00	84.13
Середній	76,86					

Найкраща модель: 1 прихований шар, 4 приховані вузли, $\alpha=0,3$, $\lambda=0,6$, $\sigma=0,9$, тест на Кон-Канаде Номер вузлів 1-го прихованого шару. 2-го прихованого шару номер вузлів точність радість Здивування Гнів відраза страх сумний Середній бази даних. Кількість даних про трейни – 2500, після сортування кількість даних про трейни зменшилась до 1475.

Кількість тестової бази даних становить 687, із медіанним фільтром (розмір = 3), без пікових кадрів трейна та дані тесту. Ось найкращий набір параметрів у таблиці 5.3.2 з найвищим показником серед інших нейронних мережах в цій роботі, також з невеликою швидкістю навчання, що означає узагальнення може бути добре, якщо змінити тестову базу даних, число імпульсу нормальний у порівнянні зі значенням 0,5, про яке повідомляють інші. Таким чином довели, що добре

Набір параметрів дуже важливий при проектуванні нейронних мереж. Ми також спробували змінити параметр функції активації від 0,9 до 1, що є

нормальним значенням для інших. В результаті, цей показник зменшиться приблизно на 2%, це підтвердило наш спосіб вибору комбінацій параметрів працює краще, ніж просто емпіричний вибір деяких комбінацій.

5.4 Результати тесту на наборі даних Кона-Канада (різні початкові ваги)

Як згадувалося в розділі 5.2, можуть спричинити різні вихідні точки ваг різні результати, але наскільки ініціалізація ваг вплине на нейронну мережу? Коливання кривої ставки в наступних таблицях є частиною відповідей, де значення n «часу ініціалізації» означає, що ми скидаємо n -й раз вихідні точки.

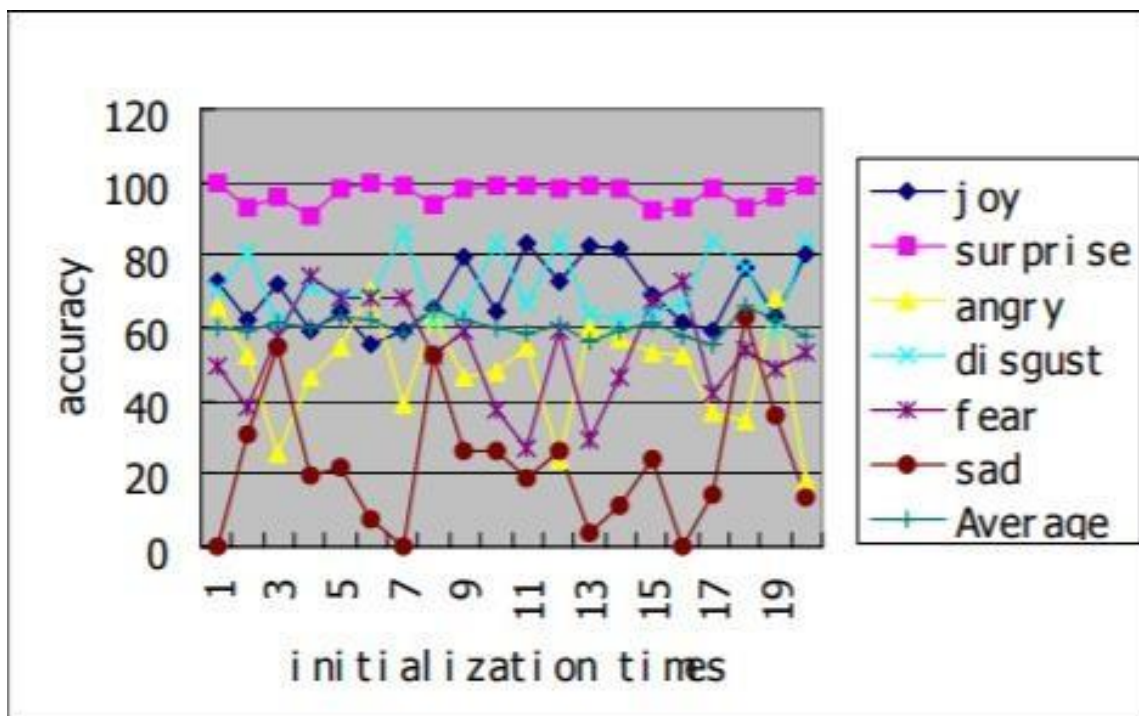


Рисунок 5.4.1 прихований шар з 10 прихованими вузлами, $\alpha=0,4$, $\lambda=\sigma=0,6$, ініціалізуйте ваги 20 разів.

Нормалізувати вихід і використовувати медіанний фільтр. Ми також бачимо, що крива курсу на рисунку 5.4.1 коливається більше, ніж на рисунку 5.4.2, де нормалізація вихідних даних не використовується в наступному. Але іноді ставка в таблиці 5.4.1 трохи вище, що означає, що легше знайти вищу ставку нормалізація.

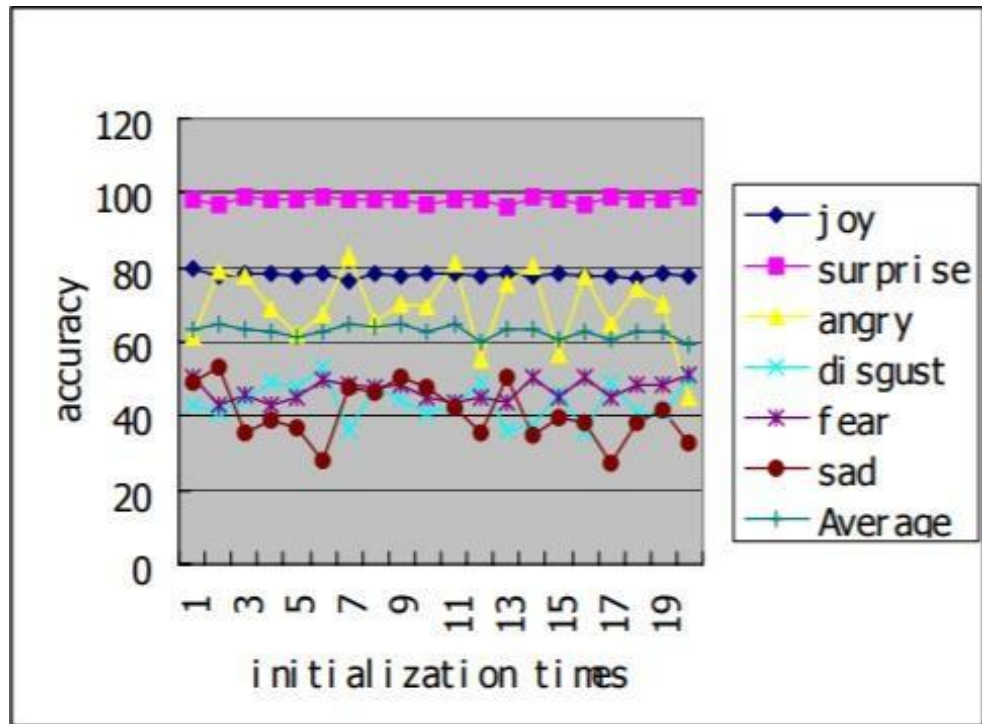


Рисунок 5.4.2 1 прихований шар з 10 прихованими вузлами, $\alpha=0,4$, $\lambda=\sigma=0,6$, ініціалізуйте ваги 20 разів. Нормалізувати вихід і використовувати медіанний фільтр.

ВИСНОВОК

Реалізація інтелектуальної системи, здатної розпізнавати емоції людини – це рішення технічно та математично складного завдання, що вимагає поглибленого вивчення предметної галузі, власного аналізу, вміння та навичок при обробці та виборі експериментальних даних, глибокі знання в галузі дискретної математики, геометрії, програмування, а також психології та багатьох інших галузях науки. Використання нейронних мереж при розробці систем розпізнавання, дозволяє структурувати погано формалізовані дані, прискорити процес обробки даних, оцінити адекватність середовища, в яке вміщено інформаційну модель нейронної мережі.

Метрична система дозволяє створити прототипи емоцій, завдяки яким у розробника з'являється можливість «взаємодії» з нейронною мережею, тобто. навчання з урахуванням виявлених прототипів емоцій. Основна мета магістерської роботи полягала у виборі інформаційної моделі нейронної мережі, а також опис алгоритму розпізнавання на вирішення практичного завдання. Було обрано інформаційну модель на основі багат шарового перцептрон.

Ця модель оптимальна за внутрішньою структурою та способом управління інформаційними потоками між нейронами Крім того, така модель здатна мінімізувати кількість вхідних елементів. Модель на основі багат шарового перцептрон є універсальною моделлю та підходить для вирішення завдань різного рівня складності, в тому числі і для розв'язання задачі розпізнавання емоцій. Основні висновки та результати дослідження:

1) На підставі вивченого матеріалу зі штучних інтелектуальним системам було виявлено найефективніші системи та точні методи розпізнавання, відмічені успішні приклади реалізації. Система, яка використовувалася у цій роботі була самонавчання.

2) Огляд основних видів інформаційних моделей нейронних мереж дозволив виділити дві універсальні моделі, застосовні для широкого кола завдань: радіальні нейронні мережі прямого поширення та односпрямовані

багатошарові мережі. Вибір інформаційної моделі для вирішення практичного завдання було зроблено на користь односпрямованих багатошарових мереж (модель багатошарового перцептрон). Ця модель підходить за внутрішньою структурою та математичним описом та передбачає різні підходи до розв'язання задачі.

3) Існуючі методи розпізнавання емоцій у магістерській роботі було розглянуто теоретично. Серед вивчених методів розпізнавання було виділено метод із використанням штучних нейронних мереж.

4) У практичній частині роботи був розглянутий математичний апарат реалізації алгоритму розпізнавання: вибір інваріантів особи, адекватної метрики, розпізнавання образів нейронною мережею. Також була представлена методика розпізнавання рухової активності особи та вирішено завдання розпізнавання емоцій у програмних продуктах аналізу та обробки даних на мові програмування Пайтон. Всі поставлені завдання були виконані в повному обсязі, отже, головної мети було досягнуто.

Ми спробували чимало способів, але все ще є подальші вдосконалення цієї роботи, тобто додати кращий інтерфейс для ручного керування параметрами під час навчання. Існує багато різних факторів, які можуть вплинути на роботу нейронамережі. Навіть для тих факторів, які ми згадували в цій дипломній роботі, ми не вичерпали всіхтакі комбінації, як-от:

- 1) Середній фільтр може розглядати більше сусідніх кадрів.
- 2) Створення різних нейронних мереж для особливостей верхньої та нижньоїграней обличчя відповідно та поєднують їхні результати.
- 3) Використовуйте імітований відпал замість набору напрямків Пауелла, щоб мінімізувати функцію помилки.
- 4) Краще справляйтеся з надмірним тренуванням.
- 5) Змінити функцію активації з сигмовидної на інші.

Щодо автентичного набору даних, то якщо емоцій 7, то порівняти зрозуміліше з результатами, які ми отримуємо з набору даних Кон-Канаде. Хоча деякі з природних емоції дуже важко отримати, тобто розлютити людей, не

сказавши їм, що це так лише експеримент.

Нині є багато інших людей, які виконують подібну роботу, використовуючи назад розмноження. Рівень розпізнавання їхніх систем становить у середньому від 67% до 100% [4,14]. Але несправедливо порівнювати результати між різними системами на різних бази даних і різні обмеження в експериментах під час отримання бази даних. У статті, яку ми досліджували, жодна з цих нейронних мереж не класифікувала емоції Набір даних Кон-Канаде або в автентичній базі даних.

Ще одна цікава річ: ми не знайшли жодного звіту про використання нейронної системи мережі, що аналізують дані Кона-Канада цілісними способами замість аналітичних, які ми тут використали. Хоча кажуть, що цілісні способи перевершують пізніші [30,31,32]. Буде цікаво, якщо ми використаємо інформацію про ціле обличчя, а не про аналітичні риси обличчя як вхід нейронної мережі. Більшість систем автоматичного аналізу емоцій класифікує людські емоції на 7 прототипних емоцій [13]. Одним з основних недоліків цього методу є те, як поводитися змішані емоції або тонкі зміни обличчя, тобто здивувати страхом або здивувати радість. Доступним рішенням може бути використання більшої кількості категорій - більше 7 вихідних вузлів в нейронних мережах. Оскільки ми не намагалися таким чином, ми не знаємо, як налаштувати категорії. Для цього потрібні результати психологічних досліджень.

Останнім часом з'явилися повідомлення, що фокусуються на використанні моделей компонентів обличчя для моделювання і відстежувати дрібнозернисті зміни виразу обличчя [13]. Коефіцієнт розпізнавання для одиниця дії на обличчя (AU) становить 93% - 96% за набором даних Кон-Канаде [13]. Але є більше 7000 комбінацій Aus[14], таким чином, навіть якщо AU є правильними визнали, як дозволити комп'ютеру зробити висновок із цих 7000 комбінацій, які люди Це буде крок до того, як комп'ютер нарешті почне спілкуватися з людиною за допомогою обличчя вирази. Однак точне розпізнавання АС є дуже важливим прогресом. Класифікація емоцій за комбінаціями АС замість зображення чи функції безпосередньо можуть призвести до більш точного розпізнавання емоцій у

майбутньому. Іншим напрямком майбутнього є використання помітних точок [33] для забезпечення точності відстеження. Разом з іншими методами виявлення серцебиття, голосу, жестів тіла визнання, зміст спілкування, буде день для людей, щоб насолоджуватися взаємодія з повністю гуманістичним комп'ютером.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Головка В.А. Нейронні мережі: навчання, організація та застосування / М.: ІПРЖР. 2008
2. Градосільська Г.В. Мережеві виміри у соціології: Навчальний посібник /М.: Видавничий дім "Новий підручник", 2004. 248 с.
3. Гундирєв К.В. Штучні нейронні мережі у задачах діагностування рейкових ланцюгів// Науч.-исл.лабор. "Комп'ютерні системи автоматики". 2005.
4. Єфремова Є.А., Дунаєв, Є.В. Застосування нейронних мереж для прогнозування фінансових часових рядів.// Доповіді ТУСУР. 2004.
5. Заєнцев І.В. Нейронні мережі: основні моделі / Навчальний посібник фіз. ф-та. Вороніж. 1999.
5. Іванов А.І. Нейромережні алгоритми біометричної ідентифікації / Вид- у: Радіотехніка. 2005. 144 с.
7. Каллан Р. Основні концепції нейронних мереж/ Вид-во: Вільямс. 2002. 287 с.
8. Кальченко Д.А. Нейронні сіті на порозі майбутнього. / Комп'ютПрес, 2005.
9. Каширіна І. Л. Нейромережні технології: Навчально-методичний посібник для вузів / Вид-во ВДУ. 2008. 72 с.
10. Князєв Б.А., Гапанюк Ю.Є. Розпізнавання аномальної поведінки людини/ Інженерний журнал.2013 512 с.
11. Козін Н. Є., Фурсов В. А. Поетапне навчання радіальних нейронних мереж / РФФІ. амер. фонд гражд. ісл. 2003.
12. Кочеткова А.С. Застосування нейронних мереж для моніторингу безпеки // Серія 9: Дослідження молодих вчених. 2007.
13. Крісілов В.А. Подання вихідних даних у нейросетевих задачах

програмування / Одеса: ОНПУ. 2003.

14. Коло П.Г. Нейронні мережі та нейрокомп'ютери: Навчальний посібник з курсу "Мікропроцесори" / Видавництво МЕІ, 2009.176 с.

15. Круглов В.В. Нечітка логіка та штучні нейронні мережі / ФІЗМАЛІТ. 2001.

15. Леван Д.М., Феоктистів Н.А. Особливості використання багат шарового перцептронну / Наукознавства. вип. 2. 2014 року.

17. Лоскутов А.Ю. Аналіз часових рядів. Курс лекцій/МДУ. 2013 року.

18. Луценко А. В. Функція Гріна та її застосування: методичний посібник з курсу «Диференціальні рівняння»/ХНУ імені В.М. Каразіна. 2013. 28с.

19. Медведєв В.С., Потьомкін В.Г. Нейронні мережі в MATLAB / М.: ДІАЛОГ- МІФІ. 2002. 496 с.

20. Местерський Л. М. Математичні методи розпізнавання образів/МДУ, курс лекцій. 2004.

21. Мінський М., Пейперт С. Перцептрони / М.: Світ. 1971.

22. Назаров А.В., Лоскутов А.І. Нейросетельні алгоритми прогнозування та оптимізації систем/СПб: Наука та техніка. 2005. 384 с.

23. Нгуен Д. Комплексне дослідження задачі класифікації із застосуванням нечітких моделей та розподілених обчислень: дис канд.тех. наук. Іркутськ. 2014 року.

24. Осовський С. Нейронні мережі для обробки інформації/М.: Вільямс. 2005.

25. Осовський С. Нейронні мережі для обробки інформації / пров. з польськ. І.Д. Рудинського. 2002. 344 с.

25. Осовський З. Розпізнавання образів Оглядова лекція./М.: Вільямс. 2007.

27. Остроух А.В. Інтелектуальні системи Навчальний посібник/ Науково-Інноваційний центр. 2015 р

28. Рассел С., Норвіг П. Штучний інтелект: сучасний підхід / Вид-во: Вільямс. 2005. 1424 с.
29. Розенблатт Ф. Принципи нейродинаміки. Перцептрони та теорія механізмів мозку/М.: Світ. 1992.
30. Рудаков А.С. Підходи до вирішення задачі прогнозування тимчасових рядів з допомогою нейронних мереж // Аналіз даних та інтелектуальні системи. №4. 2008.
31. Рутковська Д., Пілінський, М., Нейронні мережі, генетичні алгоритми та нечіткі системи/М., 2004. 452 с.
32. Семейкін В. Д., Скупченко А.В. Моделювання штучних нейронів мереж серед MATLAB // Сист. телекомунікацій. 2008.
33. Сойфер В.А. Методи комп'ютерної обробки зображень/М.: Фізматліт. 2001. 784 с.
34. Солдатова О.П., Семенов В. В. Застосування нейронних мереж для вирішення задач прогнозування / СДАУ ім. ак. С.П. Корольова. 2010 року.
35. Тархов Д. А. Нейронні мережі. Моделі та алгоритми / Вид-во: Радіотехніка. 2007. 256 с.
35. Татаренков Д.А. Аналіз методів виявлення осіб на зображенні//Технічні науки. Молодий вчений. 2015. №4 (84).
37. Уоссермен Ф. Нейрокомп'ютерна техніка: теорія та практика / пров. з англ. Ю.А.Зуєв, В.А. Точенов.1992.
38. Хайкін С. Нейронні мережі. Повний курс/М.: Вільямс, 2005.
39. Халафян А.А. Статистичний аналіз даних. 3-тє вид. навч. / Біном - Прес. 2007. 512 с.
40. Царегородцев В. Г. Обчислювальні технології// Вісник КазНУ. год 3. 2008. с. 308–315.
41. Шаміс А.Л. Поведінка, сприйняття, мислення: проблеми створення

штучного інтелекту / Вид-во: Наука. 2005. 224 с.

42. Шахнов В.А., Власов А.І., Кузнєцов А.С. Нейрокомп'ютери: архітектура та схемотехніка / М: Вид-во Машинобудування. 2000. 64 с.

43. Шітіков В.К., Розенберг Г.С., Зінченко Т.Д. Методи системної ідентифікації / Тольятті: ІЕВБ РАН. 2003. 463 с.

45. Аналітичні технології для прогнозування та аналізу даних// Нейропроект. 1999-2005. [Електронний ресурс]/URL: <http://www.neuroproject.ru/neuro.php> (дата звернення 11.02.2016).

45. База та генератор освітніх ресурсів [Електронний ресурс] // МДТУ ім. н.е. Баумана, Кафедра САПР. 2003–2015. URL: <http://bigor.bmstu.ru/?cnt/?doc=NN/base.cou> (дата звернення 11.03.2016).

47. Введення у теорію нейронних мереж [Електронний ресурс]//Основні поняття нейромереж/ URL: <http://www.ogc.ru> (дата звернення 12.04.2017)

48. Дослідження ринку систем розпізнавання емоцій мереж [Електронний ресурс] / URL: <https://habrahabr.ru/post/133686/> (дата звернення 2.02.2017).

49. Області застосування штучних нейронних мереж [Електронний ресурс] / Основні напрямки використання / URL: <http://www.neuropro.ru/> (дата звернення 9.03.2017)

50. Поняття інформації [Електронний ресурс]./URL: <http://cde.osu.ru/> (дата звернення 12.03.2017)

51. Портал штучного інтелекту [Електронний ресурс]// Штучні нейронні мережі / URL: <http://www.aiportal.ru> (дата звернення 1.04.2017)

52. Застосування нейронних мереж у робототехніці [Електронний ресурс]/URL: <http://geektimes.ru/> (дата звернення 3.04.2017)

53. Застосування нейронних мереж для класифікаційних завдань [Електронний ресурс]/URL: <http://www.basegroup.ru/> (дата звернення 5.04.2017)

54. Профайлінг [Електронний ресурс]// класифікації емоцій/

URL:<https://ua.wikipedia.org/wiki/> (дата звернення 15.04.2017)

55. Розробка та використання інтелектуальних систем [Електронний ресурс]/ URL: <http://www.studfiles.ru/preview/5809976/> (дата звернення 14.01.2017)

55. Система кодування лицьових рухів [Електронний ресурс]// основні мімічні вирази/ URL: <http://dic.academic.ru/dic.nsf/ruwiki/1628394> (дата звернення 15.04.17)

57. Емоційна напруженість методика визначення [Електронний ресурс]/URL: <http://pandia.ru/text/80/079/8545.php> (дата звернення 19.04.2017)

58. MATLAB [Електронний ресурс]//запровадження в програму/ URL: <http://matlab.ru/products/MATLAB> (дата звернення 29.04.2017)

59. MATLAB. [Електронний ресурс]// Лабораторний практикум/URL: http://elib.rshu.ru/files_books/pdf/img-503135145.pdf (дата звернення 2.05.2017)

60. MATLAB. [Електронний ресурс] // Інфоурок / URL: <https://infourok.ru> (дата звернення 10.05.2017)

61. Microsoft Excel [Електронний ресурс:] // Редагування таблиць / URL: <https://office.live.com/start/Excel.aspx?omkt=ru-RU> (дата звернення 17.05.2017)

62. MathCad [Електронний ресурс:]// Посібник користувача /URL: <http://old.exponenta.ru/soft/Mathcad/UsersGuide/0.asp> (дата звернення 18.05.2017)

Ніку Себе, Майкл С. Лью, Айра Коен, Ашутосх Гарг, Томас С.

Розпізнавання емоцій Хуанга з використанням наївного класифікатора Байєса Коші ICPR,2002 рік

63 G. Little Wort, I. Fasel, M. Stewart Bartlett, J. Movellan Повністю автоматичний кодування основних виразів із відео, Каліфорнійський університет, Сандієго

64 К. Масе. Розпізнавання виразу обличчя за допомогою оптичного потоку. ІЕІСЕТранзакції С. 3474–3483, 1991. 2

65 Maja Pantic, Leon JM Rothkrantz, Automatic Analysis of Facial Вирази: стан

мистецтва , транзакції IEEE щодо розпізнавання образів та машинного інтелекту, грудень 2000 р., стор. 1424-1444

66 Айра Коен Ніку Себе, Ларрі Чен, Ашутосх Гарг, Томас Хуанг, Розпізнавання виразу обличчя з відеопослідовностей: тимчасове та статичне Моделювання комп'ютерного зору та розуміння зображень (CVIU) спеціалізоване про розпізнавання обличчя (<http://www.ifp.uiuc.edu/~iracohen/publications.htm>)

67 А. Lanitis, СJ Taylor та TF Cootes. Єдиний підхід до кодування інтерпретація зображень обличчя. На Міжнародній конференції з комп'ютерного зору, С. 368–373, 1995.

68 Н. Себе та М. С. Лью, До покращення показників рейтингу, IEEE Транзакції з аналізу шаблонів і машинного інтелекту, с. 1132-1143, 2000.

69 W. Press, S. Teukosky, W. Vetterling, B. Flannery Numeric Recipe in C : мистецтво наукових обчислень С. 394-455

70 Магістерська робота Іро Коена з розпізнавання емоцій 2000 , Університет Іллінойсу в Urbana Champaign, стор. 1-43

71 Рафаель С. Гонсальс, Річард Е. Вудс Цифрова обробка зображень стор.571-619, Addison-Wesley Press, США, 1992

72 Лу Руцзін Штучний інтелект стор.909-932, Science Press, Пекін, Китай, 1996

73 А. Мехрабян , Спілкування без слів , Психологія сьогодні, вип. 2, №4, С. 53-56, 1968.

74 Y.Tian, T.Kanade, J.Cohn Розпізнавання одиниць дій для виразу обличчя Аналіз Carnegie_Mellon University, IEEE транзакцій за шаблоном розпізнавання та машинного інтелекту vol. 23, No 2, лютий 2001 р.97-115

75 B.Wang Розпізнавання виразів обличчя за допомогою технічних

нейронних мережзвіт 1999, Корнельський університет, США

76 Нейронні мережі Sodha Lection Nodes Department of Computer ScienceМатематика та фізика, Університет Вест-Індії, 2003(<http://scitec.uwichill.edu.bb/cmp/online/p21h/help/splash/>)

77 MJ Black та Y. Yacoob . Відстеження та розпізнавання жорстких і нежорстких обличчярухів з використанням локальних параметричних моделей руху зображення. МіжнароднийКонференція з комп'ютерного зору, стор. 374–381, 1995.

78 Ю. Якоб і Л. С. Девіс. Розпізнавання виразів обличчя людини з довгогопослідовності зображень за допомогою оптичного потоку. Транзакції IEEE щодо аналізу шаблоніві машинний інтелект, стор. 636–642, червень 1996 р .

79 Мартіньї, Багатомасштабне розпізнавання виразів обличчя за допомогою згортковогоНейронні мережі перевершили Fasel IDIAP, Швейцарія

80 П. Екман і В. В. Фрізен. Система кодування дій на обличчі: слідчогоКерівництво. Consulting Psychologists Press, Пало-Альто, Каліфорнія, 1978.

81 П. Екман. Вагомі докази універсальності у виразах обличчя: відповідь наПомилкова критика Рассела. Психологічний вісник, С. 268–287, 1994.

82 М. Розенблум, Ю. Якоб і Л. С. Девіс . Розпізнавання людського виразувід руху з використанням радіальної базисної архітектури мережі . IEEEТранзакції в нейронній мережі, стор. 1121–1138, вересень 1996 р.

83 І. А. Есса та А. П. Пентланд. Кодування, аналіз, інтерпретація та розпізнаваннявиразів обличчя. Транзакції IEEE щодо аналізу шаблонів таМашинний інтелект , с. 757–763, 1997.

84 Г. Донато, М. С. Бартлетт, Дж. К. Хагер, П. Екман і Т. Дж. Сейновський.Класифікація мімічних дій . Транзакції IEEE щодо аналізу шаблонів таМашинний інтелект, с. 974–989, 1999.

85 Т. Оцука та Дж. Охя. Розпізнавання виразів обличчя кількох людейвикористання НММ на основі автоматичного вилучення значущих кадрів із зображенняпослідовності. У Міжнародній конференції IEEE з обробки зображень, стор.546–549, 1997.

86 Т. Оцука та Дж. Охя. Дослідження трансформації міміки на основі розпізнавання виразів із тимчасових послідовностей зображень. Технічнийдоповідь, Інститут електроніки, інформації та зв'язкуІнженери (IEICE), 1997.

87 М. С. Лью, Т. С. Хуанг і К. Вонг, Навчання та вибір функцій уСтереовідповідність , транзакції IEEE на аналізі шаблонів і машиніРозвідка, вип. 16, № 9, 1994.

88 А. Нефіан і М. Хейс. Розпізнавання облич за допомогою вбудованого НММ . вIEEE Conf. на основі аудіо та відео біометричних осібАутентифікація, с. 19–24, 1999.

89 Мартінес. Отримання зображення обличчя за допомогою НММ. У семінарі IEEE наДоступ до бібліотек зображень і відео на основі вмісту, стор. 35–39, 1999.

90 Н. Олівер, А. Пентленд і Ф. Берар. ПІДНІЙ: Обличчя та губи в режимі реального часутрекер з розпізнаванням виразу обличчя. Розпізнавання образів, с. 1369–1382, 2000.

91 М. Бартлетт, Дж. Хагер, П. Екман і Т. Сейновський, Вимірювання виразів обличчя за допомогою комп'ютерного аналізу зображень , Психофізіологія, вип. 36, с.253-264, 1999.

92 Р. Брунеллі та Т. Поджіо, Розпізнавання облич: особливості проти шаблонів ,IEEE Trans. Аналіз шаблонів і машинний інтелект, vol. 15, № 10,С. 1042-1052, жовтень 1993 р.

93 Z. Zhang, Розпізнавання виразу обличчя на основі особливостей:

чутливість Аналіз та експерименти з багатошаровим перцептроном , Int'l J. Pattern Recognition та штучний інтелект, вип. 13, № 6, с. 893-911, 1999.

94 К. Тянь, Н. Себе, М. С. Лью, Е. Лупіас та Т. С. Хуанг, Пошук зображень Використання помітних точок на основі вейвлетів , J. Electronic Imaging, vol. 10, № 4, 2001 рік.



ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ



Кафедра інженерії програмного забезпечення

Магістерська робота

«РЕАЛІЗАЦІЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ
ЕМОЦІЙ ІЗ ЗАСТОСУВАННЯМ НЕЙРОННИХ МЕРЕЖТЕМА»

Виконав: студент 6 курсу, групи ПДМ-61, Кутц А. В.

Керівник: доктор філософії, доцент кафедри ІПЗ Дібрівний О. А.

Київ - 2022

Мета роботи: створення варіанту розв'язання проблеми розпізнавання емоцій обчислювальними машинами та інтелектуальними системами за допомогою нейронних мереж.

Об'єкт дослідження: створення інтелектуальної системи для розпізнавання емоцій за допомогою нейронних мереж.

Предмет дослідження: нейронна мережа для розпізнавання емоцій.

АКТУАЛЬНІСТЬ РОБОТИ

Актуальність дипломної роботи полягає в використанні нейронних мереж для розв'язання формалізованих задач в інтелектуальному аналізі даних. Зростання обсягів інформації, а також розширення кола технічно складних задач прийняття рішень вимагають систематизації наявних методів і розробки нових методик і алгоритмів рішення. У магістерській роботі розглядається можливість застосування нейронної мережі при розв'язанні задачі розпізнавання емоцій на фотографіях і відеозаписах.

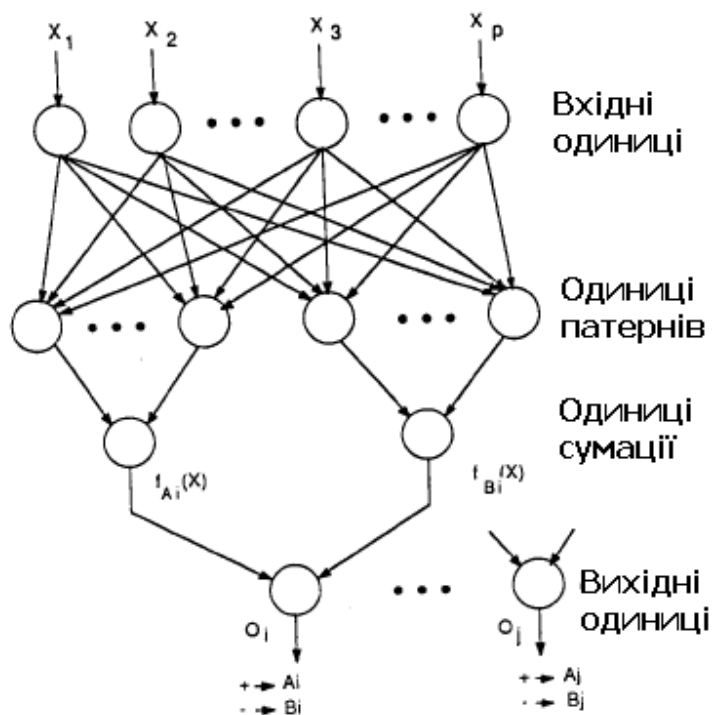
У роботі проведено аналіз наявних проблем при створенні нейронних мереж для розпізнавання емоцій та запропоновано рішення у вигляді розробленої нейронної мережі.

ІСНУЮЧІ МОДЕЛІ/МЕТОДИ/АЛГОРИТМИ ВИРІШЕННЯ ЗАДАЧІ СЧИТУВАННЯ ЄМОЦІЙ

Відомі системи аналізу даних:

- Алгоритми обмеженого перебору;
- Еволюційне програмування;
- Метод найближчого сусіда;
- Деревя рішень;
- Предметно-орієнтовані аналітичні системи;
- Статистичні пакети;
- Нейромережеві пакети.

Створення нейронної мережі



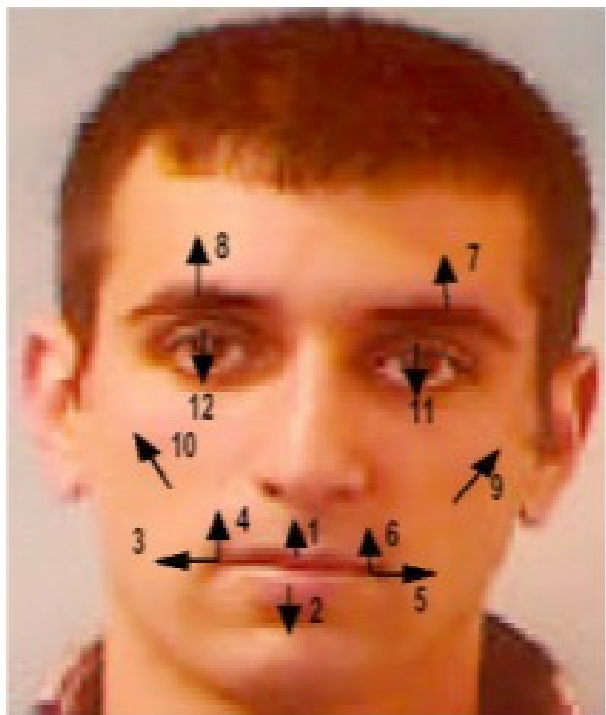
$$z_{in} = \sum_{i=0}^n x_i w_i = x_0 w_0 + x_1 w_1 + x_2 w_2 + \dots + x_n w_n \quad 5.1$$

Загальна кількість вхідних подразників до цього нейрона у вихідному шарі є

$$f(x) = \frac{1}{(1 + e^{-ax})} \quad 5.2$$

Виходом цього нейрона є функція активації

$f(z_{in})$



$$V = B[D_0 D_1 \dots D_m] \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_m \end{bmatrix} = BDP \quad 6.1$$

Лінійна комбінація MU

MU	Опис
1	Вертикальний рух центру верхньої губи
2	Вертикальний рух центру нижньої губи
3	Горизонтальний рух лівого кута рота
4	Вертикальний рух лівого кута рота
5	Горизонтальний рух правого кута рота
6	Вертикальний рух правого кута рота
7	Вертикальний рух правої брів
8	Вертикальний рух лівої брови
9	Піднімання правої щоки
10	Піднімання лівої щоки
11	Моргання правим оком
12	Моргання лівим оком



$$R(V_0 + BDP) + T$$

7.1

Формула формування кривої

Розрахунки

$$f(x, z) = (x-20) * (x-20) + (z-400) * (z-400), \quad 8.1$$

Спускний симплекс метод

$$\begin{aligned} Y(x_1, x_2, x_3, x_4, x_5, x_6) = & (x_1 - 20, 0) * \\ & * (x_1 - 20, 0) + (x_2 - 40, 0) * (x_2 - 40, 0) + (x_3 - 3, 0) * \\ & * (x_3 - 3, 0) + (x_4 - 40, 0) * (x_4 - 40, 0) + (x_5 - 30) * \\ & * (x_5 - 30) + (x_6 - 95) * (x_6 - 95) \end{aligned} \quad 8.2$$

Метод Пауелла

Результати навчання нейронної мережі з використанням спускним симплекс методом 9

	Спокій	Радість	Відраза
Спокій	100%	0,00%	0,00%
Радість	0,00%	100%	0,00%
Відраза	0,00%	0,00%	100%

До навчання

	Спокій	Радість	Відраза
Спокій	82.78	8.71	8.51
Радість	19.67	77.05	3.28
Відраза	25.00	5.82	68.18

Після навчання

Результати навчання нейронної мережі з використанням метода Пауела 10

Емоція	Радість	Здивування	Гнів	Відраза	Страх	Сум
Радість	77.87	0.00	10.66	1.64	9.84	0.00
Здивування	0.00	92.93	0.00	0.00	0.00	7.07
Гнів	0.00	0.00	70.75	27.89	0.00	1.36
Відраза	4.55	0.00	3.41	67.62	1.14	9.09
Страх	25.67	0.00	4.76	0.95	67.62	0.00
Сум	12.70	3.97	8.73	8.73	9.52	55.35

До навчання

Результати навчання нейронної мережі з використанням метода Пауела 11

Пауела

	Нейтральна	Радість	Здивування	Гнів	Відраза	Страх	Сум
Спокій	72.52	0.00	2.89	2.89	13.43	2.69	5.58
Радість	1.64	55.74	0.00	8.20	3.28	27.87	3.28
Здивування	0.00	0.00	90.91	0.00	0.00	1.01	8.08
Гнів	5.44	0.00	0.00	51.70	35.05	2.04	4.76
Відраза	3.41	3.41	0.00	14.77	57.95	2.27	18.18
Страх	11.43	12.38	1.90	0.95	12.38	57.14	3.81
Сум	25.40	0.79	15.67	5.35	12.70	2.38	35.71
Середній	64.52						

Після навчання

ВИСНОВКИ

Реалізація інтелектуальної системи, здатної розпізнавати емоції людини – це рішення технічно та математично складного завдання, що вимагає поглибленого вивчення предметної галузі, власного аналізу, вміння та навичок при обробці та виборі експериментальних даних, глибокі знання в галузі дискретної математики, геометрії, програмування, а також психології та багатьох інших галузях науки. Використання нейронних мереж при розробці систем розпізнавання, дозволяє структурувати погано формалізовані дані, прискорити процес обробки даних, оцінити адекватність середовища, в яке вміщено інформаційну модель нейронної мережі.

Метрична система дозволяє створити прототипи емоцій, завдяки яким у розробника з'являється можливість «взаємодії» з нейронною мережею, тобто. навчання з урахуванням виявлених прототипів емоцій. Основна мета магістерської роботи полягала у виборі інформаційної моделі нейронної мережі, а також опис алгоритму розпізнавання на вирішення практичного завдання. Було обрано інформаційну модель на основі багатошарового перцептрон.

Статті:

1. Кутц А. В. Державний університет телекомунікацій, Київ
ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ ЗА ДОПОМОГОЮ НЕЙРОННИХ
МЕРЕЖ // COMPUTER SYSTEMS AND INFORMATION
TECHNOLOGIES. №3, 2021.

Тези доповідей:

Кутц А.В. Інтелектуальні системи для полегшення процесу і оптимізації
завдань розпізнавання емоцій. // XIII Міжнародна науково-технічна
конференція студентства та молоді «Світ інформації та телекомунікацій».
– Київ: ДУТ, 2021.

ДЯКУЮ ЗА УВАГУ!