

Всеукраїнський конкурс студентських наукових робіт

Комп'ютерні науки

СТУДЕНТСЬКА НАУКОВА РОБОТА

На тему: «Методи структурної класифікації зображень із використанням
нейронної мережі»

Шифр: «Мережі Кохонена»

2019

ЗМІСТ

ВСТУП	3
1. Застосування нейронних мереж для структурної класифікації	5
1.1 Методи виділення ключових точок зображення	5
1.2 Аналіз особливостей мережі Кохонена для задачі класифікації	8
1.3 Прикладні аспекти та формалізація мережі Кохонена	13
2. Результати дослідження	19
2.1 Опис програмної моделі	19
2.2 Аналіз результатів тестування	21
ВИСНОВКИ	25
ПЕРЕЛІК ПОСИЛАНЬ	27

ВСТУП

На цей час у комп'ютерному зорі набули популярності та практичного застосування ряд методів розпізнавання візуальних об'єктів, що засновані на локальних прикметах зображення [1-3]. Основна проблема полягає у тому, що у більшості випадків неможливо адекватно визначити ознаки, на основі яких слід здійснювати розпізнавання. Для задач, для яких такі ознаки вдається виділити, штучні системи розпізнавання набули значного поширення і широко використовуються [1-5].

Розпізнавання образів є однією з найфундаментальніших проблем теорії інтелектуальних систем. Можливості сучасних програмних бібліотек комп'ютерного зору, наприклад, OpenCV, забезпечують вирішення важливих практичних задач у сфері комп'ютерного зору: аналіз вмісту зображень, пошук та розпізнавання заданих об'єктів, виявлення тексту, відстежування рухів об'єктів, виявлення спільних елементів на порівнюваних зображеннях, реалізація методів навчання для конкретних баз відеоданих [10].

Етап розпізнавання при застосуванні цих методів полягає в обчисленні міри релевантності між описами об'єкта та еталона, що задані у вигляді множин дескрипторів. Такий підхід дає можливість комп'ютерній програмі працювати із візуальними образами подібно людині, зір якої теж ґрунтується на локальних прикметах зображення.

У задачі розпізнавання об'ємних баз візуальних об'єктів у системах комп'ютерного зору важливими показниками є результативність і час оброблення. У зв'язку з цим при здійсненні прикладних завдань отримали розповсюдження структурні методи, що базуються на використанні нейронних мереж з метою виявлення закономірностей на множині ознак структурних описів бази зразків.

Така мережа складається з великої кількості окремих обчислювальних елементів ("нейронів"). При функціонуванні нейрон одночасно отримує багато

вхідних сигналів. Кожен вхід має свою власну синаптичну вагу, яка надає входу вплив, необхідний для функції суматора елемента обробки.

Ваги нейрону є мірою сили вхідних зв'язків і моделюють різноманітні синаптичні сили біологічних нейронів. Ваги суттєвого входу підсилюються і, навпаки, вага несуттєвого входу примусово зменшується, що визначає інтенсивність вхідного сигналу. Параметри кожного «нейрона» можуть змінюватися залежно від результатів, отриманих на попередніх наборах даних, змінюючи таким чином і порядок роботи всієї системи [7].

Становить інтерес вивчення властивостей штучних нейронних мереж при застосуванні в задачах класифікації образів, де в практиці спостерігається практично необмежена різноманітність даних, а також дослідження ефективності схем навчання мереж, що враховують ступінь близькості елементів класів у побудованому просторі ознак.

Актуальним методом для таких нейронних мереж є застосування структурної класифікації образів у ситуації, де кількість класів (еталонів) дорівнює числу кластерів, створених на множині характерних ознак (особливих точок) бази образів [8-13]. У такому випадку застосовувана кластеризація безпосередньо виконує групування множини дескрипторів стандартних зображень у фіксований набір класів, тобто вирішує задачу класифікації.

1. Застосування нейронних мереж для структурної класифікації

1.1 Методи виділення ключових точок зображення

Виділення ключових точок є основою багатьох проблем у таких задачах комп'ютерного зору, як розпізнавання об'єктів або структур на зображенні. Сучасні методи використовують різноманітні детектори для виявлення особливих точок (ОТ) і пошуку їх відповідностей на порівнюваних зображеннях.

Особливі точки – це просторові розташування або точки в зображенні, які визначають, що цікаво чи що виділяється на зображенні. Причина, по якій ці точки є особливими, полягає в тому, що незалежно від того, як змінюється зображення, які афінні або проєктивні трансформації виконуються над зображенням, ви все одно зможете знайти ті ж ключові точки в цьому зміненому зображенні в порівнянні з вхідним зображенням.

У нашій роботі використовується новітній метод обчислення бінарних дескрипторів, який називається ORB. В основі методу ORB лежить комбінація таких алгоритмів, як детектор FAST і дескриптор BRIEF з деякими покращеннями [11].

Структура алгоритму ORB показує, що він менш вимогливий до обчислювальних ресурсів у порівнянні з іншими алгоритмами. Виграш у швидкості обчислень пояснюється простою процедурою побудови дескрипторів. ORB дає за результатами тестування виграш у швидкодії при порівняльній або кращій точності, ніж його аналоги SIFT та SURF [5].

Отримання ОТ та дескрипторів в ORB проходить в кілька етапів. На першому етапі для детектування ОТ будується масштабна гаусова піраміда зображення. Потім на кожному масштабному рівні піраміди визначаються екстремуми яскравості. Для цього застосовується алгоритм FAST.

Існують і інші алгоритми детектування ОТ, такі як: алгоритм виявлення кутів Moravec, алгоритм виявлення кутів Harris & Stephens, детектор куту SUSAN. Причиною створення алгоритму FAST було створення детектора ОТ для використання в програмах реального часу на пристроях, що мають обмежені обчислювальні ресурси [4, 5].

Згідно з FAST для кожної точки зображення формується коло деякого радіуса і підраховується число пікселів, що лежать на ньому і мають значення менше або більше яскравості її центру. Якщо таких точок знайдеться більше 75% від їх можливої кількості, то центр кола вважається кандидатом на роль ОТ.

Результатом роботи детекторів є множина особливих точок, для яких необхідно побудувати математичний опис. Вхідними даними дескриптора є зображення і набір особливих точок, виділених на заданому зображенні. Виходом дескриптора є множина векторів ознак для вихідного набору особливих точок. Необхідно відзначити, що якісь дескриптори вирішують одночасно два завдання - пошук особливих точок і побудова описів цих точок.

Для побудови дескриптора отриманих точок в роботі використовується модифікація BRIEF, інваріантна до повороту за рахунок додаткових перетворень.

Використання BRIEF-дескриптора забезпечує розпізнавання однакових ділянок зображення, які були зняті з різних точок зору. При цьому максимально зменшується кількість виконуваних обчислень. Алгоритм розпізнавання зводиться до побудови масиву випадкових дерев (RCT) або наївного баєсівського класифікатора на деякій тренувальній множині зображень і подальшої класифікації ділянок тестових зображень [9]. У спрощеному варіанті може використовуватися метод найближчого сусіда для пошуку найбільш схожого ділянки в тренувальній вибірці. Невелика кількість операцій забезпечується за рахунок подання вектора ознак у вигляді бінарного рядка, а як наслідок, використання в якості міри схожості відстані Хеммінга.

Для складання дескриптора необхідно сформувати квадратне вікно, центроване щодо ключової точки і узгоджене з її орієнтацією. У цьому вікні по заданому правилу вибирається набір пар точок, значення яскравості в яких порівнюються між собою. Якщо яскравість першої точки буде вищою, в відповідний елемент дескриптора записується 1, в іншому випадку записується значення 0. Складені таким чином дескриптори можна зіставляти один з одним за нормою Хемінга.

Результат роботи методу представляється у вигляді вектора довжиною 256, що складається з результатів тестів навколо ОТ. В околиці 31×31 піксель порівнюються значення яскравості між x та y , де x, y - області 5×5 пікселів:

$$\tau(I_{x,y}) := \begin{cases} 1: I_x < I_y \\ 0: I_x \geq I_y \end{cases};$$

де I - середня яскравість обраної області.

Для досягнення інваріантності до обертання область обчислення дескриптора орієнтується по орієнтації особливої точки θ . Все $n = 256$ наборів x_i, y_i формують матрицю S розмірністю $2 \times n$. Далі S за допомогою матриці повороту R_θ орієнтується відповідно до кута θ :

$$S_\theta = R_\theta S.$$

А сам вектор дескриптора записується як:

$$g_n(I, \theta) := f(I) | (x_i, y_i) \in S_\theta,$$

$$f_n(I) := \sum_{1 \leq i \leq n} 2^{i-1} \tau(I; x_i, y_i).$$

Відзначимо, що при однакових умовах проведення експериментів на деяких тестових зображеннях точність детектування за допомогою BRIEF майже в 1.5 рази вище, ніж з використанням аналогічних дескрипторів [4, 5].

1.2 Аналіз особливостей мережі Кохонена для задачі класифікації

Історично першою публікацією, що заклала підвалини для створення нейронних мереж та штучних нейронів, які є їх складовою частиною, вважають роботу Уоррена С. Мак-Каллока та Вальтера Піттса [15]. У цій роботі було започатковано теорію, в основі якої лежав той факт, що всі аспекти нервової діяльності можна моделювати за допомогою мережі елементів, які мають два стійкі стани. Модель такого елемента отримала назву формального нейрона.

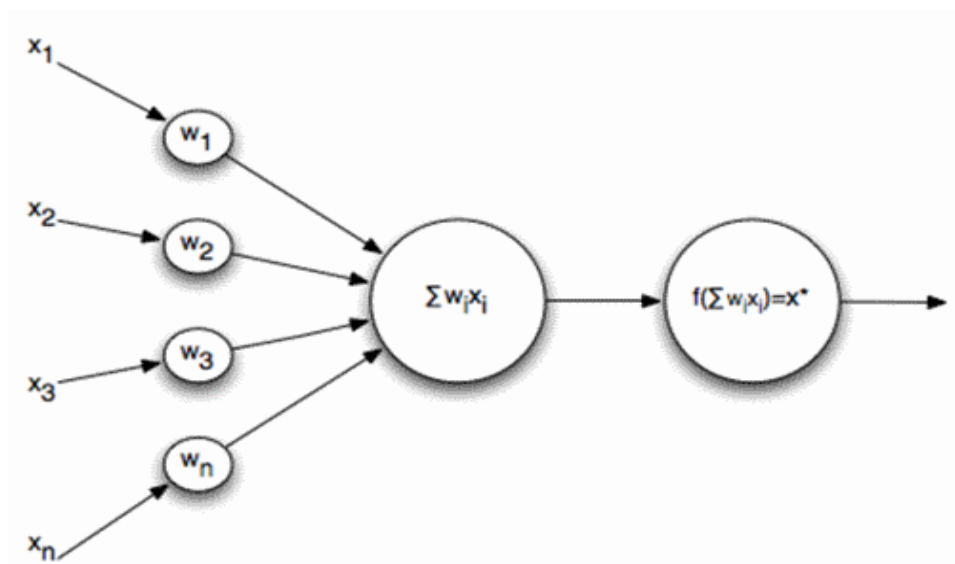


Рисунок 1.1 – Схема формального нейрона

Моделювання функцій синапсів відбувається шляхом масштабування вхідних сигналів (x_0, x_1, \dots, x_n) за допомогою вагових коефіцієнтів (w_0, w_1, \dots, w_n) .

Одержані сигнали подаються до суматора, який виконує їх обробку за формулою:

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

Змінна v є аргументом порогової активаційної функції:

$$y = f(v) .$$

У сучасній літературі зустрічається велика кількість парадигм штучних нейронних мереж, елементи яких реалізують різні активаційні функції.

При застосуванні формальних нейронів із сигмоїдальною або іншою неперервною активаційною функцією на штучних нейронних мережах доцільно розв'язувати задачі регресії. Такий поділ є до деякої міри умовним, оскільки задачі класифікації можуть використовувати також неперервні функції активації, значення яких трактуються як імовірності приналежності до відповідних класів.

Прикладом простішої активаційної функції, яка застосовується частіше, є дискретний сигмоїд із параметром, що має назву функції Хевісайда або тета-функції:

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

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

Використання дискретної активаційної функції призводить до поділу простору вихідних сигналів на деяку множину класів. Завданням нейронної

структури у цьому випадку є встановлення однозначної відповідності між векторами вхідних сигналів і даними класами. Тип задач по встановленню такої відповідності називають задачами класифікації або кластеризації.

Враховуючи особливості поставленої задачі у цій роботі використана нейронна мережа Кохонена. Нейронні мережі Кохонена – клас нейронних мереж, основним елементом яких є шар Кохонена. Шар Кохонена, в свою чергу, складається з лінійних формальних нейронів [7].

Модель мережі Кохонена характеризується структурою розподіленої пам'яті. Така структура дозволяє уникнути катастрофічної деградації у випадку відмови одного з нейронів. Ефект підвищеної живучості досягається саме завдяки розподіленій пам'яті, дія якої проявляється за рахунок того, що за класифікацію вхідного вектора відповідає не один нейрон, а кластер нейронів.

Для застосування нейронних мереж Кохонена в задачах класифікації великих баз зображень потрібна деяка формалізація. Кожен об'єкт, який потрібно класифікувати, представляється у вигляді деякого вектора, що подається на вхід нейронної мережі. Кількість нейронів у вхідному шарі визначається кількістю компонентів цього вхідного вектора. Кількість же виходів визначається кількістю класів, тобто якщо всього є M класів, то кількість нейронів у вихідному шарі теж буде M . Таким чином, кожен нейрон у вихідному шарі «відповідає» за свій клас. Значення, які приймають нейрони в вихідному шарі, відображають наскільки вектор об'єкта який класифікується на вході близький, на думку нейронної мережі Кохонена, до того чи іншого класу. Чим більше «впевненість», що об'єкт належить до того чи іншого класу, тим більше значення приймає нейрон відповідного класу. Іноді застосовують спеціальну функцію-активацію, яка робить так, що сума виходів зі всіх нейронів дорівнює одиниці [7]. В такому випадку кожен вихід можна трактувати, як ймовірність того, що об'єкт належить до даного класу.

Варто відзначити, що існує більш проста реалізація нейронної мережі Кохонена, яка називається «переможець забирає все». У такому випадку кожен

нейрон вихідного шару може приймати значення або нуль, або одиниця. При цьому для одного вхідного вектора одиниці може дорівнювати один і тільки один нейрон вихідного шару, тобто один об'єкт не може ставитися відразу до двох класів.

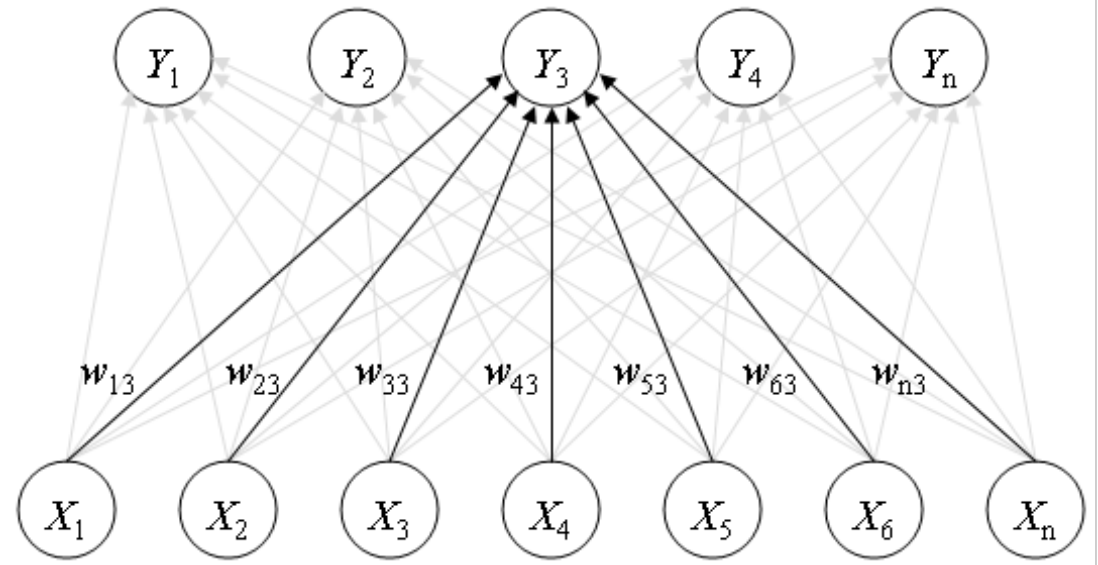


Рисунок 1.2 – Приклад мережі з п'ятьма кластерними одиницями Y і семи вхідними одиницями X

Робота мережі починається з ініціалізації карти, тобто початкового завдання векторів ваги для вузлів. Множина вагових коефіцієнтів структурована у вигляді матриці

$$W = \begin{pmatrix} W^{11}, W^{12}, \dots, W^{1m} \\ W^{21}, W^{22}, \dots, W^{2m} \\ \dots \dots \dots \\ W^{k1}, W^{k2}, \dots, W^{km} \end{pmatrix},$$

елементами якої є вектори вагових коефіцієнтів $W^{ij} = (w_1^{ij}, w_2^{ij}, \dots, w_n^{ij})$, що масштабують відповідні елементи вхідного вектора.

Відома розмірність вхідних даних, по ній певним чином будується початковий варіант карти. В процесі навчання вектори ваги вузлів наближаються до вхідних даних. Для кожного спостереження вибирається найбільш схожий по вектору ваги вузол, і значення його вектора ваги наближається до спостереження. Також до спостереження наближаються вектори ваги декількох вузлів, розташованих поруч, таким чином якщо в множині вхідних даних два спостереження були схожі, на карті їм будуть відповідати близькі вузли. Циклічний процес навчання мережі, що перебирає вхідні дані, закінчується після досягнення картою допустимої (заздалегідь заданій аналітиком) похибки, або після здійснення заздалегідь заданої кількості ітерацій.

На відміну від інших типів нейромереж, самонавчальній карті Кохонена не потрібно задавати вихідний вектор. Навчання самоорганізованої карти Кохонена здійснюється шляхом зміни компонент вектора кожного з вузлів мережі відповідно до значень компонент вхідного вектора.

Вдало обраний спосіб ініціалізації ваг карти може істотно прискорити навчання, і привести до отримання більш якісних результатів. Існують три способи ініціювання початкових ваг:

- 1) ініціалізація випадковими значеннями, коли всім ваг даються малі випадкові величини;
- 2) ініціалізація прикладами, коли в якості початкових значень задаються значення випадково вибраних прикладів з навчальною вибірки;
- 3) лінійна ініціалізація. В цьому випадку ваги ініціюються значеннями векторів, лінійно впорядкованих вздовж лінійного підпростору, що проходить між двома головних власними векторами вихідного набору даних.

В нашому випадку на початку навчання компоненти вектора ваг кожного з вузлів задаються випадковим чином. В процесі ітерацій значення ваг стабілізуються, формуючи зони, близькі за значеннями до кожного набору векторів з навчальної множини.

1.3 Прикладні аспекти та формалізація мережі Кохонена

Для застосування нейронних мереж Кохонена в задачах класифікації великих баз зображень потрібна деяка формалізація. Візьмемо за основу модель класифікації, коли для кожного класу, що задається еталоном, попередньо синтезується його деякий «центр», а класифікація здійснюється шляхом віднесення до класу з найбільш близьким центром [6].

Нехай $W = \{x \mid x \in R^n\}$, $W \subseteq R^n$ – простір структурних ознак (дескрипторів) зображень. База описів зразків (еталонів) $Z \subset W$ задана у вигляді множини $Z = \{Z^j\}_{j=1}^J$, $s_j = \text{card } Z^j$, $s = \text{card } Z$, $s = \sum_{j=1}^J s_j$. При цьому кожна ознака $x_i^j \in Z^j$ асоціюється з еталоном Z^j в плані включення до його складу, бо база зразків задана.

Розглянемо процедуру навчання мережі в наступному вигляді.

1. Ініціюємо матрицю $M = \{m_j\}_{j=1}^J$ ваг, рядками якої є формовані вектора нейронів m_j – центрів класів так, що $m_j = x_i^j$, $x_i^j \in Z^j$, де i – номер довільного вектора з класу зразків Z^j .

2. Сформуємо елементи навчальної множини у вигляді сукупності $Z = \{Z^j\}_{j=1}^J$ всіх s елементів бази еталонів.

3. Виберемо поточний елемент $z \in Z$, для кожного $j \in [1, 2, \dots, J]$ обчислимо відстань $q_j = \rho(z, m_j)$ і визначимо клас d нейрона-переможця: $d = \arg \min_j q_j$. Вид метрики визначається властивостями дескрипторів.

4. Обчислимо зміни ваг для нейронів вихідного шару мережі

$$\Delta m_j = h(j, d, t) \cdot \eta \cdot (z - m_j), \quad (1.1)$$

де η – коефіцієнт швидкості навчання що задається; $h(j, d, t)$ – значення функції околиці для нейрона j в момент часу навчання t ; зазвичай $h(j, d, t) = \exp[-\rho(j, d) / \sigma(t)]$ визначають у вигляді гаусової функції, а радіус $\sigma(t) = 1 / \exp(t^{-2})$ околиці зменшують зі збільшенням параметра t часу навчання. $t = 1, \dots, s$.

5. Коректуємо матрицю ваг $M = M + \Delta M$ на кроці t .

6. Продовжуємо навчання (п. 3-5) до завершення списку Z .

7. Перевіряємо виконання критерію зупинки навчання. Це може бути величина помилки кластеризації. При необхідності (невиконання умови зупинки) продовжуємо навчання з п. 2. Це може бути дублювання множини вхідних даних, вибір даних з множини Z в установленому або у випадковому порядку.

Дослідники відзначають, що нормування даних (приведення норми вектора до одиниці) при застосуванні мереж Кохонена прискорює збіжність мережі до локального мінімуму помилки [2]. Для мереж з незначним числом нейронів (2-3) застосування нормування вважається обов'язковим. Якщо обробляються нормовані дані, то нейрони мережі стають нормованими в ході навчання. У даній ситуації ми використовуємо ненормовані дані, прагнучи забезпечити достатню відстань нейронів між собою з метою забезпечення результативного поділу класів.

Значення помилки кластеризації оцінимо критерієм

$$E = \frac{1}{sN} \sum_{j=1}^J \sum_{v=1}^{s(j)} \rho(x_v, m_j), \quad (1.2)$$

де підсумовуються відстані елементів до центру всередині кластерів, а потім застосовується усереднення по числу s елементів навчальної множини i за розміром N дескриптора ОТ.

Іншим ключовим індексом для оцінювання якості класифікації може бути значення критерію, який підраховував частку елементів навчальної вибірки Z , які потрапили «не в свої» класи. Він обчислюється як

$$\beta = \sum_{j=1}^J (s_j - a_j) / s, \quad (1.3)$$

де a_j – число ознак із загального їх числа s_j в описі еталона Z^j , віднесених в процесі класифікації до класу j . Значення β відображають рівень помилкових рішень при класифікації ознак множини Z . Чим ближче β до нуля, тим вище досягнуто якість класифікації на навчальній вибірці.

Оптимізація роботи нейронної мережі полягає не тільки у настройці її коефіцієнтів для досягнення бажаної реакції на вхідні дані, а й в оптимізації всієї архітектури мережі. Такий підхід включає оптимізацію топології штучної нейронної мережі та узагальнення з метою забезпечення ефективного функціонування в значно більшому просторі, ніж той, що використовувався в процесі її навчання.

Отже в рамках оптимізації спочатку вивчимо в порівняльному плані три практичних варіанти побудови класифікатора, що відрізняються числом нейронів, які модифікуються в процесі навчання:

- 1) налаштовується тільки нейрон-переможець;
- 2) модифікуються три нейрона, найближчих по відстані до нейрона-переможця;
- 3) модифікується вся мережа нейронів відповідно до відстані до нейрона-переможця.

Оцінимо результативність розпізнавання з застосуванням цих варіантів навчання мережі, різних за глибиною аналізу даних. Критерієм оптимізації, як і під час навчання, є мінімізація цільової функції, в ролі якої найчастіше

використовують середньоквадратичну похибку, що поділяється на навчальну та реальну.

Обговорювані варіанти навчання класифікатора принципово відрізняються лише п. 4, де по-різному обчислюється функція $h(j,d,t)$. У той же час, для досягнення збіжності повинна бути виконана умова $h(j,d,t) \in [0,1]$.

При цьому версія 1 реалізує грубе розпізнавання (підлаштовується тільки переможець, принцип «переможець забирає все»). Версія 2 відносить елемент до трьох найбільш ближніх центрів (м'які обчислення, «переможець забирає більше»), а версія 3 здійснює найбільш тонку обробку, подібно непевному підходу відносячи аналізований дескриптор до всіх кластерів одночасно, але з різними вагами, пропорційними відстані. Версії 2 і 3 дозволяють управляти нейронами - сусідами переможця, і фактично реалізують принцип «нейронного газу», відповідно до якого здійснюють ранжування елементів мережі з околиці ознаки навчання [1, 3].

Тривіальним підходом до розширення сфери ефективного функціонування нейронної мережі є пряме застосування знань про всю множину вхідних даних на етапі формування архітектури та навчання. Але такий підхід, як правило, є недосяжним в основному з двох причин, якими є:

- 1) відсутність знань про всю можливу множину вхідних даних;
- 2) великий об'єм роботи у випадку великих масивів вхідних даних.

Але в розглянутому випадку ці причини не зупиняють такий підхід до оптимізації. Отже, знаючи тип вхідних даних та використовуючи той факт, що результативність роботи мережі Кохонена в значній мірі визначається початковими умовами її функціонування, які визначаються значеннями початкових центрів m_j ; в п. 1 процедури вивчимо такі варіанти вибору центрів:

- 1) довільно по одному дескриптору з кожного зразка;
- 2) за допомогою застосування спеціальної процедури формування центрів.

З огляду на бінарне представлення дескрипторів, для кожного Z^i визначимо вектор центру класу на підставі логічного правила

$$m_i(b) = \begin{cases} 1, & \sum_{d=1}^{s(i)} x_d(b) \geq s(i)/2, \\ 0, & \sum_{d=1}^{s(i)} x_d(b) < s(i)/2, \end{cases}, x_d \in Z^i, b=1, \dots, 256, \quad (1.4)$$

де $x_d(b)$ – біт з номером b для дескриптора s номером d в описі еталону.

Відповідно до (1.4) значення кожного з бітів центру m_i визначається значеннями відповідних бітів всіх дескрипторів, що належать еталону з номером i . Центр (1.4) відображає властивості еталона (класу).

Проаналізуємо елементи $x \in Z$ загального змісту структурних описів бази еталонів (навчальна вибірка) шляхом віднесення їх до відповідного класу v з використанням конкурентного способу

$$v = \arg \min_{i=1, \dots, J} \rho(x, m_i). \quad (1.5)$$

Відзначимо, що в (1.5), як і в п. 3 процедури навчання Кохонена, необхідно визначити метрику для обчислення відмінності двох дескрипторів. У методі Кохонена природним є застосування евклідової метрики або її квадрата.

Ця відстань обчислюється таким чином:

$$\rho(x, m_i) = \sum_{b=1}^{256} (x(b) - m_i(b))^2. \quad (1.6)$$

Ця міра може застосовуватися для обчислення відстані між об'єктами, описаними кількісними, якісними і дихотомічними ознаками. Її використання доцільно, коли ознаки однорідні за смисловим навантаженням і однаково важливі для

розв'язуваної задачі. Саме вона найбільш підходить для роботи з даними детекторів так як обробляються дійсні дані (де 256 - розмір детектора ORB).

В результаті обробки для кожного зразка відповідно до процедури Кохонена або з співвідношенням (1.5) отримуємо кластерне уявлення $h[Z^i] = (h_1^i, \dots, h_J^i)$ еталона, де h_a^i – цілі числа. Воно відповідає розподілу елементів множини Z^i за класами еталонів. Якщо розглянути набір векторів $h[Z^i]$ як рядки матриці розміром $J \times J$, то помилка (1.3) обчислюється через суму діагоналі цієї матриці, що ділиться на загальну кількість ознак s бази.

2. Результати дослідження

2.1 Опис програмної моделі

У середовищі C++, під операційною системою Windows, реалізована модель класифікації зображень на основі поелементного віднесення ОТ об'єкта до одного з центрів класів з використанням нейронної мережі Кохонена. Обчислення, аналіз і обробка дескрипторів ОТ реалізовані у вигляді програмного засобу на основі модуля features2d для крос-платформної бібліотеки OpenCV.

Функції цього модуля забезпечують у проекті зберігання даних зображення, визначення ОТ на зображенні, зберігання даних про ключові точки та дескриптори, порівняння ОТ та знаходження відповідностей для дескриптора із набору запитів та виведення зображень та даних, отриманих після їх обробки.

Для роботи алгоритму та самої програми побудовано базу зображень з еталонними зображеннями на основі яких будуть побудовані центри кластерів. За базу еталонів застосований набір з 5 зображень грошових знаків приклад яких можна побачити на рисунку 2.1.



Рисунок 2.1 - Приклади зображень з координатами виділених ОТ

З метою формування якісних структурних описів, стійких до фонових і випадковим перешкод, на основі використання стандартних функцій бібліотеки OpenCV проведена спеціальна обробка аналізованих зображень. В ході такої обробки зображення піддавалося фільтрації і бінаризації: зображення переводилося в градації відтінків сірого, обраний поріг яскравості за критерієм бінаризації Отсу, проведена бінаризація [11].

Метод Оцу є методом, заснованим на кластеризації, для автоматичного обчислення порогового зображення [3], або зведення сірого зображення до бінарного зображення. Алгоритм передбачає, що зображення містить два класи пікселів, наступної бі-модальної гистограми: пікселі переднього плану і пікселі тла, потім обчислюється оптимальний поріг, що розділяє два класи, так, що їх комбінований діапазон (дисперсія кластера) є мінімальною або рівноцінною (тому що сума попарних квадратичних відстаней постійна), так, що їх міжкластерна дисперсія є максимальною.

Застосована методика направлена на виділення контуру об'єкта і знижує вплив перешкод, пересічних областей півтонів, випадкових областей з різкою зміною інтенсивності.

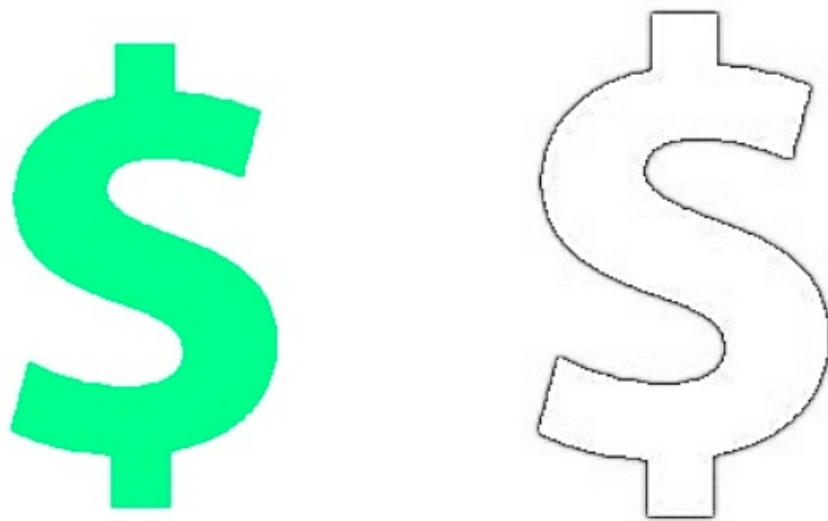


Рисунок 2.2 - Початкова обробка зображення

Дескриптори мають вигляд бінарного вектора розміром 256, кодуються в Open CV у вигляді типу `uchar` (`unsigned char` - тип даних в C++, використовують для зберігання символів, обсяг 8 біт, значення 0 ... 255), а не в бітах. Дані зберігаються в матриці, де кількість рядків дорівнює числу виявлених дескрипторів, а число стовпців дорівнює 32 (256 біт дескриптора трансформуються в 32 `uchar`).

Після отримання зазначених дескрипторів будується центр кластеризації для кожного з еталонних зображень. Оскільки дескриптор має бінарний вид потрібно просто порівняти сумарну кількість нулів та одиниць, для кожного з 256 бітів, та обрати значення, що найбільш часто зустрічаються для усього вектора в цілому.

Отримавши центри, ми можемо почати класифікацію зображень. Для цього було збудовано та проведено тестування двох алгоритмів. Перший – це SOM Кохонена, а другий алгоритм – це евристичний підхід бінарного аналізу, який описано у розділі 1.3.

Також було програмно реалізовано додаткові функції такі як побудова графіків та гістограм дескрипторів, конвекторів даних як внутрішнього використання у алгоритмах, так й для виводу інформації для користувача, що дозволяє оцінити роботу алгоритмів й порівняти їх

2.2 Аналіз результатів тестування

Використовуючи раніше зазначену базу зображень грошових знаків проведемо обробку (1.4) для множин дескрипторів кожного з еталонів і для отриманих центрів з використанням метрики Хеммінга виконаємо класифікацію (1.5). В результаті отримана матриця розподілу ОТ за класами у вигляді таблиці 2.1.

Кількість виявлених ОТ на заданій базі зображень коливалося в межах 47 ... 50. Вихідний набір даних містить 244 бінарних вектора довжини 256. Кількість ітерацій дорівнює кількості вхідних векторів. Матриця ваг - матриця дійсних чисел розміром 244 на 256.

Таблиця 2.1 – Розподіл ОТ за класами з центрами (1.4)

Еталони	Класи				
	1	2	3	4	5
Z^1	14	17	3	4	14
Z^2	9	28	7	1	8
Z^3	10	18	7	7	11
Z^4	16	15	4	12	7
Z^5	7	14	2	1	29

Як видно з таблиці 2.1, множина дескрипторів ORB для конкретної бази зображень з використанням ретушування та бінарної обробки не дозволяє однозначно класифікувати множину ОТ кожного з еталонів. Причина - число ОТ, віднесених до певного класу, не є максимальним для еталонів Z^1 , Z^3 и Z^4 . Помилка класифікації становить $E = 0,74$ за формулою (1.2) та $\beta = 0,68$ за формулою (1.3).

З використанням розробленої системи вивчені та розроблені в порівняльному плані три раніше описаних варіанти навчання класифікатора, що відрізняються числом нейронів, які модифікуються в процесі адаптації мережі. Початковий набір даних містить 244 бінарних вектора довжини 256. Кількість ітерацій дорівнює кількості вхідних векторів. Матриця ваг - матриця дійсних чисел розміром 244 на 256.

У таблиці 2.2 показані розподілення дескрипторів ОТ за класами (варіанти довільного вибору центрів і обчислення за формулою (1.4)) для версії, коли в процесі навчання налаштовується тільки нейрон-переможець.

Таблиця 2.2 – Розподіл ОТ еталонів у першому варіанті обробки

Довільний вибір центрів						Формування центрів (1.4)					
Еталони	Класи					Еталони	Класи				
	1	2	3	4	5		1	2	3	4	5
Z ¹	19	0	20	3	8	Z ¹	41	4	2	0	3
Z ²	6	14	18	5	7	Z ²	20	26	1	0	3
Z ³	6	5	18	3	15	Z ³	17	2	23	0	5
Z ⁴	6	1	14	21	8	Z ⁴	14	2	2	30	2
Z ⁵	4	4	5	20	14	Z ⁵	4	4	9	0	30

У другій половині таблиці спостерігається високий рівень розрізнення еталонів: максимуми значень знаходяться на діагоналях і значно перевищують інші елементи. Помилки в разі довільних центрів складають $E = 0,63$ та $\beta = 0,67$. У разі сформованих по (1.4) центрів помилки складають – $E = 0,38$ та $\beta = 0,34$. Як бачимо, спеціальне формування центрів (1.4) призводить до зниження помилки класифікації, а отже, поліпшується якість розрізнення об'єктів по безлічі значень їх дескрипторів. Крім того, значення помилок достатньо близькі.

У таблиці 2.3 показані варіанти модифікації трьох найближчих до переможця нейронів.

Таблиця 2.3 – Розподіл ОТ еталонів у другому варіанті обробки

Довільний вибір центрів						Формування центрів (1.4)					
Эталони	Класи					Эталони	Класи				
	1	2	3	4	5		1	2	3	4	5
Z ¹	23	0	18	2	7	Z ¹	43	3	1	0	3
Z ²	3	20	16	5	6	Z ²	17	30	1	0	2
Z ³	5	3	22	3	14	Z ³	14	1	28	0	4
Z ⁴	6	1	13	23	7	Z ⁴	12	1	2	34	1
Z ⁵	3	4	4	17	19	Z ⁵	3	3	7	0	35

Для варіанту модифікації трьох найближчих до переможця нейронів (таблиця 2.3) в разі довільних центрів помилка становить $E = 0,62$ та $\beta = 0,60$. Для сформованих у вигляді (1.4) центрів $E = 0,30$ та $\beta = 0,26$.

При модифікації всієї мережі нейронів (таблиця 2.4) помилка в разі довільних центрів становить $E = 0,58$ и $\beta = 0,61$. Для сформованих у вигляді (1.4) центрів $E = 0,22$ та $\beta = 0,20$.

Таблиця 2.4 - Розподіл ОТ еталонів у третьому варіанті обробки

Довільний вибір центрів						Формування центрів (1.4)					
Еталони	Класи					Еталони	Класи				
	1	2	3	4	5		1	2	3	4	5
Z^1	25	0	14	4	7	Z^1	48	1	1	0	0
Z^2	3	27	9	5	6	Z^2	8	41	0	0	1
Z^3	1	3	30	3	11	Z^3	6	1	37	0	3
Z^4	4	1	9	31	5	Z^4	8	0	1	40	1
Z^5	3	4	3	11	26	Z^5	2	3	4	0	39

За отриманими результатами, з розширенням числа настроюваних нейронів, особливо для варіанту спеціально сформованих центрів (1.4), помилка класифікації істотно знижується. Наприклад, динаміка зміни значень β для трьох варіантів обробки склала 0,34, 0,26 і 0,20. Це підкреслює можливості адаптації мережі Кохонена до довільних даних, особливо в разі успішного формування первинних центрів класів. Розширення набору адаптованих нейронів від 1 до загального їх числа підвищує показники успішної роботи мережі, а отже, поліпшується якість розрізнення об'єктів по множині значень дескрипторів.

ВИСНОВКИ

У роботі з використанням апарату кластеризації вирішена задача класифікації зображень на базі структурного опису у вигляді множини особливих точок, що визначаються детектором ORB. Для цього проаналізовано особливості детектору, запропоновано способи визначення центрів кластерів шляхом побітового логічного оброблення замість більш складних метрик.

Дослідження підтвердило універсальність і можливості успішної адаптації мережі Кохонена до довільних наборів візуальних даних, особливо в разі успішного формування первинних центрів класів. В ході навчання мережа Кохонена визначає ступінь розрізненості даних, покращуючи кінцеві показники класифікації. Це дозволяє вирішувати більш складний спектр завдань класифікації. Розширення набору адаптованих нейронів в процесі навчання значимо удосконалює показники дієвої роботи мережі.

Ефективність класифікації візуальних об'єктів по множині особливих точок з використанням мережі Кохонена безпосередньо залежить від наступних ключових взаємопов'язаних факторів: розпізнавана база, метод попередньої обробки, спосіб формування дескрипторів, початковий вибір центрів, метрика для порівняння дескрипторів, розмір кортежу адаптованих нейронів.

Поряд з шляхом безпосереднього віднесення дескрипторів, зображення може розпізнаватися по максимальній подібності до центрів. В класифікації можуть бути застосовані способи формування подоби з ваговим значенням кожного з класів. Ці методи менш чутливі до фактору рівномірності розподілу дескрипторів еталонів по класах.

Наукова новизна дослідження полягає в побудові методу структурної класифікації на основі модифікацій мережі Кохонена, що забезпечує необхідну результативність розрізнення та у синтезі нового продуктивного методу

структурного розпізнавання зображень на підставі бітового оброблення і побудови класифікаційних рішень у просторі кластер-еталон.

Практична значущість роботи – отримання прикладних програмних моделей для застосування та оцінювання ефективності методів структурного розпізнавання і підтвердження їх результативності в конкретних прикладах баз зображень.

Подальше поліпшення результативності розпізнавання може бути досягнуто за рахунок підходів навчання з учителем, тому що класи дескрипторів бази відомі до початку функціонування системи комп'ютерного зору. Зважаючи на бінарний вид дескрипторів ORB, є можливість для побудови результативних правил класифікації скористатися механізмами дерев рішення чи хешування. Іншим розвитком методів навчання на мережах Кохонена є застосування пакетної обробки, що має, як правило, більш високі показники якості класифікації через ретельний і багаторазовий аналіз даних.

ПЕРЕЛІК ПОСИЛАНЬ

- 1 Новотарський М.А. Штучні нейронні мережі: обчислення. [Текст] / М.А. Новотарський, Б.Б. Нестеренко // Праці інституту математики НАН України. – 2004, т. 50. – 408 с.
- 2 Осовский С. Нейронные сети для обработки информации. [Текст] / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.
- 3 Патин М. В. Сравнительный анализ методов поиска особых точек и дескрипторов при группировке изображений по схожему содержанию / М. В. Патин, Д. В. Коробов // Молодой ученый. — 2016. — №11. — С. 214. [Электронный ресурс]. – Режим доступа: <https://moluch.ru/archive/115/31188/>.
- 4 Ethan Rublee ORB: an efficient alternative to SIFT or SURF [Text] / Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski. – Computer Vision (ICCV), IEEE International Conference on IEEE – 2011, 2564 – 2571 pp.
- 5 Stefan Leutenegger. BRISK: Binary Robust Invariant Scalable Keypoints. [Text] / Stefan Leutenegger, Margarita Chli, Roland Y. Siegwart – Computer Vision (ICCV), pp. 2548 – 2555, 2011.
- 6 Борисов Е.С. Кластеризатор на основе нейронной сети Кохонена. [Электронный ресурс]. – Режим доступа: <http://mechanoid.kiev.ua/neural-net-kohonen-clusterization.html>
- 7 Кохонен Т. Самоорганизующиеся карты [Текст] / Т. Кохонен – М.: БИНОМ. Лаборатория знаний, 2013. – 655 с.
- 8 Гороховатский В.А. Структурное распознавание изображений с применением моделей интеллектуальной обработки и самоорганизации признаков [Текст] / В.А. Гороховатский, А.В. Гороховатский, А.Е. Берестовский // Радиоэлектроника, информатика, управление.–2016. – №3 (38). – С. 39–46.

9 Gorokhovatsky V.A. Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions [Text] / V.A. Gorokhovatsky // Telecommunications and Radio Engineering. – 2016, Vol. 75, No 14. – P. 1271–1283.

10 Гороховатский В.А. Структурный анализ и интеллектуальная обработка данных в компьютерном зрении [Текст] / В.А. Гороховатский. – Х.: Компания СМІТ, 2014. – 316 с.

11 Гороховатський В.О. Аналіз властивостей, характеристик та результатів застосування новітніх детекторів для визначення особливих точок зображення [Текст] / В.О. Гороховатський, Д.В. Пупченко, К. Г. Солодченко // Системи управління, навігації та зв'язку. –2018. – №1 (47). – С. 93–98.

12 Пупченко Д.В. Классификация изображений визуальных объектов по множеству дескрипторов особенных точек на основе нейронной сети Кохонена [Текст] / Д.В. Пупченко, В.О. Гороховатський // Системи управління, навігації та зв'язку. –2018. – №1 (47). – С. 68–72.

13 Пупченко Д.В. Изучение адаптационных свойств сети Кохонена в задаче классификации изображений [Текст] / Д.В. Пупченко, В.А. Гороховатский, Е.П. Путятин // Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI'2018): збірка наук. праць міжн. наук. конф., с. Залізний Порт, 21–27 травня 2018. – Херсон: ФОП Вишемирський В.С., 2018. – С. 274–276.

14 Пупченко Д.В. Применение нейронной сети Кохонена для структурной классификации изображений [Текст] / Д.В. Пупченко, В.О. Гороховатський // XXII Міжнародний молодіжний форум Радіоелектроніка та молодь у XXI столітті. Зб. Матеріалів форуму. Т.7. – Харків: ХНУРЕ. 2018. – 13–14с.

15 McCulloch W.S. A logical calculus of the ideas immanent in nervous activity [Text] / McCulloch W.S., Pitts W. – Bulletin of Mathematical Biophysics, 1943.–№ 5.– P.115-133.