

РОЗРОБКА КОМБІНОВАНОЇ МОДЕЛІ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

Волошин М. В.

1. Вступ

З кожним роком зростають обсяги інформації, що потребує формалізації і подальшої алгоритмізації процесів, які раніше виконувалися вручну. 80 % інформації людина отримує за допомогою зору, тому будь-які системи, які пов'язані з автоматичною обробкою зображень є затребуваними. Одним з ключових понять в автоматичній обробці є поняття розпізнавання об'єктів, яке являє собою область активної наукової роботи протягом останніх двадцяти років. Коли алгоритми виконують розпізнавання на рівні експерта-людини, то автоматизація веде до прискорення роботи систем обробки даних і підвищення їх ефективності.

Незважаючи на значний успіх, розпізнавання зображень тільки в деяких областях здатне за якістю результату порівнятися або перевершити когнітивну функцію сприйняття у людини. В цілому, задача розпізнавання зображення до сьогодні повністю не вирішена.

Існують різні методи для розпізнавання зображень: потенційні функції, байєсівські мережі, марківські мережі, штучні нейронні мережі, різні види асоціативної пам'яті і таке інше. Дослідження задачі розпізнавання зображень показало, що розпізнавання проводиться методами, які не повністю враховують особливості графічних об'єктів, головними з яких є малий обсяг апріорних даних відносно еталонних описів об'єктів розпізнавання. Тому актуальним є розробка універсальної моделі, що дозволяє з великою долею ймовірності віднести об'єкт розпізнавання до визначеного класу об'єктів при малій навчальній вибірці.

2. Об'єкт дослідження та його технологічний аудит

Об'єктом дослідження є процеси ідентифікації та класифікації об'єктів у задачах комп'ютерного зору. В даний час для розпізнавання зображень найкращі результати показують стохастичні моделі, а саме: підклас штучних нейронних мереж – згорткові нейронні мережі. Однак, навчання нейронних мереж – це погано обумовлена задача. Погана обумовленість означає, що навіть великий набір даних може нести в собі малий обсяг інформації про задачу, що вирішується. Тому ключова роль в процесі синтезу параметрів конкретної математичної моделі нейронної мережі належить навчальним даним. Підбір репрезентативної навчальної вибірки є однією з найскладніших задач в машинному навчанні і не завжди можливий на практиці. Отже виникає необхідність в розробці універсального методу, що дозволяє з великою долею ймовірності віднести об'єкт розпізнавання до визначеного класу об'єктів при малій навчальній вибірці.

3. Мета та задачі дослідження

Метою роботи є підвищення ефективності обробки зображень при розв'язанні технічних задач за рахунок використання нових підходів.

Для реалізації поставленої мети необхідно вирішити наступні задачі:

1. Проаналізувати розроблені методи на предмет обчислювальної складності при розв'язанні практичної задачі обробки зображень.
2. Розробити комбінований метод рішення задачі обробки зображень.
3. Обґрунтувати необхідність та переваги використання нового комбінованого методу обробки зображень.

4. Дослідження існуючих рішень проблеми

Протягом останніх декількох років, був досягнутий значний прогрес в сфері розпізнавання об'єктів при малих варіаціях в освітленні, сцені і т. ін. [1], але надійні методи розпізнавання при більш екстремальних змінах були недосяжні.

Задача розпізнавання об'єктів за допомогою класифікаторів традиційно ділиться на дві частини: виділення ключових ознак та класифікація за цими ознаками [2]. Виокремлення ознак здійснюється за рахунок апріорної інформації, внаслідок чого досягається структурна інваріантність [3].

Статистичні методи [4] використовуються для аналізу зображень, але в тих випадках, коли не потрібна висока точність розпізнавання. На вхід системи подають не лише пікселі зображення, але і деяке низькорівневе представлення зображення, наприклад, вейвлет-коефіцієнти. Машина опорних векторів є універсальним апроксиматором, яка може звести до нуля помилку навчання на навчальній вибірці, що, в свою чергу, дозволяє розраховувати на низьку помилку узагальнення на тестовій множині. Однак, практика показала, що помилка узагальнення для мереж з великою кількістю шарів (deep architecture) нижча, ніж для мереж з невеликою кількістю шарів (shallow architecture) [5]. Класичні перцептрони, RBF-мережі та SVM, мають малу кількість шарів і задача навчання мережі при розширенні кількості шарів стає досить складною.

В більшості випадків для досліджень обирають проекційні напрямки, які є ортогональними на всьому діапазоні свого класу. Метод аналізу головних компонент (PCA) заснований на лінійному проектуванні простору зображення в низькорозмірному просторі ознак. В основі даного алгоритму лежить використання фундаментальних статистичних характеристик таких, як: середні математичні очікування та коваріаційна матриця. Для кожного об'єкта обчислюють його головні компоненти. Процес розпізнавання полягає в порівнянні головних компонент невідомого об'єкта з компонентами решти об'єктів. Однак такий підхід максимізує загальний діапазон по всіх класах. PCA проекція добре підходить для відновлення з низькопросторового базису, але не може бути оптимальною з дискримінантної точки зору. Основний недолік: високі вимоги до умов зйомки зображень. Зображення повинні бути одержані в близьких умовах освітленості, однаковому ракурсі і повинна бути проведена якісна попередня обробка, що приводить зображення до стандартних умов (масштаб, поворот, центрування, вирівнювання яскравості, відсікання фону). Небажана наявність різного роду спотворень та інших внутрішньокласових варіацій.

Лінійний дискримінант – це «класичний» метод при розпізнаванні образів, в якому характеристики об’єкта для розпізнавання отримують за допомогою полярного квантування форми.

Нейромережеві структури дають досить високі результати за рахунок своєї нелінійної структури. До недоліків мереж відносять швидке зростання їх розміру при спробі налаштування на велику кількість інваріантних перетворень, використання нейронів з пороговою функцією активації та спеціальні алгоритми для їх налаштування [6].

Згорткові нейронні мережі (СНР – Convolutional Neural Network) належать до багат шарових нейронних мереж, які представляють собою модернізований багат шаровий перцептрон [7]. СНР складається з двох частин: одна частина відповідає за виділення ознак, а інша відповідає за класифікацію. Для виділення ознак використовують чергування шарів згортки (С-шари) і підвибірки (S-шари). Це один з варіантів чергування шарів складних і простих клітин, який з’явився в неокогнітроні. Нейрони в цих шарах організовані в пластини. Кожен нейрон сприймає інформацію з попереднього шару за допомогою свого рецептивного поля, яке являє собою групу нейронів, пов’язаних з поточним нейроном за допомогою параметрів налаштування. В ролі класифікатора виступає двошаровий перцептрон.

В роботах [8, 9] описані результати експериментальних досліджень методів класифікації. В ході роботи було запропоновано процес, що містить наступні методи класифікації:

- наївний байєсівський класифікатор (NB);
- метод *k*-найближчих сусідів (*k*-NN);
- дерево рішень (DT);
- машина опорних векторів (SVM);
- нейронна мережа (ANN);
- лінійний дискримінантний аналіз (LDA).

Результати дослідження зведені в табл. 1.

Таблиця 1

Точність розпізнавання на різних наборах даних

Номер набору	NB	k-NN	DT	SVM	ANN	LDA
1	86,75 %	–	83,72 %	88,07 %	–	76,5 %
2	87,02 %	–	93,17 %	85,87 %	95,16 %	–
3	–	93,80 %	–	91,13 %	94,10%	77,0 %

Примітка: побудовано на основі даних [8, 9]

З аналізу представлених в табл. 1 даних випливає, що нейромережева структура (ANN) має найнижчий відсоток помилки, але нейронні мережі мають ряд недоліків:

- більшість підходів проектування ANN є евристичними і часто не приводять до однозначних рішень;
- для побудови моделі об’єкту на основі ANN необхідно виконати багатоциклічне налаштування внутрішніх елементів і зв’язків між ними;

– проблеми підготовки навчальної вибірки пов'язані з труднощами пошуку достатньої кількості навчальних прикладів;

– значні часові витрати навчання не дозволяють застосовувати ANN в системах реального часу;

– поведінку навченої ANN неможливо однозначно передбачити, що збільшує ризик застосування ANN для керування дорогими технічними об'єктами.

Однією з ключових задач при реалізації нейромережевих структур є пошук оптимального співвідношення параметрів та їх характеристик в кожному конкретному випадку. Для ефективного вирішення такої задачі необхідний широкий спектр методів, алгоритмів і способів синтезу, що різняться обсягом обчислень, якістю результатів, часом пошуку рішень та способами представлення даних.

Таким чином, результати літературного аналізу дозволяють зробити висновок про те, що такі підходи та методи все ще залишаються надійними засобами розпізнавання, але використовуючи тільки такі підходи неможливо отримати кращі результати для різних прикладних задач [10]. Тому необхідно використовувати комбіновані підходи та методи, які орієнтовані на мінімізацію апаратних витрат при реалізації систем комп'ютерного зору.

5. Методи досліджень

Задача розпізнавання являє собою відношення:

$$Z \subseteq I^{in} \times I^{out},$$

де I^{in} – інформація про деяку множину об'єктів, розбиту на класи (їх кількість кінцева), причому спосіб представлення інформації прецедентний, а для Z відома лише частина інформації (навчальна і контрольна вибірки);

I^{out} – мітки/індекси класів. Розпізнавання об'єктів можна розділити на етапи:

- 1) виділення області спостереження;
- 2) кодування знайденої області для подальших етапів розпізнавання;
- 3) співставлення та побудова моделей класифікації;
- 4) обрахунок ймовірності приналежності до визначеного класу;
- 5) узагальнення отриманих результатів.

Перші три кроки належать до підготовчих етапів розпізнавання зображень, в ході яких формується модель придатна до розпізнавання об'єктами методом. В рамках даної роботи такі методи не будуть розглянуті. В даній роботі розглядається комбінована модель розпізнавання зображень на основі теорії несилової взаємодії. Основоположним поняттям теорії несилової взаємодії є поняття інформації. У загальноприйнятому тлумаченні під інформацією розуміють відомості, дані, знання, які отримуються і передаються в процесі взаємодії з іншими суб'єктами. Іншим основоположним поняттям є *Vir*-інтерпретація руху, яка подається через множину зміщень за або проти напрямку руху зі швидкістю, яка рівна швидкості світла у вакуумі та ймовірністю, яка визначається інтроформаційним вмістом рухомих об'єктів. Елементи теорії несилової взаємодії вперше застосовуються в якості апарату

для розпізнавання зображень. В основному теорія несилової взаємодії використовується для вирішення таких задач:

- доступ до баз даних на основі аналізу природно-мовних фонем;
- оцінка інвестиційних пропозицій;
- оцінка та прогнозування впливу шкідливих речовин у водних ресурсах на здоров'я населення;
- прогнозування результатів спортивних подій.

Модель зображення для розпізнавання інтроформаційним підходом в рамках даної роботи будується на основі підходу з голосуванням та елементів методу локальних бінарних шаблонів. Ідейною основою голосування в комп'ютерному зорі вважається перетворення Хафа, а теоретичною базою виступають теорія множин, методи обчислення оцінок, статистичний аналіз. Важливою властивістю системи, побудованої на голосуванні, є можливість отримати результуючий список фрагментів або ознак, за якими приймається рішення, а також можливість достроково зупинити розпізнавання при досягненні певних характеристик, заощадивши на обсязі обчислень.

Нехай класи об'єктів представлені еталонними описами у вигляді кінцевих множин примітивів, які обчислюються для набору даних класифікації. Нехай об'єкт класифікації $Z = \{z_i\}_{i=1}^s$ представлено кінцевою множиною характерних ознак (s – кількість елементів), еталонні описи:

$$Z(j) = \{z_i^j\}_{i=1}^{s_i}, j = 1, \dots, J,$$

де J – кінцева кількість класів. Характерні ознаки – це числовий вектор фіксованої розмірності z_i , $z_i^j \in R^n$, визначений в окремій точці зображення. Критерієм класифікації виступає кількість характерних ознак, віднесених до $Z(j)$. Механізм оцінки пов'язаний з голосуванням $z_i \in Z$. Опис Z буде віднесений до класу, який отримав більшість голосів.

Для фіксованої бази описів $Z(j)$ ймовірність $p(z, j)$ можна обчислити через часткові значення двох мір, які виражають кількість сприятливих випадків до їх загального числа:

$$p(z, j) = \frac{f(z, j)}{\sum_{i=1}^J f(z, i)}, \quad (1)$$

де $f(z, j) \geq 0$ – невід'ємна функція, яка відображає величину відповідності z класу j . Для співставлення необхідно вирахувати значення функції $f(z, j)$:

$$f(z, j) = \exp[\lambda(z, Z(j))], \quad (2)$$

де $\lambda(z, Z(j)) = \min_{z^j \in Z(j)} \chi_z(z, z^j)$ – відстань від z до $Z(j)$.

Для класифікації зображень використовується відстань між гістограмами зображень, отримана за допомогою обчислення методом локальних бінарних шаблонів. На основі даних відстаней і голосування (1) будується статистична вибірка (3), (4), яка дозволяє оцінити події прояву об'єкта (D_0), а також умови $(b_j, j = \overline{1, n})$:

$$p_0 = p(D_0), \quad (3)$$

$$p_j = p\left(\frac{D_0}{b_j}\right), j = \overline{1, n}. \quad (4)$$

За відомими ймовірностями проявлення (дії) системи (3), (4) розраховується її визначеність (5) стосовно цих проявів:

$$d_j = \begin{cases} +0,5 \sqrt{\frac{1-p}{p} + \frac{p}{1-p}} - 2, & \text{при } 0,5 \leq p < 1, \\ -0,5 \sqrt{\frac{1-p}{p} + \frac{p}{1-p}} - 2, & \text{при } 0 < p < 0,5. \end{cases} \quad (5)$$

З визначеності системи (5) вираховується її інформованість:

$$i_j = \sqrt{d^2 + 1}. \quad (6)$$

Потім обчислюється сумарний, по всім діям на систему, приріст визначеності дії системи. При цьому використовується інтроформаційне представлення зміни імпульсів об'єктів:

$$\Delta d^P = i_0 \sum_{j=1}^n d_j - d_0 \sum_{j=1}^n i_j, \quad (7)$$

або інтроформаційне представлення зміни кінетичної енергії:

$$\Delta d^E = \begin{cases} +\sqrt{\frac{\alpha^2}{2} + \sqrt{\frac{\alpha^4}{4} + \alpha^2}}, & \text{при } \alpha \geq 0, \\ -\sqrt{\frac{\alpha^2}{2} + \sqrt{\frac{\alpha^4}{4} + \alpha^2}}, & \text{при } \alpha < 0, \end{cases} \quad (8)$$

де

$$\alpha = \sum_{j=1}^n \left[\operatorname{sgn}(d_j i_0 - d_0 i_j) \cdot \frac{(d_j i_0 - d_0 i_j)^2}{\sqrt{(d_j i_0 - d_0 i_j)^2 + 1}} \right]. \quad (9)$$

Імпульсу відповідає величина впливу фрагменту зображення на реакцію. На основі інтроформаційного представлення зміни імпульсу об'єктів (7) або інтроформаційного представлення зміни кінетичної енергії (8) вираховується приріст інформованості системи:

$$\Delta i = \sqrt{\Delta d^2 + 1}. \quad (10)$$

Потім вираховується нова визначеність дії системи:

$$d_\Sigma = \Delta d \cdot i_0 + d_0 \cdot \Delta i, \quad (11)$$

та нова інформованість дії системи:

$$i_\Sigma = \sqrt{d_\Sigma^2 + 1}. \quad (12)$$

Відповідно до фізичних законів та на основі отриманих нових значень визначеності (11) та інформованості (12) вираховується ймовірність дії системи:

$$p_\Sigma = 0,5 + \frac{d_\Sigma}{2i_\Sigma}. \quad (13)$$

Отримані ймовірності (13) в подальшому необхідно узагальнити:

$$\overline{s(r_i)} = \frac{\sum_{j \neq i} \left[(p(r_i) + p(r_j)) \cdot p_{r_i r_j}(r_i) \right]}{2 \cdot \frac{C_m^2}{m}} = \frac{\sum_{j \neq i} \left[(p(r_i) + p(r_j)) \cdot p_{r_i r_j}(r_i) \right]}{m-1}, \quad (14)$$

де m – кількість можливих рішень; $\overline{s(r_i)}$ – оцінка ймовірності вибору рішення r_i ; $p(r_i)$ – безумовна ймовірність вибору рішення r_i ; $p(r_j)$ – безумовна ймовірність вибору рішення r_j ; $p_{r_i r_j}(r_i)$ – сумісна умовна ймовірність вибору рішення r_i при альтернативному рішенні r_j .

У загальному випадку побудова моделі для конкретного об'єкта вимагає за результатами вимірювань вхідного і вихідного сигналів віднесення даного

об'єкта до певного класу об'єктів. В даній роботі представлена узагальнена модель комбінованої моделі розпізнавання зображень, в якій процес обробки зображень виглядає як потік, що послідовно проходить етапи попередньої обробки, опису зображення, зіставлення та класифікації. Комбінована модель передбачає, що перехід від однієї фази обробки зображення до іншої відбувається тільки після повного і успішного завершення попередньої фази, переходів назад, вперед або перекриття фаз – не відбувається.

6. Результати досліджень

Апробація комбінованої моделі розпізнавання зображень проводилась на чотирьох різних наборах даних:

– набір зображень малих об'єктів ALOI (Amsterdam Library of Object Images) [11]. В даному наборі знаходиться 1000 класів зображень малих об'єктів по 24 екземпляри на клас. В кожному екземплярі класу систематично змінюються: кут огляду, кут освітлення та колір підсвічування. Всі знімки мають фіксований розмір 192x144 пікселі у відтінках сірого в форматі PNG;

– набір зображень рукописних символів сформований T de Campos в Microsoft Research India (The Chars 74k dataset) [12]. В даному наборі знаходиться 62 класи символів (0–9, A–Z, a–z) по 55 екземплярів на клас. Образи символів не мають фіксованих розмірів і нецентровані відносно зображення. Всі зображення строго в чорно-білій палітрі, розміром 1200x900 пікселів в форматі PNG;

– набір зображень облич Кембриджської лабораторії AT&T (ORL Database of Faces) [13]. В даному наборі знаходяться зображення облич 40 осіб по 10 екземплярів на особу. Знімки зроблені в різний час, зі зміною освітлення, міміки та деталей. Всі знімки зроблені на темному однорідному фоні з суб'єктами у строго вертикальній фронтальній позиції. Всі файли мають фіксований розмір 92x112 пікселів у відтінках сірого в форматі PGM;

– набір зображень бази CDI Set (Celebrities Data Images) [14]. В даному наборі було використано 20 відмінних осіб по 20 зображень на особу. Знімки були зроблені в різний час, зі зміною освітлення, фону, кута зйомки, деталей, міміки та положення голови. Всі файли мають різні розміри, виконані у поліхромному забарвленні, формат знімків – JPG.

Дослідження проводились на мобільних пристроях Apple, всі пристрої працювали на базі мобільної операційної системи iOS version 7.1.2, з використанням SDK XCode 5.1.1 stable release 5B1008. Для дослідження достовірності розпізнавання зображень запропонованою комбінованою моделлю було проведено декілька експериментів на наборах даних, які описані вище. В даних експериментах визначались залежності достовірності розпізнавання від налаштувань методу до опису зображень та його ключових фрагментів, залежності достовірності розпізнавання від об'єму вхідних даних і залежності часових витрат від налаштувань методу до опису зображень. Відповідно були підготовлені множини вхідних даних. Мінімальний розмір зображень, що використовувався в експериментах був 92×112 пікселів, максимальний 2000×3000 пікселів.

В ході досліджень було встановлено, що запропонована комбінована модель розпізнавання зображень інтроформаційним методом на основі локальних бінарних шаблонів з голосуванням може використовуватися в умовах наближених до умов реального часу на мобільних пристроях. Такий підхід показав високу достовірність розпізнавання (від 60 % до 95 % в залежності від набору вхідних даних) з можливістю керування достовірністю розпізнавання та часовими втратами за рахунок налаштування опису об'єктів. В порівнянні з існуючими аналогами запропонований підхід показав переважно кращі часові показники навчання та розпізнавання на великих об'ємах вхідних даних (табл. 2).

Таблиця 2

Середня достовірність розпізнавання

№	Вхідні дані	Середній відсоток класифікації за алгоритмами			
		LBP	PCA	LDA	Комбінована модель
0	Зображення предметів (ALOI)	70,861	98,154	98,702	99,429
1	ALOI + вирівнювання гістограм (h. e.)	44,607	90,948	98,769	98,722
2	Рукописні символи (цифри)	33,411	84,968	80,13	84,475
3	Рукописні символи (великі літери)	29,85	84,754	68,157	80,126
4	Рукописні символи (малі літери)	20,812	73,872	49,905	64,342
5	Рукописні символи (вся вибірка)	19,167	64,166	28,267	54,258
6	Обличчя бази зображень AT&T	85,753	93,973	91,665	94,564
7	AT&T + h. e.	84,47	87,449	90,557	91,372
8	Особи бази зображень CDI, об'єм навчальної вибірки 5 зразків на клас	34,675	26,918	29,916	34,833
9	CDI (5) + h. e.	37,3	26,018	31,633	33,463
10	CDI (5) + геометричне вирівнювання зображень (g. e.)	39,341	28,959	42,085	45,425
11	CDI (5) + h. e. + g. e.	42,713	28,836	48,265	48,075
12	CDI, об'єм навчальної вибірки 10 зразків на клас	40,009	30,61	40,828	45,635
13	CDI (10) + h. e.	41,025	33,721	44,264	47,204
14	CDI (10) + g. e.	47,996	37,307	55,969	58,673
15	CDI (10) + h. e. + g. e.	46,467	41,11	61,01	60,175

Середня достовірність класифікації запропонованим підходом склала 71,3 % (LBP – 59,9 %; PCA – 65,2 %; LDA – 65,6 %). Це в поєднанні з часовими втратами навчання та розпізнавання робить даний підхід прийнятним для застосування в системах наближених до умов реального часу на мобільних пристроях.

7. SWOT-аналіз результатів дослідження

Strengths. Як показали дослідження, запропонована комбінована модель розпізнавання зображень дозволяє проводити не лише обробку зображень, але також навчання та перенавчання в режимі близькому до реального часу при обмежених ресурсах мобільних пристроїв краще відомих підходів до розпізнавання зображень. Запропонована комбінована модель також дозволяє керувати достовірністю та часовими втратами процесу розпізнавання за рахунок редукції простору ознак.

Weaknesses. Найвищу достовірність розпізнавання запропонована модель показує лише у випадках, коли об'єкт розпізнавання «відомий» (рукописний символ, обличчя і т. ін.). В межах даної роботи проводилось дослідження та розробка тільки моделі розпізнавання. Питання класифікації та попередньої обробки в межах даної роботи не розглядалися.

Opportunities. Перспективність запропонованої комбінованої моделі забезпечена низькою обчислювальною складністю та високою достовірністю розпізнавання. Це дозволяє не лише працювати на пристроях з обмеженими апаратними ресурсами (такі як мобільні пристрої), але й будувати складні обчислювальні комплекси для обробки Big Data, які є досить перспективним напрямком сьогодення. Оскільки елементи теорії несилової взаємодії, які лежать в основі запропонованої комбінованої моделі, добре зарекомендували себе в інших задачах обробки даних, то створену комбіновану модель можна адаптувати для інших напрямків розпізнавання: звуку, поведінки і т. ін. Також її можна використовувати в задачах прогнозування різного роду подій та наслідків.

Threats. Одним з основних факторів, що впливає на успішність розпізнавання з високою достовірністю, є умови збору інформації та способи підготовки і обробки даних. Причому достовірність розпізнавання варіюється в значному діапазоні в залежності від умов збору інформації. Результати даного дослідження були отримані з наборів даних, отриманих в умовах близьких до ідеальних. Тому при роботі з реальними даними можливо треба буде додатково застосовувати методи та підходи попередньої обробки зображення, які направлені на:

- компенсацію спотворень, викликаних вібрацією камери;
- зменшення шумів, які вносяться камерою і залежать від типу та розміру її сенсору, а також умов зйомки;
- вирівнювання яскравості зображення і т. ін.

8. Висновки

1. Дослідження показують, що теорія несилової взаємодії на відміну від існуючих методів класифікації володіє такими ключовими особливостями:

- призначена для обробки великої кількості даних;
- виділяє корисну інформацію з довільного потоку;
- дозволяє природно додавати нові об'єкти;
- терпима до помилок та дозволяє швидко перепрограмувати поведінку системи.

2. На основі теорії несилової взаємодії з описом об'єктів за допомогою локальних бінарних шаблонів і підходу з голосуванням було сформовано нову модель розпізнавання графічних об'єктів у програмних системах. Особливість

даної моделі полягає в поєднанні нелінійної структури, яка притаманна нейромережевим структурам (теорія несилової взаємодії) та малої обчислювальної складності. Це дозволяє програмним системам, які побудовані на основі даної моделі, працювати в режимах близьких до реального часу при обмежених ресурсах мобільних пристроїв.

3. Модель, побудована на основі запропонованого методу розпізнавання зображень, дозволяє керувати достовірністю розпізнавання. Це досягається завдяки налаштуванню кількості секторів гістограм локальних бінарних шаблонів, які використовуються при описі зображень, та кількості фрагментів зображення, які використовуються на етапі класифікації інтроформаційним підходом. Значною мірою на час класифікації впливає кількість фрагментів зображення, оскільки, в такому випадку, потрібно враховувати попарно співставлення суміщення дій системи в кожному з можливих напрямків.

Література

1. Wagner, P. (2011). *Principal Component Analysis and Linear Discriminant Analysis with GNU Octave*. Available at: https://www.bytefish.de/blog/pca_lda_with_gnu_octave/
2. Samal, A., Iyengar, P. A. (1992). Automatic recognition and analysis of human faces and facial expressions. *Pattern Recognition*, 25 (1), 65–77. doi: [http://doi.org/10.1016/0031-3203\(92\)90007-6](http://doi.org/10.1016/0031-3203(92)90007-6)
3. Ojala, T., Pietikainen, M., Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (7), 971–987. doi: <http://doi.org/10.1109/tpami.2002.1017623>
4. Томашевич, Н. С. (2007). Статистические методы выделения признаков. *Нейрокомпьютеры в прикладных задачах обработки изображений*, 121–128.
5. Волченков, М. П., Самоненко, И. Ю. (2005). Об автоматическом распознавании лиц. *Интеллектуальные системы*, 9, 135–156.
6. Simard, P. Y., Steinkraus, D., Platt, J. C. (2003). Best practices for convolutional neural networks applied to visual document analysis. *12th International Conference on Document Analysis and Recognition*, 2, 958. doi: <http://doi.org/10.1109/icdar.2003.1227801>
7. LeCun, Y., Huang, F.-J., Bottou, L. (2004). Learning methods for generic objects recognition with invariance to pose and lighting. Los Alamitos. *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR'04)*, 2, 97–104. doi: <http://doi.org/10.1109/cvpr.2004.1315150>
8. Митрофанов, С. А. (2015). Сравнение эффективности различных методов интеллектуального анализа данных в задачах распознавания изображений. *Инновационная наука*, 12 (2), 1–2.
9. Мамонтов, Д. Ю., Карасева, Т. С. (2015). Решение задач финансового анализа с помощью интеллектуальных информационных технологий. *ИТСuT*, 1–2.
10. Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1106–1114.

11. *Amsterdam Library of Object Images (ALOI)*. Available at: <http://aloi.science.uva.nl/>
12. *The Chars74K dataset*. Available at: <http://www.ee.surrey.ac.uk/CVSSP/demos/chars74k/>
13. *The Database of Faces*. Available at: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
14. *Celebrities Data Images Set for Computer Vision*. Available at: <http://cdiset.blogspot.com/>

Не является переносимым