

**Юрій Ігорович Мишковський**

аспірант

Національний університет «Львівська політехніка», Львів, Україна

ORCID: 0009-0004-0051-026X

yurii.i.myshkovskiy@lpnu.ua

Марія Андріївна Назаркевич

доктор технічних наук, професор

професор кафедри інформаційних систем та мереж

Національний університет «Львівська політехніка», Львів, Україна

ORCID: 0000-0002-6528-9867

mariia.a.nazarkevych@lpnu.ua

**КЛАСИФІКАЦІЯ ЦИФРОВИХ МОДИФІКАЦІЙ БІОМЕТРИЧНИХ ЗОБРАЖЕНЬ
ДВОГІЛКОВОЮ ЗГОРТКОВОЮ НЕЙРОННОЮ МЕРЕЖЕЮ**

Анотація. Досліджено багатокласову класифікацію цифрових модифікацій відбитків пальців із набору SOCOFing на основі згорткових нейронних мереж. Порівняно базову CNN із двогілковою моделлю, яка поєднує ознаки зображення та локальні структурні дескриптори папілярних ліній. Структурну гілку сформовано на основі градієнтів Собеля і локального тензора структури; її внесок у спільне подання визначається навчуваним просторовим затвором, що враховує згорткові ознаки обох гілок та карту когерентності. Замість розгортання карти у вектор використано конкатенацію результатів глобального середнього й максимального агрегування. Розглянуто класи: Real, Obliteration, Central_rotation та Z_cut. Після вилучення сімейств зображень, пов'язаних із конфліктами міток або точними дублікатами між особами, отримано 54 978 зображень. Навчальну, валідаційну та тестову підвибірki сформовано за особами, щоб похідні одного відбитка не потрапляли до різних частин даних. Порівняння виконано за однакових параметрів оптимізації, роздільної здатності 96×96 пікселів, seed 44 та граничного бюджету 100 епох із достроковою зупинкою після семи епох без зростання валідаційної macro-F1. Навчання базової та двогілкової моделей завершилося після 45 і 29 епох відповідно; за валідаційною macro-F1 вибрано ваги з епох 38 і 22. На тестовій підвибірці з 8 172 зображень базова CNN досягла точності класифікації 94,20 % і macro-F1 92,74 %, а двогілкова — 95,06 % і 93,67 % відповідно. Кількість помилкових прогнозів зменшилася з 474 до 404, а число параметрів — з 4 738 052 до 43 174. Зменшення кількості параметрів пов'язане насамперед із заміною Flatten глобальним агрегуванням. F1-міра зросла для кожного класу, проте зміни Precision і Recall не були односпрямованими. Наведено матриці помилок, криві навчання та аналіз компромісів між показниками окремих класів. Результат характеризує один запуск кожної моделі на фіксованому розбитті й не встановлює статистичної значущості переваги. Внутрішню тестову підвибірku вже оцінювали раніше; вона не є новим зовнішнім тестом. Через одночасну зміну гілок, способу їх поєднання та класифікаційної частини окремий внесок цих компонентів потребує абляційної перевірки. Класи відповідають цифровим модифікаціям у SOCOFing; результати не є оцінкою виявлення фізичних підрбок або біометричної ідентифікації особи.

Ключові слова: згорткові нейронні мережі; багатокласова класифікація; комп'ютерний зір; глибоке навчання; виділення ознак; об'єднання ознак; аналіз зображень; кібербезпека

ВСТУП

Автоматизований аналіз зображень є складовою біометричного розпізнавання, яке застосовують для встановлення та підтвердження особи [1]. Згорткові нейронні мережі дають змогу формувати ознаки зображень у процесі навчання та використовувати їх для класифікації [2]. Для аналізу змінених зображень важливо враховувати характер порушень їхньої локальної структури [3]. У цій роботі класифікацію цифрових модифікацій розглянуто як окрему задачу аналізу візуальних даних у контексті кібербезпеки.



Дослідження зосереджено на поєднанні згорткових і структурних ознак та оцінюванні точності й компактності відповідної нейромережевої моделі.

Постановка проблеми. Аналіз відбитків пальців є складовою біометричних систем, у яких помилки опрацювання зображення можуть впливати на подальше прийняття рішення щодо особи [1]. Окремою задачею є розпізнавання змін папілярного рисунка. Дослідження змінених відбитків охоплюють аналіз характеру порушень та методи їх виявлення [3]. Водночас цифрове перетворення зображення, фізична зміна шкіри й атака з використанням штучного носія є різними об'єктами дослідження. У цій роботі розглянуто лише класифікацію цифрових модифікацій, представлених у наборі SOCOFing [4].

Згорткові ознаки дають змогу навчати класифікатор без ручного визначення кожного локального шаблону. Проте модель із шаром Flatten та великою повнозв'язною частиною може містити значно більше параметрів, ніж потрібно для компактного подання текстури. Для відбитків пальців доцільно дослідити поєднання таких ознак із локальними характеристиками напрямленості папілярних ліній. Практична задача полягає у перевірці того, чи забезпечує відповідна зміна архітектури кращі показники багатокласової класифікації за спільного протоколу навчання і тестування.

Аналіз останніх досліджень і публікацій. SOCOFing описано як набір відбитків та їх цифрово модифікованих варіантів [4]; застосування глибоких CNN до виявлення цих модифікацій досліджено у [5]. Раніше в авторській роботі розглянуто ідентифікацію відбитків на основі CNN [6]. Числові результати тієї публікації не використовуються як контрольні значення в цьому порівнянні, оскільки тут визначено іншу, чотирикласову задачу і власний протокол розбиття.

FingerNet демонструє включення спеціалізованих операцій опрацювання відбитків у нейромережеву систему виділення мінуцій [7]. Робота [8] розглядає тензор структури в контексті аналізу та вдосконалення CNN для біометричних задач. Ці праці обґрунтовують доцільність структурних ознак, однак не є експериментальними результатами запропонованої тут конфігурації. Основні принципи згорткового навчання викладено у [2], а глобальне середнє агрегування як складову нейромережевої архітектури розглянуто у [9]. Механізми навчуваного зважування ознак також мають відомі реалізації: Squeeze-and-Excitation використовує залежності між каналами [10], а CBAM поєднує каналну та просторову увагу [11].

У досліджуваній моделі просторовий затвор керує додаванням окремої структурної гілки до ознак зображення. Його вхід включає локальну когерентність, обчислену без навчуваних параметрів. Внесок роботи полягає в описі цієї прикладної конфігурації та її відтвореному порівнянні з базовою CNN на SOCOFing. Новизна окремих операцій згортки, тензора структури, глобального агрегування або просторового зважування не заявляється. Порівняння чисел з інших публікацій без узгодження класів, даних і розбиття не використовується для доведення переваги.

Мета статті. Оцінити точність і характер помилок двогілкової CNN зі структурними дескрипторами та навчуваним просторовим затвором порівняно з базовою CNN для чотирикласової класифікації цифрових модифікацій відбитків пальців за однакових умов підготовки даних і правил навчання.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Нехай I — нормалізоване зображення відбитка у градаціях сірого. Для побудови структурного подання обчислюються горизонтальна та вертикальна складові градієнта, позначені g_x і g_y , за допомогою оператора Собеля. Локальний тензор структури визначається через усереднені добутки цих складових:

$$J_{xx} = A_5(g_x^2), J_{yy} = A_5(g_y^2), J_{xy} = A_5(g_x g_y) \quad (1)$$

Тут A_5 позначає усереднення у вікні 5×5 із кроком 1 і доповненням SAME. Із компонентів тензора формуються два нормалізовані дескриптори:

$$q_1 = \frac{J_{xx} - J_{yy}}{J_{xx} + J_{yy} + \varepsilon}, \quad q_2 = \frac{2J_{xy}}{J_{xx} + J_{yy} + \varepsilon} \quad (2)$$

Константа $\varepsilon = 10^{-6}$ запобігає діленню на нуль. Третій канал — когерентність c — обчислюється як квадратний корінь із суми квадратів q_1 та q_2 з обмеженням результату інтервалом $[0; 1]$. Перед обчисленням кореня реалізація також обмежує його аргумент знизу нулем. Три канали $[q_1, q_2, c]$ утворюють структурне подання розміру $96 \times 96 \times 3$. Ці операції не використовують міток класів, не оцінюють статистик усього набору даних і не містять навчуваних ваг.

Основна гілка формує карту R розміру $24 \times 24 \times 64$ двома згортками з 32 та 64 фільтрами і двома операціями максимального агрегування. Структурна гілка послідовно використовує 16 та 32 фільтри, після чого проєктує ознаки згорткою 1×1 у карту S такої самої розмірності. Карту когерентності зменшено середнім агрегуванням 4×4 до $24 \times 24 \times 1$; отриману карту позначено $\bar{c}$. Затвор і поєднані ознаки визначаються так:

$$g = \sigma(\text{Conv}_{1 \times 1}([R, S, \bar{c}])), \quad F = R + g \odot S \quad (3)$$



Позначення $[\cdot]$ відповідає конкатенації каналів, σ — сигмоїді, а $\odot$ — поелементному множенню з поширенням одного значення затвора на всі 64 канали відповідної просторової позиції. Отже, карта g має один канал, а не окрему вагу для кожного каналу. Остаточний вектор отримується конкатенацією глобального середнього та максимального агрегування карти F :

$$z = [GAP(F), GMP(F)], p = \text{softmax}(W_2 \text{ReLU}(W_1 z + b_1) + b_2) \quad (4)$$

Вектор z містить 128 компонентів; прихований повнозв'язний шар також має 128 нейронів, вихідний — чотири нейрони. Прогнозований клас відповідає найбільшій вихідній імовірності. Базова модель використовує таку саму основну згорткову гілку, але після неї виконує Flatten і формує вектор із 36 864 компонентів. Порівняння архітектур наведено в табл. 1.

Таблиця 1

Структура порівнюваних моделей

Компонент	Базова CNN	Двогілкова CNN
Вхід	$96 \times 96 \times 1$	$96 \times 96 \times 1$
Гілка зображення	Conv 3×3 , 32 $\rightarrow$ Pool 2×2 $\rightarrow$ Conv 3×3 , 64 $\rightarrow$ Pool 2×2	Conv 3×3 , 32 $\rightarrow$ Pool 2×2 $\rightarrow$ Conv 3×3 , 64 $\rightarrow$ Pool 2×2
Структурна гілка	Відсутня	$[q_1, q_2, c] \rightarrow$ Conv 3×3 , 16 $\rightarrow$ Pool 2×2 $\rightarrow$ Conv 3×3 , 32 $\rightarrow$ Pool 2×2 $\rightarrow$ Conv 1×1 , 64
Поєднання гілок	Відсутнє	$F = R + g \odot S$; затвор $24 \times 24 \times 1$
Агрегування	Flatten $\rightarrow$ 36 864 ознаки	GAP + GMP $\rightarrow$ 128 ознак
Класифікатор	Dense 128, ReLU $\rightarrow$ Dense 4, Softmax	Dense 128, ReLU $\rightarrow$ Dense 4, Softmax
Параметри	4 738 052	43 174

У всіх прихованих згорткових шарах використано ReLU, за винятком згортки затвора із сигмоїдою. У просторових згортках 3×3 застосовано доповнення SAME. Пакетної нормалізації та Dropout у цих двох моделях немає. Зменшення кількості параметрів пов'язане насамперед із заміною великого входу повнозв'язного шару глобальним агрегуванням. Однак встановити окремий вплив цієї заміни на точність за порівнянням лише двох повних моделей неможливо.

МЕТОДИКА ДОСЛІДЖЕННЯ

Дані та класи. Початкова локальна копія SOCOFing містила 55 270 зображень від 600 осіб, зокрема 6 000 немодифікованих зображень. Для задачі використано мітки Real, Obliteration, Central_rotation та Z_cut. Останні три класи позначають категорії цифрової модифікації; рівні Easy, Medium і Hard, наявні в каталогах набору, не утворюють окремих вихідних класів. Позначення Real у цьому протоколі означає немодифіковане зображення, а не експериментально перевірену справжність фізичного біометричного зразка.

Перед розбиттям проведено перевірку точних збігів пікселів та узгодженості міток. Із подальшого використання виключено всі зображення 34 сімейств джерел, пов'язаних із конфліктними мітками або точними дублікатами між особами. Разом вилучено 292 зображення з урахуванням похідних, після чого залишилося 54 978. Очищення виконувалося до навчання за властивостями даних, а не за результатами моделей. Оригінальні файли не змінювалися; виключення реалізовано у маніфесті.

Для запобігання витоку даних між спорідненими зображеннями [12] розбиття виконано за ідентифікаторами осіб із split seed 2026. Усі пальці однієї особи та всі похідні відповідних зображень належать одній підвибірці. Частки осіб становлять 70 %, 15 % та 15 % для навчання, валідації й тестування відповідно. Після очищення точних піксельних дублікатів між підвибірками не виявлено. Склад даних наведено в табл. 2.

Таблиця 2

Розподіл зображень після очищення

Підвибірка	Особи	Real	Obliteration	Central_rotation	Z_cut	Разом
Навчальна	420	4 181	12 430	11 497	10 459	38 567
Валідаційна	90	896	2 669	2 450	2 224	8 239



Підвибірка	Особи	Real	Obliteration	Central_rotation	Z_cut	Разом
Тестова	90	889	2 654	2 412	2 217	8 172
Разом	600	5 966	17 753	16 359	14 900	54 978

Вхідні зображення перетворено у градації сірого, масштабовано білінійною інтерполяцією до 96×96 пікселів і нормалізовано до $[0; 1]$. Додаткову онлайн-аугментацію в цьому експерименті не застосовано. Використано всі зображення навчальної підвибірки, тобто $\text{train_fraction} = 1,0$. Цифрові модифікації, що вже входять до SOCOFing, є об'єктами класифікації, а не додатковою процедурою аугментації цього запуску.

Протокол навчання. Обидві мережі навчалися з випадкових початкових ваг за seed 44 із використанням Adam [13], швидкості навчання 0,001 та пакета з 64 зображень. Функція втрат — розріджена категоріальна перехресна ентропія; ваги класів не застосовувалися. Граничний бюджет становив 100 епох для кожної моделі. Правило дострокової зупинки — сім послідовних епох без строгого зростання валідаційної macro-F1 , $\text{min_delta} = 0$. Для тестування відновлювалися ваги з максимальною валідаційною macro-F1 . Отже, однаковими були ліміт і правило зупинки, а фактична кількість епох визначалася траєкторією навчання.

Обчислення виконано на CPU AMD Ryzen 7 7435HS із Python 3.12.14, TensorFlow 2.21.0 та Keras 3.15.1; у конфігурації ввімкнено детерміновані операції. Кожну архітектуру навчено один раз. Випадковість між повторними запусками є окремим джерелом варіації результатів [14], тому цей дизайн не дає оцінки середнього значення або стандартного відхилення за різними seed.

Метрики оцінювання. Основними показниками є Accurasy та macro-F1 . Для кожного класу k обчислюються Precision як частка правильних прогнозів серед усіх прогнозів цього класу та Recall як частка правильно розпізнаних зображень серед усіх зображень цього класу [15]. F1-міра поєднує ці показники, а macro-F1 надає однакову вагу всім чотирьом класам:

$$F1_k = \frac{2P_k R_k}{P_k + R_k}, \text{ macro-F1} = \frac{1}{4}(F1_1 + F1_2 + F1_3 + F1_4) \quad (5)$$

За нульового знаменника використовується нульове значення відповідної метрики. Додатково наведено macro-Precision , macro-Recall , зважену F1-міру, матриці помилок, кількість параметрів і тривалість навчального циклу. Остання включає валідацію та збереження контрольних точок і не є затримкою інференсу. Розміри файлів Keras не використовуються як оцінка розміру лише ваг, оскільки збережена модель може містити стан оптимізатора.

Простежуваність результатів. Експеримент зареєстровано з ідентифікатором 20261001T150303Z_dc981f86. Для обох моделей збережено конфігурації, маніфести, хеші даних і коду, історії епох, вибрані моделі та прогнози кожного тестового зображення. Метрики в таблицях повторно обчислено з прогнозів і звірено з підсумковими файлами. Хеші набору даних, розбиття та коду збігаються для обох запусків.

Тестова підвибірка не використовувалася для вибору епохи в описаних запусках. Це внутрішня тестова підвибірка SOCOFing, результати на якій уже аналізувалися раніше; вона не є новим незалежним зовнішнім тестом.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Зведені результати на спільній тестовій підвибірці наведено в табл. 3. Двогілкова модель отримала вищу точність класифікації та macro-F1 за меншої кількості параметрів. Приріст Accurasy становив 0,86 відсоткового пункту, а macro-F1 — 0,93 відсоткового пункту. Це різниці між двома конкретними запусками; довірчі інтервали та статистична значущість міжзапускової переваги не оцінювалися.

Таблиця 3

Підсумкові показники моделей

Показник	Базова CNN	Двогілкова CNN
Accurasy, %	94,20	95,06
macro-Precision , %	92,78	93,29
macro-Recall , %	92,72	94,10
macro-F1 , %	92,74	93,67
Зважена F1-міра, %	94,20	95,08
Помилкові прогнози	474	404



Показник	Базова CNN	Двогілкова CNN
Параметри	4 738 052	43 174
Виконані епохи / вибрана епоха	45 / 38	29 / 22
Навчальний цикл, хв	65,20	49,36

Навчання базової моделі завершилося після 45 епох, а двогілкової — після 29; для тестування вибрано епохи 38 та 22 відповідно. Максимальна валідаційна масго-F1 становила 92,94 % для базової та 92,90 % для двогілкової моделі. Отже, перевага двогілкової CNN на тесті не відтворюється як перевага її найкращої валідаційної масго-F1. Це додатково обмежує висновок про загальну перевагу архітектури.

Криві валідаційної масго-F1 та функції втрат показано на рис. 1. Критерій зупинки ґрунтувався на масго-F1, тому зміна валідаційної перехресної ентропії сама по собі не визначала вибір ваг. Різна тривалість циклів означає, що порівняння часу охоплює і відмінності архітектур, і кількість виконаних епох. За цими даними не можна робити висновок про швидкість оброблення одного зображення.

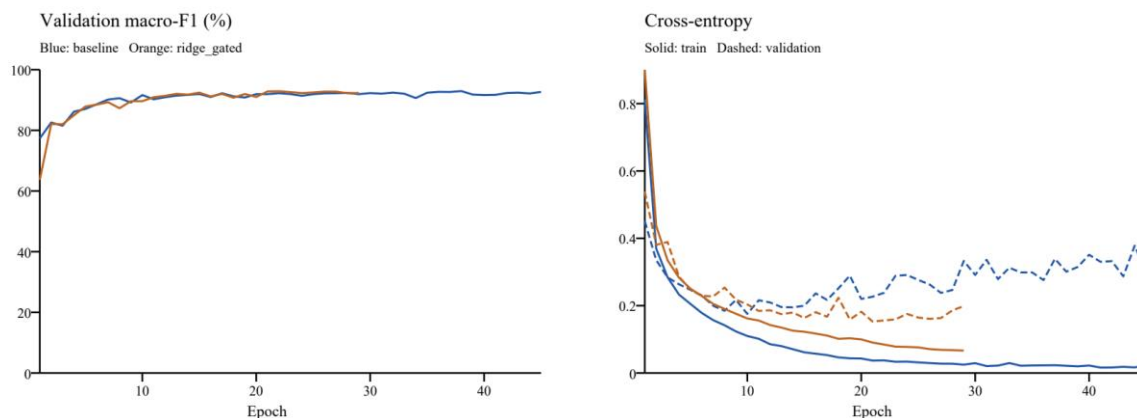


Рис. 1. Валідаційна масго-F1 та перехресна ентропія під час навчання; синя крива — базова CNN, помаранчева — двогілкова CNN, суцільні лінії втрат — навчання, штрихові — валідація

Покласові результати наведено в табл. 4. Найбільший приріст F1-міри отримано для Z_cut: з 91,24 % до 93,39 %. Для Real F1-міра зросла з 86,08 % до 87,00 %, але Precision знизилася з 86,23 % до 84,21 %, тоді як Recall зросла з 85,94 % до 89,99 %. Це означає, що модель краще знаходить немодифіковані зображення, водночас частіше відносить до Real зображення інших класів.

Таблиця 4

Метрики для окремих класів

Клас	CNN	Precision, %	Recall, %	F1, %	Зображення
Real	Базова	86,23	85,94	86,08	889
Real	Двогілкова	84,21	89,99	87,00	889
Obliteration	Базова	99,24	98,15	98,69	2 654
Obliteration	Двогілкова	98,84	99,43	99,14	2 654
Central_rotation	Базова	93,63	96,31	94,95	2 412
Central_rotation	Двогілкова	96,11	94,20	95,14	2 412
Z_cut	Базова	92,02	90,48	91,24	2 217
Z_cut	Двогілкова	94,01	92,78	93,39	2 217

Для Central_rotation Precision зросла з 93,63 % до 96,11 %, проте Recall знизилася з 96,31 % до 94,20 %. Підсумкова F1-міра цього класу змінилася лише з 94,95 % до 95,14 %. Для Obliteration також



спостерігається різноспрямована зміна Precision і Recall, хоча F1-міра зросла з 98,69 % до 99,14 %. Тому збільшення загальної точності не слід тлумачити як покращення всіх показників для кожного класу.

Матриці помилок у табл. 5 дають змогу зіставити конкретні напрями помилок. Кількість зображень Z_cut, помилково віднесених до Central_rotation, зменшилася зі 106 до 62. Для Real помилки віднесення до Z_cut зменшилися з 93 до 62. Водночас кількість Central_rotation, віднесених до Real, зросла з 14 до 53. Цей компроміс необхідно враховувати, якщо класифікатор використовується як попередній етап подальшого аналізу відбитків.

Таблиця 5

Матриці помилок: рядки — істинні класи, стовпці — прогнозовані

CNN	Істинний клас	Real	Obliteration	Central_rotation	Z_cut
Базова	Real	764	5	27	93
Базова	Obliteration	13	2605	25	11
Базова	Central_rotation	14	5	2323	70
Базова	Z_cut	95	10	106	2006
Двогілкова	Real	800	3	24	62
Двогілкова	Obliteration	7	2639	6	2
Двогілкова	Central_rotation	53	20	2272	67
Двогілкова	Z_cut	90	8	62	2057

Загальна кількість правильних прогнозів зросла з 7 698 до 7 768. У межах цього тесту це відповідає 70 додатковим правильним відповідям. Без аналізу повторних запусків та незалежного набору даних ця різниця не доводить стійкості результату до іншого випадкового початкового стану, складу осіб або способу формування модифікацій.

Межі інтерпретації. Порівняння повних архітектур одночасно змінює структурну гілку, просторовий затвор та спосіб агрегування перед класифікатором. Тому отриманий приріст не можна приписати лише тензору структури або лише затвору. Для такого висновку потрібні окремі зіставлення компактної одногілкової моделі, моделі з простим додаванням структурних ознак і варіанта затвора без карти когерентності. Ці абляційні результати в роботі не наведено.

Використані цифрові модифікації походять з одного набору даних. Оцінювання не охоплює інші сенсори, фізичні зміни папілярних ліній або атаки пред'явлення. Окремо не виміряно затримку інференсу, енергоспоживання чи кількість операцій. Менша кількість параметрів характеризує компактність навчуваної частини, але сама по собі не доводить готовності до застосування в режимі реального часу.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Описано та експериментально перевірено двогілкову CNN для класифікації цифрових модифікацій відбитків пальців. Модель поєднує згорткові ознаки зображення з дескрипторами локального тензора структури через навчуваний просторовий затвор та використовує глобальне середнє й максимальне агрегування перед класифікатором.

На очищеній версії SOCOFing із розбиттям за особами та seed 44 отримано Accurasy 95,06 % і macro-F1 93,67 % порівняно з 94,20 % і 92,74 % для базової CNN. Кількість параметрів зменшено з 4 738 052 до 43 174. F1-міра зросла для кожного з чотирьох класів, однак аналіз матриць помилок виявив компроміси між Precision і Recall, зокрема для Real та Central_rotation.

Результати підтримують доцільність подальшого дослідження компактного поєднання структурних і згорткових ознак у цій задачі. Наступні перевірки мають охоплювати повторні запуски з різними seed, абляційний аналіз компонентів, нову незалежну тестову вибірку та вимірювання часу інференсу. До виконання цих перевірок висновок обмежується описаним експериментальним протоколом і не поширюється на фізичну біометричну автентифікацію.

Декларація про використання генеративного штучного інтелекту. Для опрацювання артефактів експерименту, пошуку бібліографічних відомостей і формування рукопису використано OpenAI Codex. Числові результати взято зі збережених запусків і програмно звірено з прогнозами окремих зображень; Всі розділи перевірені та відредаговані автором особисто. Висновки та результати є власним інтелектуальним внеском

**ПОДЯКИ**

Дослідження здійснено завдяки грантовій підтримці Національного фонду досліджень України за проектом «Методи аналізу та оптимізації мультимодальних даних для моделей глибинного навчання у військовій сфері», реєстраційний номер проекту 2025.07/0017 від 24/12/2025.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20. <https://doi.org/10.1109/TCSVT.2003.818349>
2. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
3. Yoon, S., Feng, J., & Jain, A. K. (2012). Altered fingerprints: Analysis and detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(3), 451–464. <https://doi.org/10.1109/TPAMI.2011.161>
4. Shehu, Y. I., Ruiz-Garcia, A., Palade, V., & James, A. (2018). *Sokoto Coventry Fingerprint Dataset* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.1807.10609>
5. Shehu, Y. I., Ruiz-Garcia, A., Palade, V., & James, A. (2018). Detection of fingerprint alterations using deep convolutional neural networks. In V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, & I. Maglogiannis (Eds.), *Artificial neural networks and machine learning – ICANN 2018* (Lecture Notes in Computer Science, Vol. 11139, pp. 51–60). Springer. https://doi.org/10.1007/978-3-030-01418-6_6
6. Мишковський, Ю., & Назаркевич, М. (2024). Метод ідентифікації відбитків пальців на основі згорткових нейронних мереж. *Інформаційні системи та мережі*, (15), 1–14. <https://doi.org/10.23939/sisn2024.15.001>
7. Tang, Y., Gao, F., Feng, J., & Liu, Y. (2017). FingerNet: An unified deep network for fingerprint minutiae extraction. In *2017 IEEE International Joint Conference on Biometrics (IJCB)* (pp. 108–116). IEEE. <https://doi.org/10.1109/BTAS.2017.8272688>
8. Hernandez-Diaz, K., Bigun, J., & Alonso-Fernandez, F. (2024). *Understanding and improving CNNs with complex structure tensor: A biometrics study* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2404.15608>
9. Lin, M., Chen, Q., & Yan, S. (2014). Network in network. In *2nd International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1312.4400>
10. Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2020). Squeeze-and-excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(8), 2011–2023. <https://doi.org/10.1109/TPAMI.2019.2913372>
11. Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Eds.), *Computer vision – ECCV 2018* (Lecture Notes in Computer Science, Vol. 11211, pp. 3–19). Springer. https://doi.org/10.1007/978-3-030-01234-2_1
12. Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, 4(9), Article 100804. <https://doi.org/10.1016/j.patter.2023.100804>
13. Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1412.6980>
14. Bouthillier, X., Delaunay, P., Bronzi, M., Trofimov, A., Nichyporuk, B., Szeto, J., Mohammadi Sepahvand, N., Raff, E., Madan, K., Voleti, V., Ebrahimi Kahou, S., Michalski, V., Arbel, T., Pal, C., Varoquaux, G., & Vincent, P. (2021). Accounting for variance in machine learning benchmarks. *Proceedings of Machine Learning and Systems*, 3, 747–769. https://proceedings.mlsys.org/paper_files/paper/2021/hash/0184b0cd3cfb185989f858a1d9f5c1eb-Abstract.html
15. Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>

**Yurii I. Myshkovskyi**

PhD student

Lviv Polytechnic National University, Lviv, Ukraine

ORCID ID: 0009-0004-0051-026X

yurii.i.myshkovskyi@lpnu.ua

Mariia A. Nazarkevych

Doctor of Technical Sciences, Professor

Professor of the Department of Information Systems and Networks

Lviv Polytechnic National University, Lviv, Ukraine

ORCID ID: 0000-0002-6528-9867

mariia.a.nazarkevych@lpnu.ua

CLASSIFICATION OF DIGITAL ALTERATIONS IN BIOMETRIC IMAGES USING A TWO-BRANCH CONVOLUTIONAL NEURAL NETWORK

Abstract. This study investigates multiclass classification of digitally altered fingerprint images from SOCOFing using convolutional neural networks. The objective is to compare a baseline CNN with a two-branch model that combines image features and local descriptors of fingerprint ridge structure. The structural branch uses Sobel gradients and a local structure tensor. Its contribution to the fused representation is controlled by a trainable spatial gate conditioned on both feature streams and a coherence map. Global average and maximum pooling replace spatial flattening in the classification head. Four classes are considered: Real, Obliteration, Central_rotation, and Z_cut. Before splitting the data, source-image families associated with conflicting labels or exact duplicates across subjects were excluded, leaving 54,978 images. Training, validation, and test subsets were formed by subject so that derivatives of the same fingerprint remained in a single subset. Both models used 96×96 grayscale inputs, identical optimization settings, random seed 44, and a maximum budget of 100 epochs. Early stopping was triggered after seven epochs without an improvement in validation macro-F1. On the shared test subset of 8,172 images, the baseline achieved 94.20% classification accuracy and 92.74% macro-F1, while the two-branch model achieved 95.06% and 93.67%, respectively. The number of incorrect predictions decreased from 474 to 404, and the parameter count decreased from 4,738,052 to 43,174. Class-specific F1 improved for all four classes, although precision and recall did not improve uniformly. Confusion matrices and learning curves are reported to show both gains and error trade-offs. Training stopped after 45 epochs for the baseline and 29 epochs for the modified model; checkpoints were selected at epochs 38 and 22 using validation macro-F1. The results describe one run per model on a fixed split and do not establish statistical significance or robustness across random seeds. Because the feature streams, fusion mechanism, and pooling head were changed together, their individual contributions cannot be determined without ablation experiments. The internal SOCOFing test split had been evaluated previously and does not constitute a newly reserved external evaluation set. The task concerns digital alteration categories in SOCOFing rather than physical presentation-attack detection or fingerprint identity recognition. The findings support further evaluation of compact structural feature fusion under repeated runs and independent test conditions.

Keywords: convolutional neural networks; multiclass classification; computer vision; deep learning; feature extraction; feature fusion; image analysis; cybersecurity

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20. <https://doi.org/10.1109/TCSVT.2003.818349>
2. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
3. Yoon, S., Feng, J., & Jain, A. K. (2012). Altered fingerprints: Analysis and detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(3), 451–464. <https://doi.org/10.1109/TPAMI.2011.161>
4. Shehu, Y. I., Ruiz-Garcia, A., Palade, V., & James, A. (2018). *Sokoto Coventry Fingerprint Dataset* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.1807.10609>



5. Shehu, Y. I., Ruiz-Garcia, A., Palade, V., & James, A. (2018). Detection of fingerprint alterations using deep convolutional neural networks. In V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, & I. Maglogiannis (Eds.), *Artificial neural networks and machine learning – ICANN 2018* (Lecture Notes in Computer Science, Vol. 11139, pp. 51–60). Springer. https://doi.org/10.1007/978-3-030-01418-6_6
6. Myshkovskiy, Yu., & Nazarkovich, M. (2024). Metod identyfikatsii vidbytkiv paltsiv na osnovi zghortkovykh neironnykh merezh [Fingerprint identification method based on convolutional neural networks]. *Informatsiini systemy ta merezhi*, (15), 1–14. <https://doi.org/10.23939/sisn2024.15.001>
7. Tang, Y., Gao, F., Feng, J., & Liu, Y. (2017). FingerNet: An unified deep network for fingerprint minutiae extraction. In *2017 IEEE International Joint Conference on Biometrics (IJCB)* (pp. 108–116). IEEE. <https://doi.org/10.1109/BTAS.2017.8272688>
8. Hernandez-Diaz, K., Bigun, J., & Alonso-Fernandez, F. (2024). *Understanding and improving CNNs with complex structure tensor: A biometrics study* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2404.15608>
9. Lin, M., Chen, Q., & Yan, S. (2014). Network in network. In *2nd International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1312.4400>
10. Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2020). Squeeze-and-excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(8), 2011–2023. <https://doi.org/10.1109/TPAMI.2019.2913372>
11. Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Eds.), *Computer vision – ECCV 2018* (Lecture Notes in Computer Science, Vol. 11211, pp. 3–19). Springer. https://doi.org/10.1007/978-3-030-01234-2_1
12. Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, 4(9), Article 100804. <https://doi.org/10.1016/j.patter.2023.100804>
13. Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1412.6980>
14. Bouthillier, X., Delaunay, P., Bronzi, M., Trofimov, A., Nichyporuk, B., Szeto, J., Mohammadi Sepahvand, N., Raff, E., Madan, K., Voleti, V., Ebrahimi Kahou, S., Michalski, V., Arbel, T., Pal, C., Varoquaux, G., & Vincent, P. (2021). Accounting for variance in machine learning benchmarks. *Proceedings of Machine Learning and Systems*, 3, 747–769. https://proceedings.mlsys.org/paper_files/paper/2021/hash/0184b0cd3cfb185989f858a1d9f5c1eb-Abstract.html
15. Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>

Отримано редакцією журналу / Received: 30.05.26

Прорецензовано / Revised: 16.06.26

Схвалено до друку / Accepted: 26.10.26

