

РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ ОСІБ НА ОСНОВІ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

О.О. Яковенко, Н.І. Кушніренко, І.С. Дорофєєва, А.Р. Євтушенко

Одеський національний політехнічний університет
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: dorofeeva283@gmail.com

Захищеність і цілісність особистих даних є актуальною проблемою в сучасному світі, так як останнім часом стало більш випадків зломів паролів, банківських аккаунтів і рахунків, невірної ідентифікації особистостей, подробиць особистих даних. Такі способи автентифікації, як доступ за паролем, використання електронних перепусток, одноразові повідомлення з кодом є ненадійними на відміну від біометричних методів. Пароль можна забути, втратити, підібрати або вкрасти, а біометричні системи контролю доступу зручні для користувачів тим, що носії інформації знаходяться завжди при них, не можуть бути загублені або вкрадені. Біометричний контроль доступу вважається більш надійним, так як ідентифікатори не можуть бути передані третім особам, скопійовані. З усіх видів біометрії (ідентифікація за відбитками пальців, по райдужній оболонці ока, по голосу, по геометрії руки) ми вирішили зупинитися на розпізнаванні облич. Проаналізувавши існуючі методи класифікації зображень, був вибраний оптимальний варіант – згортальна нейронна мережа, успіх якої обумовлений можливістю обліку двовимірної топології зображення, на відміну від багаточасового перцептрона. Технології розпізнавання осіб застосовуються в найрізноманітніших сферах: забезпечення безпеки в місцях великого скупчення людей; системах охорони; фейс-контроль в сегменті громадського харчування та розваг, пошук підозрілих і потенційно небезпечних відвідувачів; верифікація банківських карт; онлайн-платежі. В роботі розроблена система класифікації облич на основі згорткової нейронної мережі та система інтерпретації результатів класифікації, що дозволяє задавати співвідношення помилок 1-го та 2-го роду та обирати поріг детектування на основі цього співвідношення. На відміну від стандартного рішення, система не використовує критерій максимальної правдоподібності, що дозволяє отримувати більше інформації від класифікатора та зменшити рівень помилок системи. Вихідними даними для нашого дослідження є 840 фотографій, на яких зображені автори статті. Розроблений нами метод є інноваційним і дозволяє поліпшити комплексні захисні системи.

Ключові слова: згорткова нейронна мережа, розпізнавання обличчя, машинне навчання, класифікація зображень, управління помилками.

Вступ

Захищеність і цілісність особистих даних є актуальною проблемою в сучасному світі, так як останнім часом стало більш випадків зломів паролів, банківських аккаунтів і рахунків, невірної ідентифікації особистостей, подробиць особистих даних. Такі способи автентифікації, як доступ за паролем, використання електронних перепусток, одноразові повідомлення з кодом є ненадійними на відміну від біометричних методів. Пароль можна забути, втратити, підібрати або вкрасти, а біометричні системи контролю доступу зручні для користувачів тим, що носії інформації знаходяться завжди при них, не можуть бути загублені або вкрадені. Біометричний контроль доступу вважається більш надійним, так як ідентифікатори не можуть бути передані третім особам, скопійовані. З усіх видів біометрії (ідентифікація за відбитками пальців, по райдужній оболонці ока, по голосу, по геометрії руки) було прийнято рішення зупинитися на розпізнаванні облич. Проаналізувавши існуючі методи класифікації зображень, було обрано оптимальний варіант – згорткова нейронна мережа, успіх якої обумовлений

можливістю обліку двовимірної топології зображення, на відміну від багаточарового перцептрона. Технології розпізнавання осіб застосовуються в найрізноманітніших сферах: забезпечення безпеки в місцях великого скупчення людей; системах охорони; фейс-контроль в сегменті громадського харчування та розваг, пошук підозрілих і потенційно небезпечних відвідувачів; верифікація банківських карт; онлайн-платежі. Було розроблено згорткову нейронну мережу, та метод інтерпретації її результатів з подальшою можливістю вибору порога детектування, що надає можливість управляти помилками 1-го і 2-го роду. Математична модель згорткової нейронної мережі була побудована в системі Matlab. Вихідними даними для дослідження є 840 фотографій, на яких зображені автори статті. Розроблений метод базується на сучасних технологіях, та поєднує переваги загортальних нейронних мереж та розповсюджені методи управління помилками.

Актуальність дослідження полягає в тому, що багато напрямків науки, техніки і виробництва в значній мірі орієнтуються на розвиток систем, в яких інформація несе характер поля (зображення). При обробці такої інформації виникає ряд складних наукових, технічних і технологічних проблем. Однією з найскладніших на сьогоднішній момент з них є обробка і розпізнавання зображень. Про важливість цієї проблеми свідчить той факт, що дослідження з розпізнавання образів, аналізу зображень та мови включені в перелік пріоритетних напрямів розвитку науки і техніки.

Розпізнавання зображень знаходить широке застосування в різних сферах діяльності людини – це може бути контроль топології друкованих плат, текстури тканини, робототехніка (інтелектуальні системи). В інформатиці – контроль доступу до інформації щодо ідентифікації особи (біометрична ідентифікація). Спеціальне застосування – це контроль доступу до об'єктів обмеженого доступу, оперативний пошук в картотеці зображень, дактилоскопія та ін. Широко використовуються ці методи для класифікації історичних джерел на папері, а також у фізиці, хімії, біології та ін. галузях науки [1].

В даний час все більш широке поширення набувають біометричні системи ідентифікації людини. Цей напрямок знайшов відображення у роботах В.В. Моліцького, І.І. Данилюка, В.В. Карпінєцького, А.В. Приймака, Ю.Є. Яремчука, О.І. Костюченка, М.О. Войтка, І.А. Терейковського. Традиційні системи ідентифікації вимагають знання пароля, наявності ключа, ідентифікаційної картки, або іншого ідентифікуючого предмету, який можна забути або втратити. На відміну від них біометричні системи ґрунтуються на унікальних біологічних характеристиках людини, які важко підробити і які однозначно визначають конкретну людину. До таких характеристик відносяться відбитки пальців, форма долоні, візерунок райдужної оболонки, зображення сітківки ока. Особа, голос і запах кожної людини так само індивідуальні [2].

Мета роботи

Метою дослідження є розробка згорткової нейронної мережі з постобробкою результатів, що полягає в встановленні порога. Це дає можливість краще розпізнавати зображення і керувати помилками 1-ого та 2-ого роду.

Основна частина

Штучну нейронну мережу (ШНМ) також називають просто «нейронна мережа» (НМ) – це математична модель, побудована за принципом роботи біологічних нейронних мереж – мереж нервових клітин живого організму. ШНМ складається з пов'язаної групи штучних нейронів і обробляє інформацію, використовуючи коннективізм. У більшості випадків – це адаптивна система, яка змінює свою структуру,

грунтуючись на обробці вхідної або вихідної інформації, яка тече через мережу під час фази навчання.

Машинне навчання – великий підрозділ штучного інтелекту, що вивчає методи побудови алгоритмів, здатних навчатися.

Згортовка нейронна мережа – спеціальна архітектура ШНМ, запропонована Яном Лекуном [3] і націлена на ефективне розпізнавання зображень, входить до складу технологій глибокого навчання. Використовує деякі особливості зорової кори, в якій були відкриті так звані прості клітини, що реагують на прямі лінії під різними кутами, і складні клітини, реакція яких пов'язана з активацією певного набору простих клітин. Таким чином, ідея згорткових нейронних мереж полягає в чергуванні згорткових шарів і субдискретизуючих шарів.

Переваги згорткової нейронної мережі перед звичайною ШНМ:

- в порівнянні з повнозв'язною нейронною мережею – набагато менша кількість налаштованих ваг, так як одне ядро ваг використовується цілком для всього зображення, замість того, щоб робити для кожного пікселя вхідного зображення свій персональний ваговий коефіцієнт. Це підштовхує ШНМ при навчанні до узагальнення демонстрованої інформації, а не до попиксельного запам'ятовування кожної показаної картинки у великій кількості вагових коефіцієнтів, як це робить перцептрон [4];
- зручне розпаралелювання обчислень, а також можливість реалізації алгоритмів роботи і навчання мережі на графічних процесорах;
- відносна стійкість до повороту і зсуву зображення, яке розпізнається;
- навчання за допомогою класичного методу зворотного поширення помилки.

Метод гнучкого порівняння на графах [5]. Суть методу зводиться до еластичного порівняння графів, що описують зображення осіб. Особи представлені у вигляді графів зі зваженими вершинами і ребрами. На етапі розпізнавання один з графів – еталонний, залишається незмінним, в той час як інший деформується з метою найкращої підгонки до першого. У подібних системах розпізнавання графи можуть являти собою як прямокутну решітку, так і структуру, утворену характерними точками особи. Недоліки: висока обчислювальна складність процедури розпізнавання, низька технологічність при запам'ятовуванні нових еталонів, лінійна залежність часу роботи від розміру бази даних осіб.

Нейронні мережі [6]. В даний час існує безліч різновидів нейронних мереж. Одним з найбільш широко використовуваних варіантів є мережа, побудована на багат шаровому перцептроні, яка дозволяє класифікувати подане на вхід зображення/сигнал відповідно до попереднього навчання мережі. Недоліки: додавання нової еталонної особи в базу даних вимагає повного перенавчання мережі на всьому наявному наборі (досить тривала процедура, в залежності від розміру вибірки від 1 години до декількох днів). Проблеми математичного характеру, пов'язані з навчанням: потрапляння в локальний оптимум, вибір оптимального кроку оптимізації, перенавчання і тому подібне. Важко формалізується етап вибору архітектури мережі (кількість нейронів, шарів, характер зв'язків).

Приховані Марковські моделі (ПММ) [7]. Одним з статистичних методів розпізнавання осіб є приховані Марковські моделі з дискретним часом. ПММ використовують статистичні властивості сигналів і враховують безпосередньо їх просторові характеристики. Елементами моделі є: безліч прихованих станів, безліч спостережуваних станів, матриця перехідних ймовірностей, початкова ймовірність станів. Кожному відповідає своя Марковська модель. При розпізнаванні об'єкта перевіряються згенеровані для заданої бази об'єктів Марковської моделі і шукається максимальна із спостережуваних ймовірність того, що послідовність спостережень для даного об'єкта згенерована відповідною моделлю. На сьогоднішній день не вдалося знайти приклад комерційного застосування ПММ для розпізнавання осіб. Недоліки: необхідно підбирати параметри моделі для кожної бази даних.

Метод головних компонент (Principal Component Analysis, PCA) [8]. Одним з найбільш відомих і опрацьованих є метод головних компонент, заснований на перетворенні Карунена-Лосва. Спочатку метод головних компонент почав застосовуватися в статистиці для зниження вимірності простору ознак без істотної втрати інформації. У задачі розпізнавання осіб його застосовують головним чином для представлення зображення особи вектором малої розмірності (головних компонент), який порівнюється потім з еталонними векторами, закладеними в базу даних. Недоліки: метод головних компонент добре зарекомендував себе в практичних додатках. Однак в тих випадках, коли на зображенні особи присутні значні зміни в освітленості або виразі обличчя, ефективність методу значно падає. Вся справа в тому, що PCA вибирає підпростір з такою метою, щоб максимально апроксимувати вхідний набір даних, а не виконати дискримінацію між класами осіб.

Активні моделі зовнішнього вигляду (Active Appearance Models, AAM) і активні моделі форм (Active Shape Models, ASM) [9].

Активні моделі зовнішнього вигляду (AAM) – це статистичні моделі зображень, які шляхом різного роду деформацій можуть бути підігнані під реальне зображення. Даний тип моделей в двовимірному варіанті було запропоновано Тімом Кутсом і Крісом Тейлором в 1998 році. Спочатку активні моделі зовнішнього вигляду застосовувалися для оцінки параметрів зображень облич. Активна модель зовнішнього вигляду містить два типи параметрів: параметри, пов'язані з формою, і параметри, пов'язані зі статистичною моделлю пікселів зображення або текстурою. Суть методу ASM полягає в обліку статистичних зв'язків між розташуванням антропометричних точок на наявній вибірці зображень осіб, знятих в анфас. На зображенні експерт розмічає розташування антропометричних точок. На кожному зображенні точки пронумеровані в однаковому порядку. Недоліки: моделі AAM і ASM призначені для того, щоб точно локалізувати ці антропометричні точки на зображенні особи, а не класифікувати його.

Згорткові нейронні мережі [10]. Найкращі результати в області розпізнавання осіб (за результатами аналізу публікацій) показала Convolutional Neural Network або згорткова нейронна мережа (далі – ЗНМ). Успіх обумовлений можливістю обліку двовимірної топології зображення, на відміну від багат шарового перцептрона.

Саме тому для вирішення нашої задачі ми будемо використовувати ЗНМ.

Відмінними рисами ЗНМ є локальні рецепторні поля (забезпечують локальну двовимірну зв'язність нейронів), загальні ваги (забезпечують детектування деяких рис в будь-якому місці зображення) і ієрархічна організація з просторовими семплінгом. Завдяки цим нововведенням ЗНМ забезпечує часткову стійкість до змін масштабу, зсувів, поворотів, змін ракурсу і інших спотворень.

Особливістю згорткової нейронної мережі є операція згортки. Згортка – це операція над парою матриць A (розміром $n_x \times n_y$) та B (розміром $m_x \times m_y$), результатом якої є матриця $C = A \times B$ розміру $(n_x - m_x + 1) \times (n_y - m_y + 1)$. Кожен елемент результату обчислюється як множення матриці B і деякої підматриці A такого ж розміру (підматриця визначається положенням елемента в результаті).

Логічний сенс згортки полягає в наступному: чим більше величина елемента згортки, тим більше ця частина матриці A буде схожа на матрицю B (схожа в контексті скалярного множення). Тому матрицю A називають зображенням, а матрицю B – фільтром або зразком (kernel).

У згортковій нейронній мережі виходи проміжних шарів утворюють матрицю (зображення) або набір матриць (кілька шарів зображення). Так, наприклад, на вхід згорткової нейронної мережі можна подавати три шари зображення (R-, G-, B-канали зображення). Основними видами шарів в згортковій нейронній мережі є згорткові шари, пулінгові шари і повнозв'язні шари.

Згорткові нейронні мережі працюють на основі фільтрів, які займаються розпізнаванням певних характеристик зображення (наприклад, прямих ліній). Фільтр – це колекція кернелів; іноді в фільтрі використовується один кернел (рис. 1). Кернел – це звичайна матриця чисел, що називаються вагами, які навчаються з метою пошуку на зображеннях певних характеристик. Фільтр зміщується уздовж зображення і визначає, чи присутня деяка шукана характеристика в конкретній його частині [11]. Для отримання відповіді такого роду здійснюється операція згортки, яка є сумою добутків елементів фільтра і матриці вхідних сигналів.

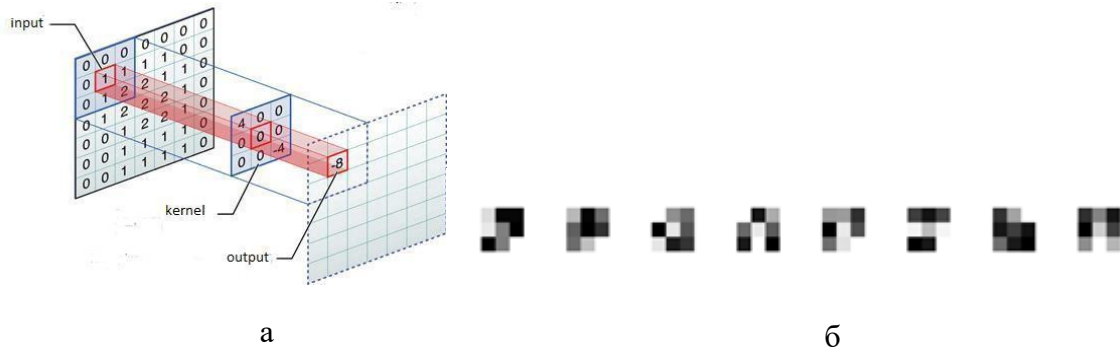


Рис. 1. Приклад роботи фільтра: а – візуалізація процесу згортки; б – приклад кернелів, отриманих у результаті експерименту

Архітектура згорткової нейронної мережі

Згорткова нейронна мережа складається з декількох типів шарів (рис. 2).

Вхідні дані (Image Input) – це шар, у якому зображення подається на ЗНМ.

Двовимірна згортка (Convolution 2D) – двовимірний згортковий шар застосовує ковзаючи згорткові фільтри до вхідних даних. Шар згортає вхідні дані, переміщаючи фільтри уздовж вхідних даних по вертикалі і горизонталі і обчислюючи точкове множення ваг і вхідних даних, а потім додаючи член зміщення.

Пакетна нормалізація (Batch Normalization) – шар нормалізації партії, нормалізує кожен вхідний канал в міні-партії. Щоб прискорити навчання згорткових нейронних мереж і знизити чутливість до ініціалізації мережі, шари пакетної нормалізації використовуються між згортковими шарами та лінійними шарами, такими як *ReLU*. Шар спочатку нормалізує активації кожного каналу, віднімаючи середнє значення міні-партії і ділячи його на стандартне відхилення міні-партії. Потім шар зрушує вхідний сигнал на навчене зміщення β і масштабує його на той, який навчає масштабний коефіцієнт γ .

Шар-випрямляч (*ReLU*) – виконує порогову операцію для кожного елемента вводу, де будь-яке значення менше нуля встановлюється в нуль.

Шар вибору максимального елемента (*Max Pooling 2D*) виконує знижувальну дискретизацію шляхом ділення вхідних даних на прямокутні області пулу і обчислення максимуму кожної області.

Повнозв'язний шар (*Fully Connected*) – повністю пов'язаний шар примножує вхідні дані на матрицю ваг, а потім додає вектор зміщення.

Softmax – застосовує функцію *softmax* до входу.

Результат класифікації (*Classification Output*) – шар класифікації, обчислює крос-ентропійну втрату для задач мультикласової класифікації з взаємовиключними класами. Шар виводить кількість класів з вихідного розміру попереднього шару.

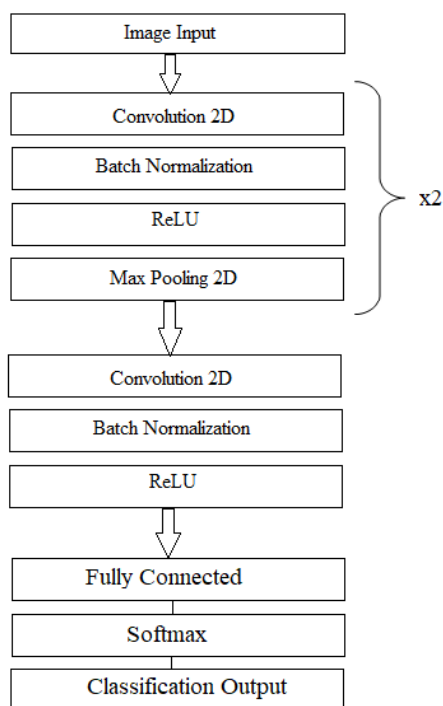


Рис. 2. Структурна схема ЗНМ

Метод постобробки при класифікації зображень.

На рисунку 3 зображено: 2 класи зображень (A і B), зображення X і визначається близькість зображення до одного з класів. У такому виді дані отримуються зі загорткової мережі.

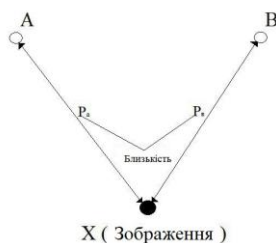


Рис. 3. Графічне зображення поняття близькості до класу, A та B – класи зображень, P_A та P_B – близькість зображення до класів A та B відповідно

У таблиці 1 показані дані, в яких відображені значення параметра P (близькість, *proximity*).

Таблиця 1.

Значення параметра близькості зображень до класів A та B для перших трьох зображень

	P_A	P_B
1	0.9862	0.0138
2	0.9964	0.0036
3	0.9932	0.0068

Використовувались два класи зображень A і B , які в реальності являли собою фотографії двох людей, а саме авторів статті. Використання стандартного класифікатора Matlab визначає найближчий клас A або B .

За замовчуванням при класифікації зображень Matlab використовує принцип максимальної правдоподібності. Цей принцип не досконалий, так як не дозволяє гнучко налаштовувати помилки першого і другого роду.

Для того, щоб керувати помилками 1-го і 2-го роду потрібно використовувати близькість до класу. Наприклад, якщо $P_A = 0.9$, а $P_B = 0.89$, то система фактично не впевнена, але проста класифікація повертає клас A . В роботі розроблена система, яка враховує цей недолік. Пропонується використовувати параметр вірогідності класу A – L_A (*likelihood of A*), який дозволяє вирішити цей недолік (1). Цей параметр базується на різності параметрів близькості до класів A та B :

$$P_A - P_B.$$

Якщо зображення ближче до класу A , то його індекс $P_A - P_B$ більше нуля. Для наочності приведемо графік (рис. 4), де по горизонталі номери зображень, а по вертикалі індекс поточного зображення. Перші 320 зображень – реальний клас A , а другі – реальний клас B .

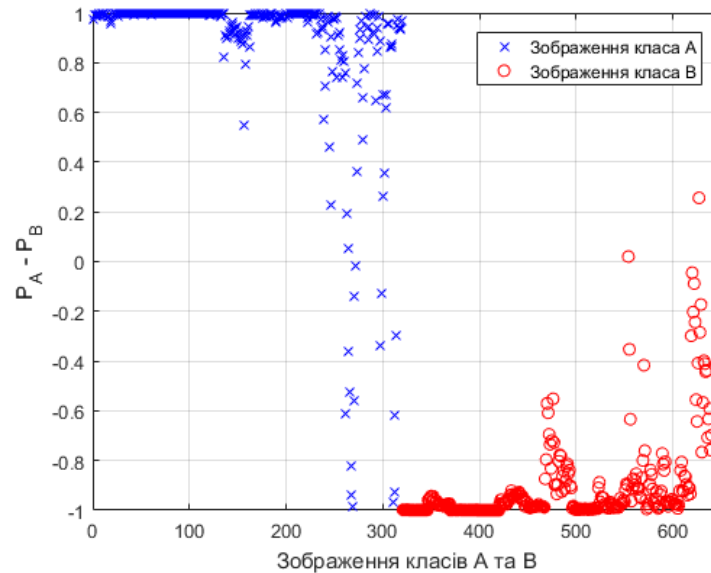


Рис. 4. Графік $P_A - P_B$

Для того, щоб перетворити індекс в діапазон $[0,1]$ перетворимо вище наведену формулу таким чином:

$$L_A = (P_A - P_B + 1)/2 \quad (1)$$

Тепер значення близькості до класу розподіляються від 0 до 1. 1 – клас A , 0 – клас B . На графіку видно (рис. 5), що в останніх 80 зображеннях програмі важко визначити належність до класу A або класу B .

Параметр L_A знаходиться в діапазоні $[0,1]$. Виникає питання встановлення значення порога t таким чином, щоб отримати контрольовані значення ймовірності помилок першого (P_{FP}) і другого (P_{FN}) роду.

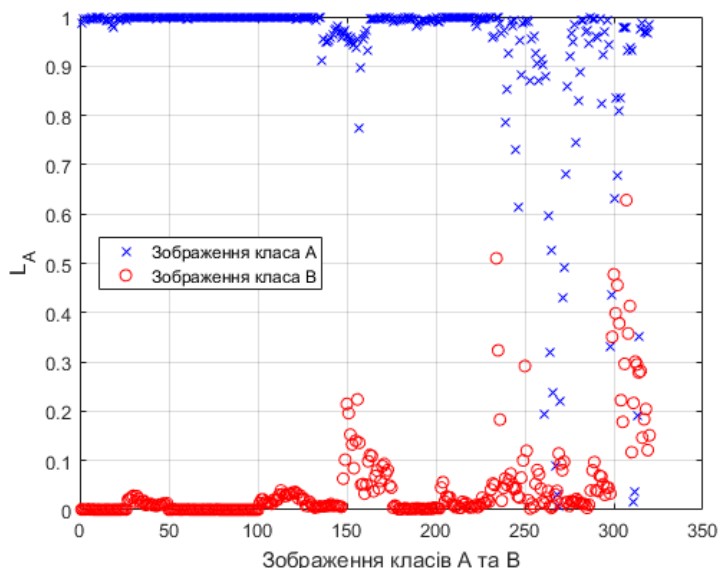


Рис. 5. Значення індексу L_A для зображень обох класів

Для цього використовуються отримані значення L_A для зображень обох класів, графіки помилок отримаємо наступним чином:

- перебираються значення порога t від 0 до 1 з деяким кроком (наприклад, 0.001);
- для кожного значення t знаходиться відсоток зображень класу A , для яких $L_A < t$, цей відсоток є P_{FN} для цього значення порога;
- аналогічно, відсоток значень класу B , для яких $L_A > t$ є P_{FP} для цього значення порога.

Таким чином, на графіку можна бачити залежність помилок від порога, і вибирати поріг відповідно з необхідним рівнем обох помилок. Перетин графіків помилок першого і другого роду відповідає рівності помилок ($P_{FP} = P_{FN}$). Клас A на графіку позначається х-подібними символами, клас B – кружками.

На графіку (рис. 6) зображені помилки 1-го і 2-го роду в залежності від значення порога детектування t . Помилка першого роду (також *false positive*, FP) – це помилка, при якій реальний елемент класу B детектується як клас A . Помилка другого роду (також *false negative*, FN) – це помилка, при якій реальний елемент класу A детектується як клас B .

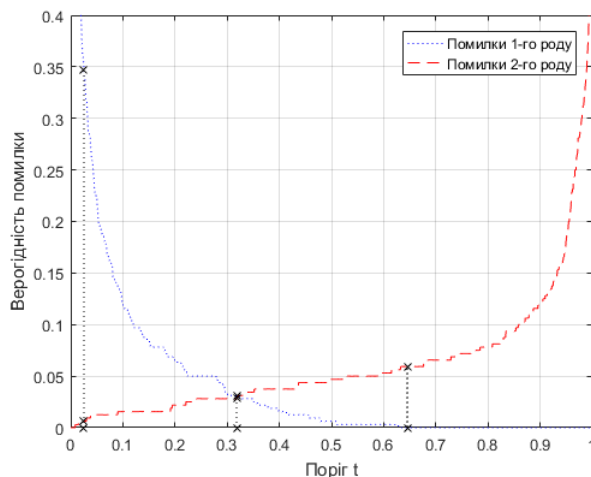


Рис. 6. Помилки 1-го і 2-го роду в залежності від порога детектування

Наприклад, є пропускний пункт. Клас A – співробітники організації, клас B – сторонні. При виборі порога $t = 0.318$, система буде в рівному ступені помилятися при визначенні близькості і до класу A , і до класу B , то ймовірність пропустити не співробітника дорівнює ймовірності не пропустити співробітника, і дорівнює $p = 0.04$.

Якщо вибрати поріг $t = 0.6452$, то система буде схильна до того, щоб детектувати елементи класу A як клас B . Іншими словами, ніколи не буде пускати сторонніх, але іноді, буде ймовірність не пустити співробітника $P_{FN} = 0.06$.

При виборі порога $t = 0.025$ система буде схильна до того, щоб детектувати елементи класу B як клас A . Тобто, система пропустить співробітників в 99% випадках, але з ймовірністю 35 % пропустить сторонніх.

Таким чином, розроблена система дозволяє адміністратору вибрати оптимальний варіант пропускної системи відповідно до потреби.

Особливості тренування ЗНМ.

Для дослідження було використано по 420 зображень кожного класу, на тренування довільного вибиралися по 100 зображень з кожного класу. Для тренування і аналізу зображення проходили попередню обробку, а саме знебарвлення і зменшення до розміру 25×25 пікселів (рис. 7). Зменшення зображень обумовлено продуктивністю.

Для тренування були використані наступні параметри:

- *sgdm* – оптимізатор стохастичного градієнтного спуску з імпульсом (*Stochastic Gradient Descent with Momentum*, SGDM [12]), можна вказати значення імпульсу, використовуючи аргумент пари ім'я-значення 'Momentum';
- *InitialLearnRate* – початкова швидкість навчання, яка використовується для навчання, вказується у вигляді розділеної запитом пари, що складається з *InitialLearnRate* і позитивного скаляра; значення за замовчуванням становить 0,01 для оптимізатора *sgdm* і 0,001 для вирішувачей *rmsprop* і *adam*, якщо швидкість навчання занадто низька, то навчання займає багато часу, якщо висока, то навчання може досягти неоптимального результату або розходитися;
- *Shuffle, every-epoch* – заново перемішувати тренувальні дані в кожній епохі;
- *ValidationFrequency* – валідація проводиться кожні 30 епох.

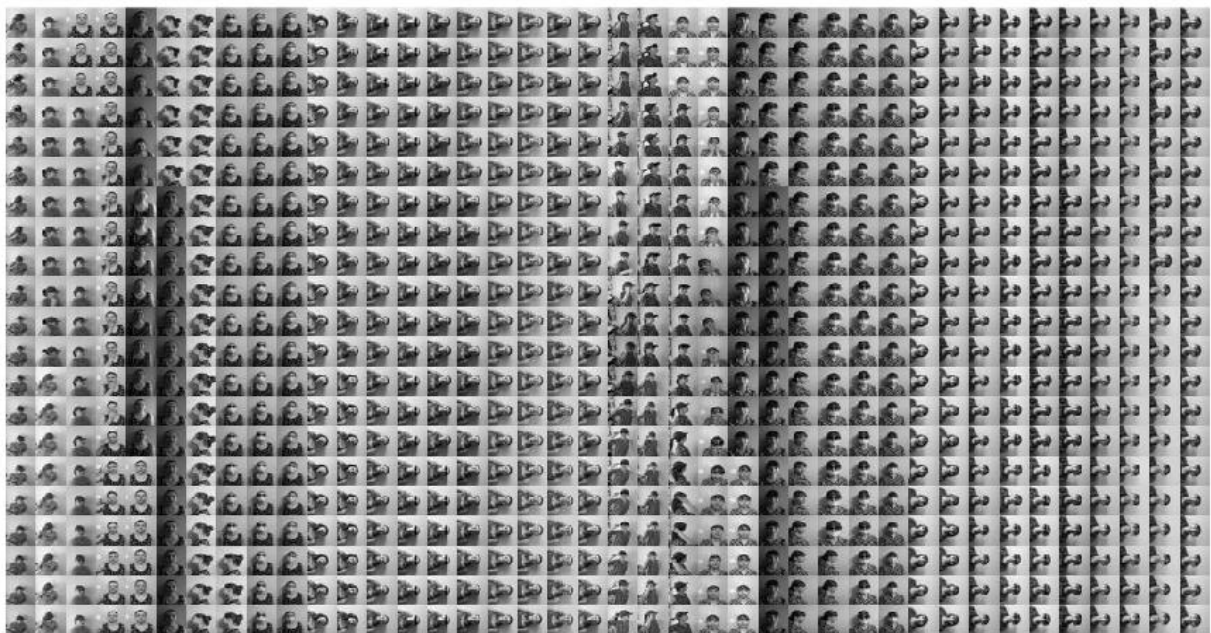


Рис. 7. Всі зображення (тренувальні та перевіральні)

Висновки

У даній роботі були проаналізовані існуючі рішення в області класифікації зображень. В результаті проведення експерименту була створена згортальна нейронна мережа з постобробкою результатів, були проаналізовані різні значення порога класифікації, в залежності від якого можуть завдання реалізовуватися різні політики безпеки. Розроблений метод є інноваційним і дозволяє покращити комплексні захисні системи. Даний напрямок буде об'єктом подальших досліджень.

Список літератури

1. Галушкин, А.И. Нейронные сети. Основы теории / А.И. Галушкин. – М.: Горячая линия – Телеком, 2010. – 480 с.
2. Панканти, Ш. Биометрия: будущее идентификации / Ш. Панканти, Р.М. Болле, Э. Джейн // Открытые Системы. – 2000. – № 3. – С. 51-63.
3. Santaji, G. Neural networks for facerecognition using SOM / G. Santaji, G. Jayshree, M. Shaml, G. Dhanaji // IJCSST. – 2010. – 287с.
4. Субботін, С.О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень / С.О. Субботін. – Запоріжжя: ЗНТУ, 2008. – 341 с.
5. Степанов, В.Н. Дискретная математика: графы и алгоритмы на графах / В.Н. Степанов. – ОмГТУ, 2010. – 120 с.
6. Савельев, А.В. На пути к общей теории нейросетей. К вопросу о сложности / А.В. Савельев // Нейрокомпьютеры: разработка, применение. – 2006. – С. 4-14.
7. Yu, S. Hidden semi-Markov models / S. Yu // Artificial Intelligence. – 2010. – Vol.174, No 2. – Pp. 215-243.
8. Jolliffe, I.T. Principal component analysis, series: springer series in statistics / I.T. Jolliffe. – 2nd ed. Springer, 2002. – 487 p.
9. Русай, А.Н. Биометрическая аутентификация диктора в MATLAB / А.Н. Русай. – Учебное пособие. – М.: Русайнс, 2017. – 512 с.
10. Барский, А.Б. Логические нейронные сети / А.Б. Барский. – М.: Интернет-университет информационных технологий, бином. Лаборатория знаний, 2007. – 352 с.
11. Fukushima, K. Artificial Vision by Multi-Layered Neural Networks: Neocognitron and its Advances / K. Fukushima // Neural Networks. – 2013. – Pp. 103-119.
12. Qian, N. On the momentum term in gradient descent learning algorithms. Neural Networks / N. Qian // The Official Journal of the International Neural Network Society. – 1999. – 12(1). – Pp. 145–151.

РАЗРАБОТКА СИСТЕМЫ РАСПОЗНАВАНИЯ ЛИЦ НА ОСНОВЕ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

А.А. Яковенко, Н.И. Кушниренко, И.С. Дорофеева, А.Р. Евтушенко

Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: dorofeeva283@gmail.com

Защищенность и целостность личных данных является актуальной проблемой в современном мире, так как в последнее время участились случаи взломов паролей, банковских аккаунтов и счетов, неверной идентификации личностей, подделки личных данных. Такие способы аутентификации, как доступ по паролю, использование электронных пропусков, одноразовые сообщения с кодом являются ненадежными в отличие от биометрических методов. Пароль можно забыть, потерять, подобрать или украсть, а биометрические системы контроля доступа удобны для пользователей тем, что носители информации находятся всегда при них, не могут быть утеряны либо украдены. Биометрический контроль доступа считается более надежным, т. к. идентификаторы не могут быть переданы третьим лицам, скопированы. Из всех видов биометрии (идентификация по отпечатку пальца, по радужной оболочке глаза, по голосу, по геометрии руки) мы решили остановиться на распознавании лиц. Проанализировав существующие методы классификации изображений, был выбран оптимальный вариант – сверточная нейронная сеть, успех

которой обусловлен возможностью учета двумерной топологии изображения, в отличие от многослойного персептрона. Технологии распознавания лиц применяются в самых разнообразных сферах: обеспечение безопасности в местах большого скопления людей; системы охраны, избежание незаконного проникновения на территорию объекта, поиск злоумышленников; фейс-контроль в сегменте общепита и развлечений, поиск подозрительных и потенциально опасных посетителей; верификация банковских карт; онлайн-платежи. В работе разработана система классификации лиц на основе сверточной нейронной сети и система интерпретации результатов классификации, что позволяет задавать соотношение ошибок 1-го и 2-го рода и выбирать порог детектирования на основе этого соотношения. В отличие от стандартного решения, система не использует критерий максимального правдоподобия, что позволяет получать больше информации от классификатора и уменьшить уровень ошибок системы. Исходными данными для нашего исследования являются 840 фотографий, на которых изображены авторы статьи. Разработанный нами метод является инновационным и позволяет улучшить комплексные защитные системы.

Ключевые слова: сверточная нейронная сеть, распознавание лица, машинное обучение, классификация изображений, управление ошибками.

DEVELOPING OF THE FACE RECOGNITION SYSTEM ON THE BASIS OF CONVOLUTIONAL NEURAL NETWORK

O.O. Iakovenko, N.I. Kushnirenko, I.S. Dorofieieva, A.R. Yevtushenko

Odessa National Polytechnic University,

1, Shevchenko Avenue, Odessa, 65044, Ukraine; e-mail: dorofeeva283@gmail.com

Security and integrity of personal data is an important issue in the modern world, as cases of password cracks, bank accounts and accounts, incorrect identification of individuals, falsification of personal data have recently become more frequent. Such methods of authentication as password access, use of electronic passes, one-time messages with a code are unreliable in contrast to the biometric methods. The password can be forgotten, lost, cracked or stolen, thus biometric access control systems are convenient for users because data storage devices are always with them, cannot be lost or stolen. Biometric access control is considered more reliable, because identifiers cannot be transferred to third parties or be copied. Among all the types of biometrics (identification by fingerprint, by iris, by voice, by hand geometry) we decided to choose the face recognition. After analyzing the existing methods of image classification, we have chosen the best option - a convolutional neural network, the success of which is due to the possibility of taking into account the two-dimensional image topology, in contrast to the multilayer perceptron. Face recognition technologies are used in a wide variety of areas: security in places with large concentrations of people; security systems; face control in the catering and entertainment segment, searching for suspicious and potentially dangerous visitors; verification of bank cards; online payments. We have developed a convolutional neural network, on the basis of which a post-processing method was created with the subsequent possibility of choosing a threshold, which makes it possible to manage type I and type II errors. The mathematical model of our convolutional neural network was built on the Matlab system. The initial data for our study are 840 photos, on which are depicted the authors of the article. The method developed by us is innovative and allows for improving protective complex systems.

Keywords: convolutional neural network, face recognition, machine learning, image classification, error management.