

УДК 004

В.Юрта, магістр гр. КН-23М,

Центральноукраїнський національний технічний університет

ДОСЛІДЖЕННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ ІНТЕЛЕКТУАЛЬНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ МЕТОДОМ ПОТЕНЦІЙНИХ ФУНКЦІЙ

У статті розроблено програмне забезпечення, яке призначено для системи інтелектуального розпізнавання образів методом потенційних функцій. Метою розробки є дослідження та програмна реалізація системи інтелектуального розпізнавання образів методом потенційних функцій. Об'єктом дослідження є процес інтелектуального розпізнавання образів методом потенційних функцій. Предметом дослідження є методи інтелектуального розпізнавання образів методом потенційних функцій. Методи дослідження базуються на методах теорії розпізнавання образів, методах математичної статистики, методах розробки програмного забезпечення. Результат роботи – програмна реалізація системи інтелектуального розпізнавання образів методом потенційних функцій. В процесі роботи над програмною моделлю виконано аналіз існуючих апаратних та програмних засобів. В повній мірі описані всі компоненти розробленого програмного забезпечення.

Постановка проблеми. Формально, розпізнавання є не що інше, як ухвалення рішення про приналежність розпізнаваного об'єкта або його стану до певного класу (класів) [2].

Із цього випливає внутрішній і органічний зв'язок методів розпізнавання образів і прийняття рішень. Аналітичний огляд дозволяє зробити вивід, що найбільш глибока основа цього зв'язку полягає в тому, що й розпізнавання образів, і прийняття рішень є насамперед зняттям невизначеності.

Розпізнавання знімає невизначеність у питанні про те, до якого класу відноситься розпізнаваний об'єкт. Якщо до розпізнавання існувала невизначеність у питанні про те, до якого класу відноситься розпізнаваний об'єкт або його стан, то в результаті розпізнавання ця невизначеність зменшується, причому можливо й до нуля (коли об'єкт ідентифікується однозначно).

Ухвалення рішення (вибір) також знімає невизначеність у питанні про те, яке з можливих рішень буде прийнято, якщо існувало кілька альтернативних варіантів рішень і приймається одне з них.

Для строгого дослідження процесів зняття невизначеності оптимальним є застосування апарата теорії інформації, що як би спеціально створена для цієї мети.

Із цього безпосередньо витікає можливість застосування методів теорії інформації для рішення завдань розпізнавання й прийняття рішень.

Таким чином, теорія інформації може розглядатися як єдина основа методів розпізнавання образів і прийняття рішень.

Аналіз останніх досліджень і публікацій. При аналізі останніх досліджень і публікацій [1-10] було виявлено певні прогалини у забезпеченні системи інтелектуального розпізнавання образів методом потенційних функцій.

Мета й завдання дослідження. Метою роботи є дослідження та програмна реалізація системи інтелектуального розпізнавання образів методом потенційних функцій.

Для досягнення поставленої мети визначена програма дослідження, що складається з наступних завдань:

- Огляд існуючих систем інтелектуального розпізнавання образів методом потенційних функцій.
- Дослідження системи інтелектуального розпізнавання образів методом потенційних функцій.
- Програмна реалізація системи інтелектуального розпізнавання образів методом потенційних функцій.

Об'єктом дослідження є процес інтелектуального розпізнавання образів методом потенційних функцій.

Предметом дослідження є методи інтелектуального розпізнавання образів методом потенційних функцій.

Методи дослідження базуються на методах теорії розпізнавання образів, методах математичної статистики, методах розробки програмного забезпечення.

Виклад основного матеріалу. Розпізнавання зображень – це процес, за допомогою якого система аналізує зображення та визначає, що на ньому зображено. Це може бути:

- Ідентифікація об'єкта (наприклад, «на цьому фото зображений кіт»).
- Розпізнавання обличчя.
- Розпізнавання тексту (OCR).
- Класифікація сцени (місто, природа, інтер'єр тощо).

Отже, це специфічне завдання в межах ширшої галузі комп'ютерного зору. Воно включає всі технології та методи, які дозволяють комп'ютеру «бачити» та інтерпретувати візуальні дані, як людина:

- Розпізнавання зображень та об'єктів.
- Відстеження руху.
- Сегментація (визначення меж об'єкта).
- Глибинний аналіз (3D).
- Аналіз відео в режимі реального часу та багато іншого.

Тобто, розпізнавання фотографій за допомогою штучного інтелекту – це як мозок, який каже: «На фотографії дівчина». А комп'ютерний зір – це весь процес: око бачить, мозок обробляє, аналізує та приймає рішення.

Як працює розпізнавання зображень?

Все починається з надходження зображення в систему. Система розпізнавання зображень працює за принципом «від простого до складного»:

1. Попередня обробка. Покращується якість зображення, видаляються шуми та зайві деталі, нормалізуються яскравість та контрастність.
2. Вилучення ознак. Комп'ютер шукає важливі деталі: краї, кути, текстури, кольори або контури об'єктів, щоб визначити, що може бути релевантним на зображенні.
3. Класифікація. На основі витягнутих ознак система використовує алгоритми (наприклад, нейронні мережі або SVM), щоб визначити, до якого класу належить об'єкт: кішка, собака, людина тощо.

Після розпізнавання майже будь-яка програма для класифікації зображень може відображати мітку на екрані, передавати дані іншій програмі або запускати автоматичну дію (наприклад, відкриття воріт під час розпізнавання номерного знака автомобіля).

Методи розпізнавання зображень

1. Класичні методи (до ери нейронних мереж: приблизно до 2012 року)

Раніше програми для розпізнавання зображень покладалися на ручне вилучення ознак.

- Метод опорних векторів (SVM). Зображення розділяється на блоки, обчислюються гістограми кольору/текстури. На основі цього SVM класифікує об'єкти (наприклад, «кіт» проти «собака»).
- SIFT, SURF, HOG – вилучення ознак. Використовуються для розпізнавання об'єктів з різними масштабами та перспективами. Добре працюють у завданнях, що вимагають геометричної стійкості (наприклад, розпізнавання логотипів або орієнтирів).

– KNN (k-найближчих сусідів). Порівнює нове зображення з відомими. Чим ближче воно за характеристиками до своїх «сусідів», тим вища ймовірність того, що воно належить до того ж класу. Простий у реалізації, але не масштабований.

2. Методи на основі нейронних мереж (сучасний підхід)

– ЗГН (згорткові нейронні мережі) – золотий стандарт комп'ютерного зору. Використовується в більшості сучасних систем (Google Lens, камери смартфонів, системи спостереження). Програми для ідентифікації зображень на основі ЗГН «навчаються» автоматично виявляти важливі ознаки. Мережеві шари поступово виявляють прості форми (лінії, кути) до складних об'єктів (обличчя, предмети).

Популярні архітектури для створення програмного забезпечення для розпізнавання зображень:

- LeNet – перша успішна CNN (1998), розпізнавання цифр.
- AlexNet (2012) – спричинив бум штучного інтелекту.
- VGG, ResNet, Inception – глибші, точніші та швидші мережі.
- RCNN / YOLO / SSD (для виявлення об'єктів). Визначають не лише те, що зображено, але й де знаходиться об'єкт на зображенні. Використовуються в системах спостереження, автономних транспортних засобах, AR-додатках.
- GAN (Генеративно-змагальні мережі). Використовуються не стільки для розпізнавання, скільки для генерації зображень. Але їх можна використовувати для покращення якості, "врізання" або розуміння складних об'єктів.

Специфічні для галузі варіанти використання

Ми буквально оточені програмним забезпеченням для розпізнавання зображень. Дістаєте телефон? Привіт, Face ID або Android Face Unlock. Відкрийте Snapchat, щоб застосувати фільтр: камера розпізнає ваше обличчя, очі, губи в режимі реального часу (технологія FaceMesh). Відвідуєте Китай і не можете зрозуміти назву вулиці? «Фоторежим» Google Translate на допомогу. Навіть штрих-коди та QR-коди – це класичне мобільне розпізнавання зображень.

Але це повсякденні приклади. Давайте розглянемо випадки використання розпізнавання зображень у більш «серйозних» сферах.

Розпізнавання зображень у медичній діагностиці

Медицина є однією з найчутливіших і найважливіших галузей застосування. Нейронні мережі навчаються виявляти пухлини, аномалії та інші патології на МРТ, рентгенівських знімках та ультразвукових дослідженнях. Спеціальне програмне забезпечення для розпізнавання зображень допомагає виявити ризик меланому на ранній стадії, аналізуючи звичайну фотографію родимки. У проектах Google Health штучний інтелект визначав рак молочної залози з вищою точністю, ніж досвідчені рентгенологи.

Застосування в сільському господарстві та моніторингу сільськогосподарських культур

Дрони та супутники створюють зображення сільськогосподарських угідь, які потім обробляються нейронними мережами. Ці системи можуть виявляти ділянки з поганим ростом врожаю, ознаками хвороб, посухи або зараження шкідниками. Раніше такі проблеми виявляли візуально, часто занадто пізно. Сьогодні розпізнавання зображень за допомогою штучного інтелекту може виявити проблеми ще до того, як люди їх помітять.

Крім того, штучний інтелект допомагає контролювати стадії росту рослин, прогнозувати врожайність, контролювати регулярність поливу та навіть автоматично сортувати фрукти та овочі за стиглістю та якістю. Наприклад, камери на сільськогосподарських об'єктах визначають, які яблука потрібно зібрати сьогодні, а які мають дозріти далі.

Промисловий контроль якості за допомогою розпізнавання зображень

Камери зі штучним інтелектом та алгоритмами розпізнавання зображень контролюють якість продукції: виявляють тріщини, відхилення форми, помилки маркування. Наприклад, на заводах BMW такі системи перевіряють відповідність компонентів вимогам

перед складанням. На лініях розливу Coca-Cola камери виявляють навіть невеликі дефекти пляшок і автоматично їх бракують.

Розпізнавання зображень у логістиці та управлінні запасами

Сучасне програмне забезпечення для розпізнавання зображень незамінне в цій галузі. На складах та в розподільчих центрах камери на базі штучного інтелекту розпізнають штрих-коди, QR-коди та навіть текст на упаковці – у режимі реального часу та без участі людини. Це дозволяє швидше сортувати, відстежувати товари та перевіряти цілісність упаковки. Камери виявляють пошкодження вантажу, неправильне маркування або порушення умов зберігання (наприклад, відкриту коробку або відсутню пломбу).

У транспортній логістиці розширене розпізнавання зображень використовується для розпізнавання номерних знаків вантажівок, контролю в'їзду та виїзду з об'єктів, а також автоматичного реєстрування часу прибуття та відправлення. Це усуває людські помилки та підвищує точність обліку – ключову перевагу сучасних...розробка логістичного програмного забезпечення працює на базі розпізнавання зображень.

Під час доставки кур'єрські служби використовують програми для розпізнавання зображень, щоб перевірити адреси, документи та навіть особу одержувача. Деякі компанії використовують камери у фургонах для моніторингу дорожніх умов, поведінки водіїв та запобігання аваріям.

Розпізнавання зображень у системах безпеки та відеоспостереження

У сфері безпеки та правоохоронної діяльності програми для зіставлення зображень допомагають знаходити підозрюваних у натовпах, виявляти забуті або небезпечні предмети в громадських місцях та аналізувати поведінку на масових заходах.

Наприклад, лондонська поліція використовує розпізнавання обличчя у реальному часі під час демонстрацій та концертів. В аеропортах США біометричні камери швидко зіставляють обличчя пасажирів з їхніми паспортними даними без контакту чи ліній.

Інструменти розпізнавання зображень для роздрібної торгівлі та електронної комерції

У роздрібній торгівлі та електронній комерції сервіси розпізнавання зображень використовуються переважно для автоматизації та оптимізації. Наприклад, системи допомагають контролювати вітрини товарів та відповідність планогамі, швидко проводити інвентаризацію та контролювати наявність товарів на полицях. Розпізнаючи цінніки та штрих-коди, магазини можуть автоматично оновлювати інформацію про товар, зменшуючи кількість помилок та пришвидшуючи роботу персоналу.

По-друге, технології візуального пошуку залишаються сильною конкурентною перевагою та, ймовірно, стануть стандартом. Додатки для візуального розпізнавання дозволяють покупцям знаходити товари за фотографіями, а функції віртуальної примірки роблять онлайн-шопінг зручнішим та персоналізованішим.

Аналіз поведінки клієнтів за допомогою комп'ютерного зору допомагає оптимізувати розміщення товару та покращити маркетингові стратегії. Зрештою, ці технології роблять роздрібну торгівлю ефективнішою, а покупки – простішими та приємнішими.

Найкраще програмне забезпечення для розпізнавання зображень у 2025 році

У 2025 році ринок технологій розпізнавання зображень пропонує широкий спектр рішень: від потужних хмарних платформ до гнучких бібліотек з відкритим кодом. Найкраще програмне забезпечення визначається не лише точністю моделі, але й тим, наскільки легко його можна інтегрувати в бізнес-процеси та адаптувати до конкретних завдань. Ось лідери, яких варто відзначити:

1. Google Cloud Vision AI – один із найпопулярніших інструментів, потужний механізм розпізнавання зображень. Підтримує розпізнавання об'єктів, тексту, облич, логотипів та класифікацію зображень. Відомий масштабованістю та високою точністю. Ідеально підходить для аналізу великих обсягів даних та легко інтегрується через API.

2. Amazon Rekognition – хмарне рішення від AWS. Ідеально підходить для розробників, яким потрібна надійна та гнучка система. Це більше, ніж просто модуль для

програм розпізнавання фотографій. Rekognition може розпізнавати обличчя, емоції, об'єкти, сцени та текст. Містить вбудовані інструменти для моніторингу в режимі реального часу та аналізу відео. Широко використовується в безпеці та роздрібній торгівлі.

3. Комп'ютерний зір Microsoft Azure пропонує широкий набір функцій: оптичне розпізнавання символів (OCR), розширений аналіз зображень, описи зображень, метадані та просторове розуміння. Azure легко масштабується та підтримує готові до використання сценарії для застосувань в охороні здоров'я, транспорті та виробництві.

4. OpenCV + TensorFlow (або PyTorch). Для команд, які віддають перевагу повній гнучкості та контролю, найкращим вибором є стек з відкритим кодом. OpenCV обробляє базові зображення, тоді як TensorFlow/PyTorch дозволяють навчати та розгортати нейронні мережі. Це вимагає більшої експертизи, але пропонує майже необмежені можливості налаштування.

5. Viso Suite Повнофункціональна no-code/low-code платформа для створення систем комп'ютерного зору. Дозволяє компаніям швидко запускати рішення без глибоких знань програмування. Підтримує всі ключові функції: від штучного інтелекту для розпізнавання зображень до аналізу поведінки в режимі реального часу.

6. Clarifai – універсальна платформа, що пропонує як готове до використання рішення для ідентифікації об'єктів у додатку, так і інструменти для навчання власних моделей. Працює із зображеннями та відео, може використовуватися для візуального пошуку, фільтрації контенту, контролю якості тощо. Особливо цінується за зручний інтерфейс та добре розроблений API.

7. IBM Watson Visual Recognition. Підходить для великих підприємств та проектів з високими вимогами до надійності та аналітики. Здатний класифікувати зображення, виявляти аномалії, аналізувати дефекти та відстежувати тенденції на основі візуальних даних.

Вибір між бібліотеками з відкритим кодом та комерційними платформами є одним із ключових рішень під час початку будь-якого проекту комп'ютерного зору. Відкритий код означає свободу, гнучкість та контроль. Комерційні рішення пропонують швидкість, простоту та підтримку. Але якщо ви хочете все це, єдина відповідь – це розробка програмного забезпечення для розпізнавання зображень на замовлення.

Нові тенденції в розпізнаванні зображень

У 2025 році програми для виявлення об'єктів виходять за рамки серверних систем. Все більше рішень переходять на периферійні обчислення. Це означає, що фотографії та відео обробляються безпосередньо на смартфонах, камерах, дронах або роботах без необхідності надсилати дані в хмару. Це прискорює аналіз, економить пропускну здатність і особливо важливо для конфіденційних завдань: від охорони здоров'я до безпеки.

Водночас, ШІ стає глибоко персоналізованим. Алгоритми все більше адаптуються до конкретних користувачів або компаній. Вони враховують поведінку, уподобання, контекст та історію взаємодії.

У сфері споживчих додатків набирають обертів фото- та відеосервіси на базі штучного інтелекту. Вони можуть аналізувати зовнішність, емоції, стиль, вік і навіть настрій людини в режимі реального часу.

Проблеми та міркування

Незважаючи на вражаючий прогрес, технології розпізнавання зображень все ще стикаються з обмеженнями. Системи штучного інтелекту можуть помилятися за поганого освітлення, незвичайних ракурсів, зображень низької якості або коли візуальних даних недостатньо для точної класифікації. Розпізнавання облич за різними етнічними ознаками залишається особливо складним, що впливає на точність та викликає занепокоєння громадськості.

Етичні питання та конфіденційність залишаються на передньому плані. Використання візуального штучного інтелекту в громадських місцях, магазинах або онлайн-додатках часто

відбувається без явної згоди користувачів. Вкрай важливо дотримуватися законів про захист даних (таких як GDPR) та забезпечувати прозорість: хто, де і як використовує візуальні дані.

Зрештою, ключовим завданням є забезпечення точності та об'єктивності моделі. Алгоритми навчаються на наборах даних, і якщо ці набори даних містять упередженість або недостатню представленість певних груп, результати будуть спотвореними. Саме тому розробники все частіше впроваджують механізми аудиту штучного інтелекту, а компанії інвестують не лише в технології, а й в етику їх використання.

Глобальний ринок розпізнавання зображень стрімко зростає. За даними Fortune Business Insights, він досягне 58,56 мільярда доларів у 2025 році та зросте до 163,75 мільярда доларів до 2032 року зі середньорічним темпом зростання 15,8%. Це зростання зумовлене впровадженням штучного інтелекту в різних галузях, включаючи охорону здоров'я, роздрібну торгівлю та виробництво.

Розпізнавання образів

Розпізнавання образів передбачає штучне розпізнавання навколишніх об'єктів за допомогою машинного навчання та алгоритмів розпізнавання образів. Іншими словами, це процес виявлення тенденцій у заданому шаблоні. У просторі машинного навчання (ML) розпізнавання шаблонів демонструє використання надійних алгоритмів для виявлення закономірностей у заданому наборі даних. Така техніка дозволяє вирішити різні проблеми, такі як розпізнавання міміки, комп'ютерного зору, розпізнавання мови та багато інших.

Відповідність шаблону

Зіставлення шаблонів передбачає процес зіставлення функцій цілі з попередньо визначеним шаблоном та ідентифікацію цілі через проксі. Це техніка, яка використовується для пошуку підзображення цільового зображення, яке нагадує шаблонне зображення.

Ця техніка використовується для визначення подібності між двома подібними об'єктами. Коли шаблон збігається, він зберігається в шаблонах, і шаблону надається гнучкість для скалярних і обертальних варіацій. Можливість цього методу залежить від уже збережених шаблонів у базі даних.

Нечітка модель

Нечітка логіка – це форма багатозначної логіки, у якій значення істинності змінних лежать між 0 і 1 включно. Він використовується для обробки часткової істини, де значення істинності коливається між повністю істинним і повністю хибним. Відомо, що цей метод розпізнавання образів дає найкращі кінцеві результати порівняно з іншими методами через те, що тут моделюється для невизначених областей і компонентів для розпізнавання.

Гібридна модель

Коли кілька алгоритмів розгортаються разом, цей метод відомий як модель гібридного алгоритму. Наприклад, поєднання нейронних мереж із нечіткими системами чи прихованими марковськими моделями чи поєднання символічного та підсимволічного навчання – це все гібридні стратегії розпізнавання образів. Найкраща частина використання такого підходу полягає в тому, що він використовує переваги всіх задіяних методів і є більш придатним для людського сприйняття.

Методи розпізнавання образів мають ряд переваг. Вони не тільки корисні для аналізу тенденцій і прогнозування, але й для визначення об'єктів на різних відстанях і під різними кутами. Вони можуть бути застосовані до різних областей, включаючи фондовий ринок, галузі медицини, фінансову галузь тощо, щоб робити цінні прогнози. Оскільки вони високоавтоматизовані, вони пропонують ефективні рішення проблем у реальному часі.

Розробка структурної схеми

У роботах М.Айзермана, Е.Бравермана, Л.Розоноера були розроблені і досліджені алгоритми розпізнавання, в основу яких покладені так звані потенційні функції. Пропоновані авторами алгоритми базуються на основній гіпотезі про характер функцій, що розділяють множини, що відповідають різним образам.

Геометрична інтерпретація методу

Для спрощення, але без обмеження спільності, розглянемо завдання розпізнання образів (РО) для двох класів V_1, V_2 .

Будемо вважати, що в просторі вхідних описів кожному вхідному зображенню $S(x)$ відповідає єдина крапка простору X . Допустимо, що класи V_1 й V_2 не перетинаються. Це означає, що в просторі вхідних описів X існує, принаймні, одна функція, що повністю розділяє множини зображень, що належать різним образам. Ця функція $\Phi(x)$ повинна приймати значення $\Phi(x) > 0$ в крапках, що відповідають зображенням образу V_1 й $\Phi(x) < 0$ для $x \in V_2$. У загальному випадку таких поділяючих функцій може бути багато. У процесі навчання системи, що розпізнає, послідовно пред'являють зображення, яким відповідають крапки простору X . При цьому відомо, до якого класу V_1 або V_2 належать пропоновані зображення. Метод потенційних функцій зв'язаний з наступною процедурою. При показі (пред'явленні) у ході навчання деякого зображення $S_k(x)$, якому відповідає крапка x_k в просторі X , з ним зв'язується функція $U(x, x_k)$, задана на всьому просторі X і залежна від x_k як від параметра. Така функція називається потенційною. Навчальній послідовності $S = \{S_1, S_2, \dots, S_k\}$ й крапкам $x_1, x_2, \dots, x_k$ простору X відповідає послідовність потенційних функцій $U(x, x_1), U(x, x_2), \dots, U(x, x_k)$, що використовується для побудови поділяючої функції $\Phi(x)$ за допомогою визначальних правил. Правила формування функцій повинні бути такими, щоб у міру збільшення числа пред'явлених зображень функція $\Phi(x)$ прагнула б до однієї з поділяючих функцій.

Метод потенційних функцій припускає існування в просторі X системи функцій $\varphi_i(x), i = 1, 2, \dots$, що дозволяють для кожної пари поділяючих множин знайти таке N , при якому поділяючу функцію можна було б представити у вигляді:

$$\Phi(x) = \sum_{i=1}^N c_i \varphi_i. \quad (1)$$

Якщо в просторі X існує повна система функцій, то $\varphi_i(x)$ можна вважати її елементами й будь-якою функцією цієї системи може бути представлена у вигляді нескінченного ряду

$$\Phi(x) = \sum_{i=1}^{\infty} c_i \varphi_i. \quad (1a)$$

Однак нам бажано, щоб поділяюча функція (згідно (1)) розкладалася в ряд з кінцевим числом членів N . Із цього треба, що поділяючі функції в просторі X повинна бути досить гладкими й не мати великого числа екстремумів у малій області. Використовуючи (1) можна ввести в розгляд N -мірний простір Z , на яке відображається вихідний простір X . Кожній крапці простору X відноситься у відповідність крапка простору Z відповідно до співвідношення $Z_i = \varphi_i(x), i=1, 2, \dots, N$.

У силу умови (1) поділяюча функція $\Phi(x)$ відбивається в лінійну функцію $\sum_{i=1}^N a_i Z_i$ в просторі Z і має наступну властивість:

$$\Phi(\mathbf{x}) = \sum_{i=1}^N a_i Z_i(\mathbf{x}) > 0, \text{ если}$$

$$\Phi(\mathbf{x}) = \sum_{i=1}^N a_i Z_i(\mathbf{x}) < 0, \text{ если} . \quad (2)$$

У зв'язку з тим, що функції, що розкладаються за системою $\varphi_i(\mathbf{x})$ лінеаризуються в просторі Z , останнє називають зпрямляючим простором.

Алгоритм розпізнавання, заснований на методі потенційних функцій

Як потенційна функція приймається скалярна функція двох векторних аргументів виду

$$\Phi(\mathbf{x}) = \sum_{i=1}^{\infty} \alpha_i^2 \cdot \varphi_i(\mathbf{x}) \cdot \varphi_i(\mathbf{x}^*), \quad (3)$$

де $\varphi_i(\mathbf{x})$ – лінійно-незалежна система функцій; α_i – дійсні числа, відмінні від 0 для всіх i ; $\mathbf{x}^*$ – крапка, що з'являється в ході навчання.

Передбачається, що $\varphi_i(\mathbf{x})$ й $U(\mathbf{x}, \mathbf{x}^*)$ обмежено для $\mathbf{x} \in V_1 \cup V_2$. Нехай у процесі навчання пред'являються зображення $S_1, S_2, \dots, S_k$, яким відповідають крапки $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ в просторі X . Кожна із крапок належить V_1 або V_2 . Будемо вважати безліч V_1 позитивним, а V_2 – негативним. З появою першої крапки $\mathbf{x}_1$ будувється потенційна функція $U_1(\mathbf{x})$, рівна потенціалу, що відповідає крапці $\mathbf{x}_1$, узятому з відповідним знаком.

$$U_1(\mathbf{x}) = \begin{cases} U(\mathbf{x}, \mathbf{x}_1), & \text{если } \mathbf{x} \in V_1 \\ -U(\mathbf{x}, \mathbf{x}_1), & \text{если } \mathbf{x} \in V_2 \end{cases} \quad (4)$$

Далі нехай після i -го пред'явлення побудований потенціал $U_i(\mathbf{x})$.

На наступному $(i+1)$ -м кроці навчання показується крапка $\mathbf{x}_{i+1}$. У результаті можливі 4 випадки:

- $\mathbf{x}_{i+1} \in V_1$; $U_i(\mathbf{x}_{i+1}) > 0$;
 - $\mathbf{x}_{i+1} \in V_2$; $U_i(\mathbf{x}_{i+1}) < 0$;
 - $\mathbf{x}_{i+1} \in V_1$; $U_i(\mathbf{x}_{i+1}) < 0$;
 - $\mathbf{x}_{i+1} \in V_2$; $U_i(\mathbf{x}_{i+1}) > 0$;
- (5)

У випадках 1 і 2 алгоритм правильно класифікує зображення S_{i+1} . У цьому випадку приймається

$$U_{i+1}(\mathbf{x}) = U_i(\mathbf{x}) \quad (6)$$

В 3 і 4 випадках є помилка класифікації й необхідна корекція потенційної функції.

Для випадку 3 приймаємо:

$$U_{i+1}(\mathbf{x}) = U_i(\mathbf{x}) + U(\mathbf{x}, \mathbf{x}_{i+1}) \quad (7)$$

Для випадку 4 приймаємо:

$$U_{i+1}(\mathbf{x}) = U_i(\mathbf{x}) - U(\mathbf{x}, \mathbf{x}_{i+1}) \quad (8)$$

Побудований після i -го кроку потенціал можна записати в такий спосіб:

$$U_i(\mathbf{x}) = \sum_{\mathbf{x}_k, \mathbf{x}_k \in V_1} U(\mathbf{x}, \mathbf{x}_k) + \sum_{\mathbf{x}_k, \mathbf{x}_k \in V_2} (-U(\mathbf{x}, \mathbf{x}_k)) \quad (9)$$

де $\mathbf{x}_{k(-)}$ – крапки, що належать образу V_1 , підстановка яких у попередній потенціал приводила до помилки класифікації; аналогічно M^* .

У роботах авторів було розроблено кілька алгоритмів РО, заснованих на методі потенційних функцій.

Розходження між варіантами алгоритму зводиться в основному до вибору законів корекції поділяючої функції від кроку до кроку.

Перший алгоритм

Будемо вважати, що побудовано функцію $\Phi_i(x)$, а на $(i+1)$ -м кроці пред'явлена т. x_{i+1} , для якої відомо дійсне (необхідне) значення функції $\Phi_{i+1}(x)$. Воно повідомляється вчителем. Тоді функція $\Phi_{i+1}(x)$ будується за наступним правилом:

$$\Phi_{i+1}(x) = \Phi_i(x) + \alpha_{i+1} \cdot \text{sign} [\Phi(x_{i+1}) - \Phi_i(x_{i+1})] \cdot U(x, x_{i+1}), \quad (10)$$

де:

Φ – дійсне значення поділяючої функції в т. x_{i+1} ;

α_{i+1} – будь-яка послідовність чисел, що задовольняє наступним умовам:

$$\alpha_i \rightarrow 0; \quad \sum_{i=1}^{\infty} \alpha_i = \infty; \quad \sum_{i=1}^{\infty} \alpha_i^2 < \infty$$

Наприклад, $\alpha_i = c/i$, $c = \text{const}$.

Збіжність алгоритму обґрунтовується наступною теоремою.

Теорема 1. Нехай x_i – послідовність незалежних випадкових крапок з X , а $P(x_i)$ – послідовність ймовірностей їхньої появи.

Нехай $\Phi(x)$ – функція, що задовольняє умові (1)

$$\Phi(x) = \sum_{i=1}^N c_i \varphi_i \quad (10)$$

Тоді послідовність функцій $\Phi_i(x)$ ($i=1,2,\dots$), обумовлених рекурентними співвідношеннями (10) при $i \rightarrow \infty$ задовольняє умові:

$$P \left\{ \lim_{i \rightarrow \infty} \int |\Phi(x) - \Phi_i(x)| \cdot P(x) dx = 0 \right\} = 1 \quad (11)$$

Другий алгоритм

Як і колись, думаємо $\Phi_0(x) \equiv 0$.

$$\Phi_{i+1}(x) = \Phi_i(x) + \frac{1}{\lambda} \cdot [\Phi(x_{i+1}) - \Phi_i(x_{i+1})], \quad (12)$$

де λ – довільна позитивна константа, що задовольняє умові:

$$\lambda = \frac{1}{\gamma} \max U(x, x_{i+1}), \quad (13)$$

де $\Phi(x_{i+1})$ – дійсне значення поділяючої функції на т. x_{i+1} .

Збіжність цього алгоритму підтверджується наступною теоремою:

Теорема 2. Нехай виконуються ті ж умови, що й у теоремі 1. Тоді послідовність функцій $\Phi_i(x)$ ($i=1,2,\dots$), обумовлених рекурентним співвідношенням (12) при $i \rightarrow \infty$ задовольняє умовам:

$$P \left\{ \lim_{i \rightarrow \infty} \int |\Phi(x) - \Phi_i(x)|^2 \cdot P(x) dx = 0 \right\} = 1$$

Наведені алгоритми можна використовувати й для послідовних наближень коефіцієнтів c_j у поданні поділяючої функції

$$\Phi(x) = \sum_{i=1}^N c_i \varphi_i$$

за формулами

$$C_{j,i+1} = C_{j,i} + \alpha_{i+1} \cdot \text{sign} \left[\Phi(x_{i+1}) - \sum_{i=1}^N c_i \varphi_i(x_{i+1}) \right], j=1,2,\dots,n \quad (15)$$

$$C_{j,i+1} = C_{j,i} + \frac{1}{\lambda} \cdot \left[\Phi(x_{i+1}) - \sum_{i=1}^N c_i \varphi_i(x_{i+1}) \right], j=1,2,\dots,n \quad (16)$$

Дані алгоритми є алгоритмами статистичної оптимізації, окремими випадками методу СКГ. Описаним алгоритмам можна дати зручну геометричну інтерпретацію використовуючи для цього процес навчання. Будемо вважати, що в просторі X існує поділяюча поверхня $\Phi(x)$, яку можливо представити розкладанням (1), причому така, що виконуються умови (2). Тоді в зпрямляючому просторі існує минаюча через початок координат поділяюча площина з напрямним вектором V :

$$V^T Z = \sum_{i=1}^{\infty} v_i z_i,$$

така що

$$V^T Z = \begin{cases} > 0, & \text{если } x \\ < 0 & \text{если } x. \end{cases} \quad (17)$$

Відобразивши множину V_2 симетрично відносно початку координат, одержимо множину $V = V_1 \cup V_2^*$, де V_2^* – відображення множини V_2 . (див. рис. 1)

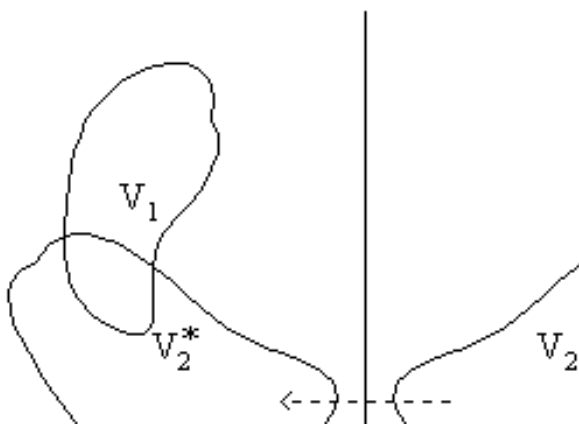


Рисунок 1 – Відображення множин

Множини V_1 й V_2 роздільні площиною з напрямним вектором V ($V^T Z = 0$) за умови, що:

$$V^T Z = \sum_{i=1}^{\infty} v_i z_i > 0, \text{ при } .$$

Іншими словами множини V_1 й V_2 розділяються цією площиною за умови, що об'єднана область V лежить по одну сторону від її. Поставимо у відповідність послідовність M крапок $\{x_1, x_2, \dots, x_p\}$ із простору X , що належать множинам V_1 і V_2 , і послідовність M крапок $\{z_1, z_2, \dots, z_p\}$ з $V = V_1 \cup V_2^*$. Потенційна функція (3) може бути представлена в зпрямляючому просторі Z як скалярний добуток двох векторів

$$Z_i = \alpha_i \varphi_i(x) \text{ и } Z_i^* = \alpha_i \varphi_i(x^*) \quad (18)$$

Тоді вираження (3) для потенційної функції можна представити в наступному виді:

$$U(x, x^*) = Z \cdot Z^* \quad (19)$$

Тепер вираження (9) можна переписати так:

$$U_p(Z) = \sum_{Z_i \in M^*} Z_i \quad (20)$$

де Z_i – крапки з послідовності M^* , пред'явлення яких у процесі навчання привело до виправлення помилок.

Якщо з послідовності крапок M^* вилучити всі крапки, які не привели до виправлення помилок, а ті крапки, що залишилися, необхідні для виправлення помилок, перенумерувати підряд, то вираження (20) можна переписати в такий спосіб:

$$U_p(Z) = Z \cdot \sum_{m=1}^{k(p)} Z_m, \quad Z_m \quad (20a)$$

де M^{**} – множина крапок, що включають у себе тільки ті крапки, які супроводжувалися виправленням помилок, які відбулися протягом перших p пред'явлень.

Виправлення помилки в т. $Z \in V$ буде відбуватися за умови $U_p(Z) < 0$.

На підставі цього можна затверджувати, що $(k+1)$ -е виправлення помилки відбудеться, якщо:

$$Z_{k+1} \cdot \sum_{m=1}^k Z_m \quad (21)$$

Тепер роботу алгоритму можна пояснити в такий спосіб. Поява першої крапки Z_1 з множини M^* приводить до побудови в зпрямляючому просторі площини $U_1(Z) = (Z, Z_1)$ з напрямним вектором Z_1 . У випадку, якщо наступної крапки з M^* лежить у тім же підпросторі (просторі), куди орієнтований напрямний вектор Z_1 , то помилка відсутня й при цьому положення поділяючої площини й вектор Z_1 не змінюються, і виробляється пред'явлення наступного зображення. Як тільки запропонована крапка потрапить у протилежний півпростір, відбувається виправлення помилки.

При цьому напрямний вектор площини V , побудований до цього кроку складається з вектором крапки, що зажадала виправлення й сумарний вектор приймається за новий напрямний вектор (див. рис 2, де показаний випадок, коли виправлення помилки треба було вже на 2-му кроці. Тут новий напрямний вектор дорівнює $Z_1 + Z_2$). У загальному випадку, після k виправлень помилок напрямний вектор поділяючої площини дорівнює

$$\sum_{m=1}^k Z_m \text{ (где } Z_m \in \cdot \text{)} \quad (22)$$

Після певного числа пред'явлень зображень із навчальної послідовності одержали площину, що проходить через початок координат зпрямляючого простору приймається за поділяючу площину. На цьому навчання закінчується. У процесі розпізнавання для кожної нової крапки підраховується значення $U(x)$, а рішення приймається за знаком $U(x)$.

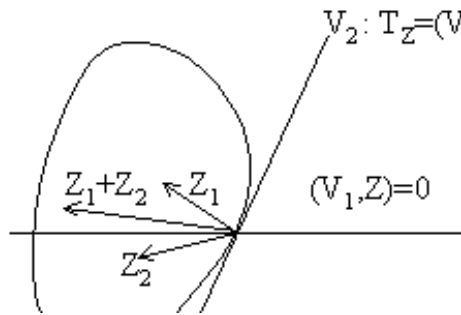


Рисунок 2 – Випадок, коли виправлення помилки треба було вже на 2-му кроці

Збіжність алгоритму потенційних функцій і умова його зупинки.

В [2] доведені 2 теореми, у яких затверджуються дуже важливі властивості алгоритмів, заснованих на методі потенційних функцій. Перша з них затверджує, що число виправлень помилок в алгоритмах звичайно, а в другий доводиться збіжність цих алгоритмів.

Теорема 3 Нехай M – довільна нескінченна послідовність крапок $\{x_1, x_2, \dots, x_p\}$ простору X , що належать множинам V_1 і V_2 . Припустимо, що існує функція $\Phi(x)$ така, що:

$$\Phi(x) = \begin{cases} > \xi, & \text{если } x \\ < -\xi & \text{если } x \end{cases} \quad (24)$$

Крім того, ця функція представима розкладанням (1). Нехай $U(x, x^*)$ обмежена на $V_1 \cup V_2$. Тоді існує ціле число k , що не залежить від вибору послідовності M , таке, що при використанні методу потенційних функцій число виправлень помилок не перевершує k , де:

$$k \leq \left\lceil \frac{\sup_{x \in V_1 \cup V_2} \sqrt{U(x, x^*)}}{\inf_{x \in V_1 \cup V_2} |\varphi(x)|} \sqrt{\sum_{i=1}^{\infty} \frac{1}{i^2}} \right\rceil. \quad (25)$$

Наведена теорема не доводить збіжності функції $U_p(x)$, формованої відповідно до алгоритму поділяючої функції, тому що вона не накладає ніяких обмежень на статистику пред'явлення зображень у процесі навчання.

Насамперед, для правильного поділу множин V_1 і V_2 необхідна досить представницька навчальна послідовність. Це означає, що крапки, що відповідає зображенням з навчальної послідовності повинні бути досить розкидані по множинах V_1 і V_2 .

Теорема 4. Нехай множини V_1 й V_2 у просторі X такі, що існує поділяюча функція, що задовольняє умові (24), представима розкладанням (1). Нехай функція $U(x, x^*)$ обмежена при $x \in V_1 \cup V_2$. Припустимо, що поява зображень із навчальної послідовності суть незалежні випадкові події й, яким би не було n , до n -му кроку алгоритму є строго позитивна ймовірність виправлення помилки за умови, що до цього кроку не відбулося повного поділу множин V_1 і V_2 . Тоді з імовірністю, рівної 1, для кожної реалізації алгоритму знайдеться таке кінцеве число m , що

$$U_m(x) = \begin{cases} > 0, & \text{если } x \\ < 0 & \text{если } x \end{cases} \quad (26)$$

Тобто процес поділу множин з імовірністю $P=1$ здійснюється за кінцеве число кроків.

Структурна схема системи наведена на рисунку 3.

Структурно система складається з наступних частин:

1. Відеокамера спостереження, з якої поступає інформація систему розпізнання образів.
2. Система розпізнання образів, яка складається з наступних блоків:
 - Блок читання картинки з відеокамери.
 - Блок аналізу та виділення ознак за допомогою багатопрохідної схеми інтелектуального розпізнавання образів методом потенційних функцій на основі кластерного аналізу.
 - Блок класифікації та опису об'єкта.
3. База даних журналювання розпізнаних образів.
4. База даних образів облич.
5. База даних образів цифр та букв на номерах автомобілів.



Рисунок 3 – Структурна схема системи

Висновки. У статті наведені теоретичне узагальнення й рішення наукового завдання дослідження методів інтелектуального розпізнавання образів методом потенційних функцій.

Рішення даного завдання полягало у вирішенні наступних задач:

- Був проведений огляд існуючих систем інтелектуального розпізнавання образів методом потенційних функцій.
- Досліджена система інтелектуального розпізнавання образів методом потенційних функцій.
- На основі отриманих результатів досліджень створена програмна реалізація системи інтелектуального розпізнавання образів методом потенційних функцій.

Розроблені під час виконання випускної кваліфікаційної роботи за другим (магістерським) рівнем вищої освіти алгоритми дозволяють успішно вирішувати завдання

інтелектуального розпізнавання образів методом потенційних функцій. Проведено аналіз предметної галузі в ході якого були виявлені об'єкти, взаємодія яких носить істотний характер для функціональної діяльності предметної галузі, і їхні основні характеристики; побудована алгоритм і вибраний середовище розробки.

Список літератури

1. Smirnov O., Fedorov E., Neskorođieva A., Neskorođieva T. «Intellectual Classification method of Gymnastic Elements Based on Combinations of Descriptive and Generative Approache». CEUR Workshop Proceedings Volume 3664, 2024, Pages 11-23.
2. Malyukov V., Bebeško B., Lakhno V., Smirnov O., Malyukova I., Mohylnyi H. «Managing the Purchase-Sale Process of Digital Currencies Under Fuzzy Conditions». Lecture Notes in Networks and Systems, 2023, 729 LNNS, pp. 104–112.
3. Al-Mudhafar Aqeel, A.M., Smirnova, T., Buravchenko, K., Smirnov, O. «The method of assessing and improving the user experience of subscribers in software-configured networks based on the use of machine learning». Advanced Information Systems, 2023, 7(2), pp. 49-56.
4. Smirnov, O., Sydorenko, V., Aleksander, M., Zhyharevych, O., Yenchey, S. «Simulation of the cloud IoT-based monitoring system for critical infrastructures». CEUR Workshop Proceedings, Volume 3530, 2023, pp. 256-265.
5. Smirnov, O., Karapetyan, A., Fedorov, E., «Creating Neural Network and Single Solution Human-Based Metaheuristic Methods of Solving the Traveling Salesman Problem». CEUR Workshop Proceedings, Volume 3312, 2022, pp. 47-58.
6. Smirnov, O., Neskorođieva, T., Fedorov, E., Rudakov, K., Neskorođieva, A. «Method Detection Audit Data Anomalies on Basis Restricted Cauchy Machine» CEUR Workshop Proceedings, Volume 3187, 2022, pp. 1-12.
7. Smirnov O., Smirnova T., Anas M. Al-Oraiqat, Drieiev O., Polishchuk L., Sheroy Khan, Yassin M. Y. Hasan, Aladdein M. Amro, Hazim S. AlRawashdeh «Method for Determining Treated Metal Surface Quality Using Computer Vision Technology». Sensors (Basel, Switzerland) Volume 22, Issue 16, 6223, 2022.
8. Smirnov, O., Lakhno, V., Akhmetov, B., Chubaievskyi, V., Khorolska, K., Bebeško, B. «Selection of a Rational Composition of Information Protection Means Using a Genetic Algorithm». In: Rajakumar, G., Du, K.L., Vuppapapati, C., Beligiannis, G.N. (eds) Intelligent Communication Technologies and Virtual Mobile Networks. Lecture Notes on Data Engineering and Communications Technologies, vol 131. 2023. Springer, Singapore. pp. 21-34.
9. Kuznetsov, A., Oleshko, I., Chernov, K., Bagmut, M., Smirnova, T. «Biometric authentication using convolutional neural networks». Lecture Notes in Networks and Systems. Volume 152, 2021, Pages 85-98.
10. Smirnov O., Kuznetsov A., Kryvinska N., Kiian A., Kuznetsova K. «Full Non-Binary Constant-Weight Codes». SN Computer Science, Vol 2, 337, 2021. <https://doi.org/10.1007/s42979-021-00739-w>.
11. Smirnov O., Neskorođieva T., Fedorov E., Rymar P. «Neural Network Modeling Method of Transformations Data of Audit Production with Returnable Waste». CEUR Workshop Proceedings Volume 3101, 2021, Pages 192-207.
12. Smirnov, O., Kuznetsov, A., Potii, O., Poluyanenko, N., Stelnyk, I., Mialkovsky, D. «Combining and filtering functions in the framework of nonlinear-feedback shift register». International Journal of Computing; 2020, Volume 19, Issue 2 – Research Institute for Intelligent Computer Systems – 2020. – P. 247-256.
13. Smirnov O., Kuznetsov A., Kiian A., Kuznetsova T. «Non-binary constant weight coding technique». CEUR Workshop Proceedings. Volume 2740, 2020, Pages 102-114.
14. Smirnov O., Kuznetsov A., Kiian A., Cherep A., Kanabekova M., Chepurko I. «Testing of code-based pseudorandom number generators for post-quantum application». 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT), Ukraine, Kyiv, May 14-18. 2020. P. 172-177.
15. Smirnov O., Kuznetsov A., Pushkar'ov A., Serhiienko R., Babenko V., Kuznetsova T., «Representation of Cascade Codes in the Frequency Domain». In: Radivilova T., Ageyev D., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 48. Springer, Cham. 2021. pp 557-587.
16. Smirnov, O., Drieieva, H., Drieiev, O., Polishchuk, Y., Brzhanov, R., Aleksander, M. «Method of fractal traffic generation by a model of generator on the graph». CEUR Workshop Proceedings Volume 2616, 2020, Pages 366-379.
17. Smirnov, O., Drieieva, H., Drieiev, O., Simakhin, V., Bondar, S., Odarchenko, R. «Managing multifractal properties of the binary sequence generated with the Markov chains», CEUR Workshop Proceedings Volume 2608, 2020, Pages 633-645.
18. Smirnov O. Kuznetsov A., Zaichenko Yu., Pastukhov M., Oleshko O., Kuznetsova K., «Formation of Discrete Signals with Special Correlation Properties». International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2019; Odessa; Ukraine; 9-13 September 2019. P.22-28.
19. Smirnov, O., Kuznetsov, A., Kolovanova, I., Kuznetsova, T., «Noise immunity of the algebraic geometric codes». International Journal of Computing; 2019, Volume 18, Issue 4 – Research Institute for Intelligent Computer Systems – 2019. – P. 393-407.

20. Smirnov, O., Kuznetsov, A., Reshetniak, O., Ivko, N., Katkova, T., Kuznetsova, T., «Generators of Pseudorandom Sequence with Multilevel Function of Correlation». 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), Kyiv, Ukraine, 8 – 11 October 2019. P.517-522.
21. Smirnov, O., Ulichev, O., Meleshko, Y., Khokh, V., Goncharenko, I. «Method of Choosing Objects for Informational Influence in Social Networks during Information Campaign Based on the Analytic Hierarchy Process». CEUR Workshop Proceedings, Vol 2588, P. 215-227, 2019.
22. Smirnov, O., Krasnobayev, V., Yanko, A., Kuznetsova, T. «Methods of nulling numbers in the system of residual classes». CEUR Workshop Proceedings, Vol 2588, P. 90-106, 2019.
23. Smirnov, O., Kuznetsov, A., Kovalchuk, D., Averchev, A., Pastukhov, M., Kuznetsova, K., «Formation of Pseudorandom Sequences with Special Correlation Properties», 2019 3rd International Conference on Advanced Information and Communications Technologies, AICT -2019/ Lviv, Ukraine, 2-6 July, 2019, P. 395-399.
24. Smirnov, O., Kuznetsov, A., Kiian, A., Zamula, A., Rudenko, S., Hryhorenko, V., «Variance Analysis of Networks Traffic for Intrusion Detection in Smart Grids», 2019 IEEE 6th International Conference On Energy Smart Systems (2019 IEEE ESS), Kyiv, Ukraine April 17-19, 2019 P. 353-358.
25. Smirnov, O., Kuznetsov, A., Kavun, S., Babenko, B., Nakisko, O., Kuznetsova, K., «Malware Correlation Monitoring in Computer Networks of Promising Smart Grids», 2019 IEEE 6th International Conference On Energy Smart Systems (2019 IEEE ESS), Kyiv, Ukraine April 17-19, 2019 P. 347-352.
26. Smirnov, O., Kuznetsov, A., Kovalchuk, D., Pastukhov, M., Kuznetsova, K., Prokopovych-Tkachenko, D., «Discrete Signals with Special Correlation Properties», CEUR Workshop Proceedings Volume 2353, CEUR Workshop Proceedings 2019, Pages 618-629.