

Т. В. Дем'янець, Д. В. Федасюк

Національний університет "Львівська політехніка", м. Львів, Україна

ЗАСТОСУВАННЯ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ВИЯВЛЕННЯ МЕЛАНОМИ ЗА ЗОБРАЖЕННЯМ НОВОУТВОРЕННЯ НА МОБІЛЬНОМУ ПРИСТРОЇ

Розглянуто сучасні методи для виявлення меланоми за зображенням новоутворення, проаналізовано їх переваги та недоліки. У статті продемонстровано прототип мобільного застосунку для виявлення меланоми за зображенням родимки на підставі згорткової нейронної мережі, котрий розробляється для операційної системи Android. Доповнено набір даних на підставі HAM10000 для навчання задля усунення незбалансованості класів та покращення показників точності мережі. Проведено пошук наявних нейронних мереж, котрі забезпечують високу точність, та обрано для дослідження нейромережі VGG16, MobileNet та NASNetMobile. До наведених нейронних мереж застосовано техніки передавання навченості та точного налаштування для адаптування мереж до задачі класифікації уражень шкіри. Встановлено, що використання згаданих технік дає змогу отримати високі показники точності нейронної мережі для вказаної задачі. Описано процес конвертування згорткової нейронної мережі в оптимізований формат Flatbuffer засобами TensorFlow Lite для розміщення та використання на мобільному пристрої. Оцінено характеристики швидкодії обраних нейронних мереж на мобільному пристрої за часом класифікації на центральному і графічному процесорах та проведено порівняння обсягу пам'яті, котру займає файл окремої мережі. Проведено порівняння розміру файлу нейронної мережі до конвертування та після. Продемонстровано, що використання конвертера TensorFlow Lite значно зменшує розмір файлу нейронної мережі шляхом використання оптимізованого формату. Результати дослідження свідчать про високу швидкість у застосунку та компактність мереж на пристрої, а використання графічного пришвидшення дає змогу значно скоротити тривалість класифікації зображення новоутворення. На підставі проаналізованих параметрів обрано NASNetMobile як оптимальну нейронну мережу для застосування у мобільному застосунку виявлення меланоми.

Ключові слова: Android додаток; зображення новоутворень; штучний інтелект; машинне навчання; налаштування нейронної мережі.

Вступ

Меланома – це злоякісне утворення на шкірі (рідше на слизових оболонках або сітківці ока), що розвивається з клітин, званих меланоцитами. Меланома становить тільки 4 % від всіх видів раку шкіри, проте 80 % смертей від раку шкіри спричинені саме меланою. Ознаками розвитку меланоми є зміна кольору родимки, збільшення у розмірах, зміна форми, зміна поверхні родимки (можуть появлятися горбистість, кровоточивість). Найсильнішими факторами ризику є сімейна історія меланоми, множинні доброякісні або нетипові невуси. Також вразливими є люди зі світлою шкірою, вплив ультрафіолетового випромінювання та люди, старші 50 років [9].

Для профілактики меланоми та самодіагностування існують критерії ABCDE [1], які дають змогу обстежити родимку на ознаки асиметрії (англ. *Asymmetry*), нерівності меж (англ. *Border irregularity*), зміна кольору (англ. *Color variegation*), діаметр родимки понад 6 мм (англ. *Diameter*) та еволюція родимки (англ. *Evolving*), а саме ріст родимки, поява кірочок, тріщин. Використання критеріїв дає змогу виявити меланому на ранніх стадіях та вчасно отримати лікування.

Проблема раннього виявлення меланоми є надзвичайно важливою для подальшого прогнозу лікування, тому на сьогодні існує низка рішень для її виявлення, які базуються на різних підходах. З таких рішень є алгоритми, засновані на ABCDE правилі [8], нейронні ме-

режі різних типів [11], [15], [21], опорно-векторні машини [3].

За останні роки обчислювальна потужність смартфонів зросла і це дозволило запускати потужні алгоритми машинного навчання на мобільних пристроях. Тому актуальним є використання вищеприказаних алгоритмів на мобільному пристрої, а саме згорткової нейронної мережі, адже цей тип мереж показує високу точність у задачах класифікації зображень.

Об'єкт дослідження – класифікація зображень новоутворень, отриманих смартфоном за допомогою штучного інтелекту.

Предмет дослідження – моделі на підставі штучних нейронних мереж та мобільний додаток для класифікації зображень новоутворень.

Мета роботи – дослідження конфігурацій згорткових нейронних мереж з застосуванням технік transfer learning та тонкого налаштування для вирішення задачі класифікації уражень шкіри на мобільному пристрої.

Для досягнення поставленої мети необхідно виконати наступні завдання дослідження:

- проаналізувати конфігурації згорткових нейронних мереж;
- застосувати техніки передавання навченості та тонкого налаштування нейронних мереж;
- оцінити показники точності класифікації;
- конвертувати файли мереж за допомогою TensorFlow Lite для використання їх в мобільному застосунку виявлення меланоми;

- оцінити швидкодію процесу класифікації на центральному процесорі та з використанням графічного процесора у мобільному застосунку.

Наукова новизна отриманих результатів – запропоновано модель на підставі згорткової мережі, яка описує процес конвертування, використання та розміщення нейромережі для виявлення меланоми за допомогою мобільного пристрою.

Практична значущість результатів дослідження – можливість застосування досліджених нейронних мереж для завдання класифікації зображень утворень на шкірі з допомогою мобільного пристрою за рахунок прийнятної швидкодії та достовірності виявлення меланоми.

Матеріали та методи дослідження. Для дослідження меланоми за зображенням новоутворення використано нейронні мережі, зокрема згорткові нейронні мережі, навчання яких здійснювалось на підставі відкритих вхідних наборів даних зображень меланом і родимок. Для конвертування нейронної мережі в формат Flatbuffer на мобільний пристрій використано конвертор TensorFlowLite, що забезпечило зменшення розміру файлу нейронної мережі і функціонування застосунку на мобільному пристрої.

Тестування та дослідження мобільного застосунку проведено на мобільному пристрої Xiaomi Mi 9 Life з 8 ядрами, оперативною пам'яттю 6 ГБ та графічним процесором Adreno 616 з частотою 750 МГц.

Аналіз останніх досліджень та публікацій. Сучасний стан проблеми та методи раннього виявлення меланоми розглянуто в дослідженні [14], де вказані переваги та недоліки, а також оцінено точність цих рішень. Розглянуто реалізацію ABCD правила для витягнення ознак з фото родимки, діагностику на підставі ABCD правила для витягнення ознак та використання нейронної мережі для класифікації цих ознак на дві категорії: меланома або доброякісне утворення, а також проаналізовано згорткові нейронні мережі та виявлення меланоми за допомогою опорно-векторних машин. Найвищі результати точності показують опорно-векторні машини (80-90 %) та згорткові нейронні мережі (60-75 %).

Suleiman Mustafa та інші [10] продемонстрували застосування опорно-векторних машин з використанням ABCD правила та алгоритму GrabCut для сегментації ураження у задачі класифікації ураження на злоякісне та доброякісне. Отримано показники точності 80 %, чутливості 71.4 % та специфічності 55.36 %.

Виявлення меланоми за допомогою згорткових мереж демонструється у дослідженні, де авторам вдалось отримати значення чутливості 81 % та специфічності 80 %. У дослідженні запропоновано впровадження системи глибокого навчання на сервері, оснащеному блоком графічного оброблення (GPU). Система використовує клінічні (недерматоскопічні) зображення, до яких попередньо застосовуються алгоритми усунення надмірного освітлення та шумів, а далі зображення передаються на вхід нейронної мережі [11].

У наступному дослідженні продемонстрована конфігурація згорткової нейронної мережі, котра складається з 3 шарів згортки, 3 шарів агрегації та 4 пов'язаних шарів. Навчання проводилось у 25 епох та кінцеве значення точності мережі склало 74 % [12].

Ще одним аспектом дослідження є розміщення та запуск нейронної мережі на пристрої. Andrej Ignatov та

інші [7] у дослідженні розглядають сучасний стан глибокого навчання в екосистемі Android, описуються наявні фреймворки, моделі та обмеження щодо використання штучного інтелекту на смартфонах. Також автори представили AI Benchmark для оцінки продуктивності у різних задачах, пов'язаних з запуском глибоких нейронних мереж: класифікація, розпізнавання облич, видалення розмиття на зображенні, сегментація та інші. Застосунок використовує TensorFlow Lite для запуску нейронних мереж. Результати згруповано за моделями смартфонів, системами на кристалі (англ. *System on a chip*), наприклад Qualcomm Snapdragon 845, Samsung Exynos 8895 Octa та інші. Нарахування загальної кількості балів здійснено шляхом зваженої суми балів, набраних у 9 тестах. Вагові коефіцієнти відкалібровані на підставі результатів Google Pixel 2.

На підставі згаданого вище дослідження, для конвертування та взаємодії з нейронною мережею на мобільному застосунку обрано TensorFlow Lite. TensorFlow Lite – це набір інструментів, що допомагає розробникам запускати моделі TensorFlow на мобільних, вбудованих та пристроях інтернету речей [19]. Ключовими особливостями є підтримка різних платформ, висока продуктивність шляхом використання апаратного пришвидшення, інструменти оптимізації моделей, використання ефективного формату моделі з використанням FlatBuffer, котрий оптимізований для невеликого розміру та портативності.

Такий підхід до зберігання та використання моделі дасть змогу використовувати застосунок без підключення до Інтернету, зменшити затримки за рахунок відсутності потреби у надсиланні даних та отримання їх з сервера та підвищить конфіденційність, адже дані залишатимуться на пристрої.

Виходячи з сказаного вище, для дослідження обрано згорткову нейронну мережу, адже ці мережі демонструють хорошу точність у задачі класифікації зображень [18] та можуть використовувати розпаралелення обчислень на графічному процесорі для проведення навчання. Мережі такого типу проявляють відносну стійкість до зсуву об'єкту на зображенні за рахунок проходження по всьому зображенню, а саме положення родимки на зображенні.

Результати дослідження та їх обговорення

Для отримання високих показників точності нейронної мережі, необхідно навчити її на дуже великому наборі даних, тому для випадку, коли набір даних не є достатньо великим, використовується техніка машинного навчання передавання навченості нейронної мережі. Техніка дає змогу перенести досвід попередньо навченої нейронної мережі на великому наборі даних для однієї задачі на розв'язання іншої задачі [6].

Використання техніки передавання навченості для виявлення акральної меланоми продемонстровано у дослідженні [21]. Автори використовують попередньо навчену на підставі бази даних ImageNet нейронну мережу VGG-16, яка складається з 16 шарів (13 згорткових та 3 пов'язаних). Завдяки техніці Transfer learning, нейронну мережу вдалось донавчити для класифікації уражень на дві категорії (меланома або не меланома). У цьому дослідженні отримано значення чутливості 92 % та значення специфічності 75 %. Схожі резуль-

Для навчання нейронної мережі використано набір вхідних даних HAM10000, котрий налічує 10 тисяч зображень утворень, поділених на 7 категорій діагнозів: актинічний кератоз, базаліома, себореїний кератоз, дерматофіброма, невус, меланома, судинні ураження шкіри [20]. HAM10000 містить файл метаданих, де надано інформацію про зображення утворення: назва зображення, тип ураження, вік, стать пацієнта та локалізація ураження. Для дослідження відібрано зображення уражень з двох категорій: меланома та невус.

Аналіз характеристик нейронних мереж. Для визначення характеристик нейронних мереж використано матрицю невідповідностей та метрики точності, повноти, чутливості, специфічності, f_1 -метрика [13]. Матриця невідповідностей має такий вигляд (табл. 1). З матриці відповідності можна розрахувати обрані метрики.

		Справжній клас	
		Позитивний	Негативний
Прогнозований клас	Позитивний	Істинно позитивний (ІП)	Хибно позитивний (ХП)
	Негативний	Хибно негативний (ХН)	Істинно Негативний (ІН)


$$\text{Точність} = \frac{III + IH}{III + XPI + XH + IH}$$
$$\text{Влучність}(\textit{precision}) = \frac{II}{II + XII}$$
$$Повнота(recall) = \frac{ПП}{ПП + ХН} = Чутливість$$
$$\text{Специфічність} = \frac{IH}{XII + IH}$$
$$F1-метрика = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

Згорткову нейронну мережу VGG16 запропонували K. Simonyan та A. Zisserman у роботі [17] для змагання ImageNet Challenge 2014. Кількість параметрів мережі становить 138.2 млн, а top-1 точність рівна 71.3 % та top-5 точність 90 %. Конфігурація містить 16 шарів (13 шарів згортки та 3 повноз'язні шари). Розмір фільтру є сталим та становить 3×3 (рис. 1,а). Для адаптації нейронної мережі до задачі виявлення меланоми, класифікатор представленої нейронної мережі замінено на наступний: перший повноз'язний шар зі 128 нейронами, другий повноз'язний шар з 256 нейронами та вихідний шар з двома вихідними нейронами.

У продемонстрованій вище (рис. 1,б) конфігурації кількість параметрів становить 16 386 754. Заморожено згорткові шари нейронної мережі, а саме у процесі навчання, параметри у цих шарах не будуть оновлюватись та проведено навчання класифікатора. Навчання мережі складало 30 епох з розміром підвибірки 32, оскільки набір даних для навчання відрізняється сильно від набору даних, на котрому навчалась нейронна мережа VGG16.

Наступним до мережі застосовано техніку fine-tuning, тобто розморожено всі згорткові шари та проведено навчання всієї конфігурації згорткової мережі тривалістю 15 епох. Порівняння метрик після застосування техніки transfer learning та fine-tuning продемонстровано у табл. 2.

Табл. 2. Порівняння метрик після застосування техніки передавання навченості та техніки точного налаштування до нейронної мережі VGG16

Метрика	Точність класифікації	Точність (precision)	Повнота, чутливість	Специфічність	F1-метрика
Передавання навченості	78 %	68 %	77 %	85 %	72 %
Точне налаштування	86 %	82 %	78 %	88 %	80 %

З таблиці видно, що використання передавання навченості у поєднанні з тонким налаштуванням нейронної мережі дає змогу покращити показники точності VGG16. Наступною мережею для дослідження обрано мобільну нейронну мережу MobileNetV2, котру представили Sandler M. та інші [16]. Така конфігурація при значеннях $\alpha=1$ та $\rho=1$ демонструє top-1 точність 72 % та top-5 точність 92.2 % і містить 3 538 984 параметри. Аналогічно до VGG16, модель завантажено без класифікатора та додано власний, котрий складається з шару global average pooling та виходу з 2 нейронами: меланома або невус. Навчання з новим класифікатором проводилось у 20 епох з розміром підвибірки 32. Розмірність вхідного зображення становить 160×160 пікселів, а кількість каналів становить 3 (червоний, зелений, синій). Отримано наступні показники метрик: точність класифікації 67 %, влучність 52 %, чутливість 73 %, специфічність 80 % та F1-метрика 61 %.

Табл. 3. Порівняння показників метрик навчання нейронної мережі NASNetMobile

Метрика	Точність класифікації	Влучність (precision)	Повнота, чутливість	Специфічність	F1-метрика
Передавання навченості	70.9 %	58 %	67 %	79.7 %	62 %
Точне налаштування	81.6 %	75 %	74 %	85.5 %	74 %

Застосування техніки тонкого налаштування мережі не дало змоги покращити показники нейронної мережі. NASNetMobile – це ще одна назва нейронної мережі, котра призначена для застосування в мобільних пристроях, продемонстрована у дослідженні Zoph B. та інших [22]. Класифікатор мережі замінено на наступний: шар global average pooling, перший повнозв'язний шар з 256 нейронами, шар регуляризації, другий повнозв'язний шар зі 128 нейронами, шар регуляризації, третій повнозв'язний шар з 64 нейронами та вихідний шар з 2 нейронами, котрі відповідають конкретному класу

(меланома або невус). Навчання класифікатора тривало 30 епох з розміром підвибірки 32. Застосування тонкого налаштування після донавчання мережі з класифікатором показало значне покращення метрик нейронної мережі.

Метод конвертування нейронної мережі та запропонований мобільний застосунок. Наступним кроком після проведення перенесення навчання та тонкого налаштування мережі, необхідно конвертувати мережі для використання на мобільному пристрої з допомогою TensorFlow Lite.



Рис. 2. Схема процесу конвертування нейронної мережі та додавання метаданих

Конвертер TensorFlow Lite отримує файл моделі TensorFlow або Keras та генерує модель TensorFlow Lite в оптимізованому форматі FlatBuffer. Після цього процесу важливим є додавання метаданих до моделі для спрощення генерації коду на пристроях, де буде використовуватись нейронна мережа.

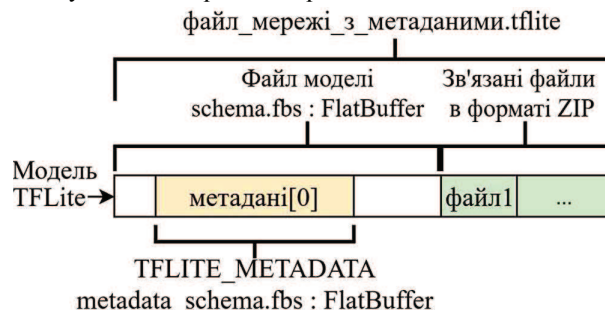


Рис. 3. Модель нейронної мережі з метаданими та зв'язаними файлами

Як видно з рисунка метадані моделі містяться в полі метаданих під назвою TFLITE_METADATA. Деякі моделі можуть містити зв'язані файли та вони приєднуються в кінець файлу моделі у вигляді ZIP-архіву. Такими файлами можуть бути файли класифікаційних міток, в яких міститься номер мітки та назва класу [2].

Прототип мобільного застосунку для виявлення меланоми. Застосунок розроблено для ОС Android, а нейронна мережа розміщується на пристрої користувача. Функціями мобільного додатку є створення нового обстеження, перегляд попередніх обстежень, наявність галереї зображень з обстежень з мітками локалізації для відстеження динаміки росту новоутворення. Додаток буде повністю автономним та не вимагатиме підключення до мережі Інтернет, адже нейронна мережа для класифікації розміщуватиметься на пристрої користувача. Для більш наочного розуміння застосунку, на рисунку у наведено діаграму випадків використання.



Рис. 4. UML-діаграма випадків використання

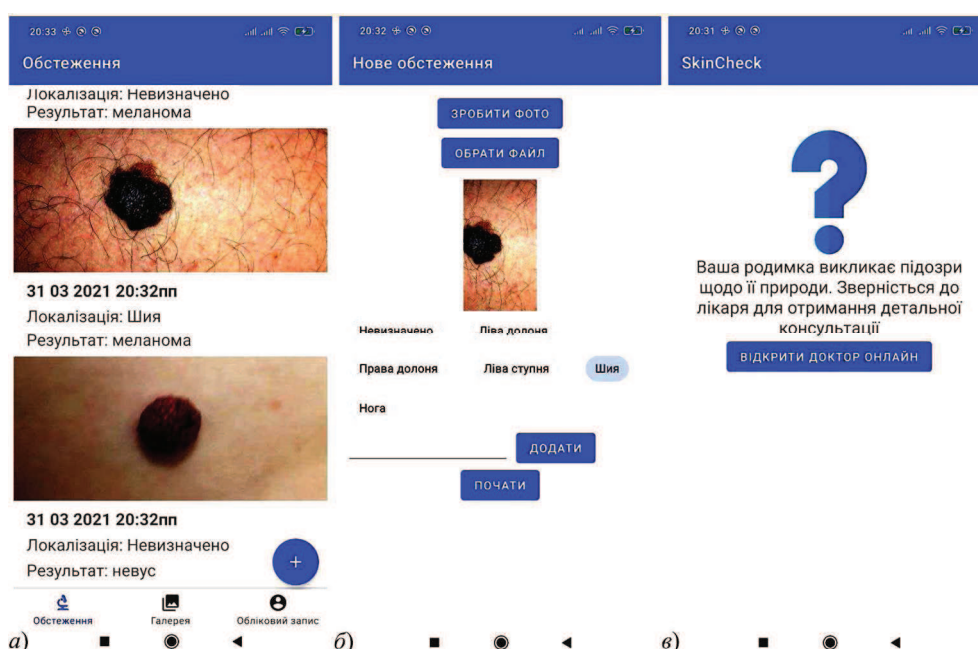


Рис. 5. Інтерфейс екрану "Мої обстеження" (а) та "Додавання нового обстеження" (б), скріншот екрану у разі виявлення меланому (в)

На рис. 5 наведено скріншоти інтерфейсу додатку для виявлення меланоми та у разі виявлення меланоми є можливість відкрити застосунок "Доктор Онлайн" для отримання подальшої консультації.

Спосіб зберігання та використання мережі на пристрої користувача забезпечується шляхом конвертації повноцінної моделі нейронної мережі в оптимізовану для конкретного пристрою шляхом використання конвертера TensorFlow Lite. Ще однією перевагою є зменшення затримки отримання результату класифікації, адже не потрібно відправляти дані на сервер та чекати їх отримання.

Обговорення результатів дослідження. Проведено оцінювання параметрів продуктивності процесу класифікації зображень з використанням нейронних мереж на мобільному пристрої.

Оцінка проводилась на мобільному пристрої Xiaomi Mi 9 Lite, оснащеного процесором Qualcomm Snapdragon 710 з 8 ядрами (6 ядер Kryo 360 [Cortex-A55] з частотою до 1.7 ГГц та 2 ядра Kryo 360 [Cortex-A75] з час-

тотую до 2.2 ГГц), 6 ГБ оперативної пам'яті та графічним процесором Adreno 616 з частотою 750 МГц.

Як параметри для оцінювання використано: обсяг пам'яті, котру займає модель, тривалість класифікації з використанням процесора та тривалість класифікації з використанням графічної карти.

Табл. 4. Оцінювання параметрів запропонованих мереж в мобільному застосунку

Нейронна мережа	VGG16	Mobile-NetV2	NASNet-Mobile
Розмір файлу моделі до конвертації (КБ)	192 166	9 144	51 149
Розмір файлу моделі після конвертації (КБ)	64 023	8 672	17 897
Тривалість класифікації з використанням центрального процесора (мс)	1148	86	191
Час класифікації з використанням графічного процесора (мс)	229	21	-

Важливим параметром для мобільного пристрою є обсяг пам'яті, котрий може займати застосунок на пристрої користувача, адже розмір файлу нейронної мережі прямо впливає на обсяг пам'яті, котру буде займати застосунок. З табл. 4 видно, що конвертація дає змогу значно зменшити розмір файлу нейронної мережі. Наприклад, файл мережі VGG16 після конвертації зменшився втричі, NASNetMobile у 2.85 рази, проте майже незмінним залишився розмір файлу мережі MobileNetV2. Тривалість класифікації є наступним важливим параметром саме тому, що зменшення затримок у застосунку позитивно впливає на досвід його використання.

Виходячи з отриманих результатів, встановлено, що найшвидшою за часом класифікації та найменшою за розміром мережею є MobileNetV2. Однак, отримані показники чутливості та специфічності є найнижчими серед зазначених моделей. Нейронна мережа VGG16 демонструє найкращу точність з представлених, проте розмір її файлу є найбільшим. Варто відзначити, що використання графічного процесора пришвидшує тривалість класифікації у 4-5 разів.

Для мережі NASNetMobile не вдалось оцінити тривалість класифікації з використанням графічного прискорення, проте її показники точності, розміру та часу класифікації є оптимальними для використання на мобільному пристрої для задачі класифікації уражень шкіри.

Висновки

Запропоновано метод, який полягає у використанні техніки передавання навченості, налаштуванні, конвертуванні, зберіганні та використанні нейромережі для задачі виявлення меланоми за зображенням родимки на мобільному пристрої. Таке рішення дає змогу покращити швидкодію та автономність додатку шляхом розміщення файлу мережі безпосередньо на смартфоні.

Проведено аналіз наявних нейронних мереж та застосовано техніки transfer learning і fine tuning для використання їх у задачі класифікації уражень шкіри. Оцінено метрики класифікації та встановлено, що використання технік передавання навченості та точного налаштування дає змогу отримати високі показники точності нейронної мережі, чутливості, специфічності.

Конвертовано обрані нейронні мережі для використання їх у застосунку. Обчислено показники продуктивності та визначено, що оптимальною нейронною мережею за критеріями точності, розміру файлу та часу на очікування результату класифікації є NASNetMobile.

Реалізовано прототип мобільного застосунку для виявлення меланоми за зображенням новоутворення з використанням конвертованої та розміщеної на пристрої користувача нейронної мережі, який підтверджує припущення щодо можливого застосування нейронних мереж у мобільних пристроях для даної задачі.

References

- [1] Abbasi, N. R., Shaw, H. M., Rigel, D. S., Friedman, R. J., McCarthy, W. H., Osman, I., Kopf, A. W., & Polsky, D. (2004). Early Diagnosis of Cutaneous Melanoma. *JAMA*, 292(22), 2771–2776. <https://doi.org/10.1001/jama.292.22.2771>
- [2] Adding metadata to TensorFlow Lite models. (2021). TensorFlow. <https://www.tensorflow.org/lite/guide>
- [3] Bakheet, S. (2017). An SVM Framework for Malignant Melanoma Detection Based on Optimized HOG Features.

- Computation*, 5(4), 4. <https://doi.org/10.3390/computation5010004>
- [4] Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., Smith, J.R. (2015). Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images. In: Zhou L., Wang L., Wang Q., Shi Y. (Eds) Machine Learning in Medical Imaging. MLMI 2015. *Lecture Notes in Computer Science*, vol. 9352. Springer, Cham. https://doi.org/10.1007/978-3-319-24888-2_15
- [5] Deng, J., Dong, W., Socher, R., Li, L. J., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition, 248–255. <https://doi.org/10.1109/cvpr.2009.5206848>
- [6] Hussain, M., Bird, J. J., Faria, D. R. (2019). A Study on CNN Transfer Learning for Image Classification. In: Lotfi A., Bouchachia H., Gegov A., Langensiepen C., McGinnity M. (Eds) Advances in Computational Intelligence Systems. UKCI 2018. *Advances in Intelligent Systems and Computing*, vol. 840. Springer, Cham. https://doi.org/10.1007/978-3-319-97982-3_16
- [7] Ignatov, A., et al. (2019). AI Benchmark: Running Deep Neural Networks on Android Smartphones. In: Leal-Taixé L., Roth S. (Eds.) Computer Vision – ECCV 2018 Workshops. ECCV 2018. *Lecture Notes in Computer Science*, vol. 11133. Springer, Cham. https://doi.org/10.1007/978-3-030-11021-5_19
- [8] Kasmi, R., & Mokrani, K. (2016). Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule. *IET Image Processing*, 10(6), 448–455. <https://doi.org/10.1049/iet-ipr.2015.0385>
- [9] Miller, A. J., & Mihm, M. C. (2006). Melanoma. *New England Journal of Medicine*, 355(1), 51–65. <https://doi.org/10.1056/nejmra052166>
- [10] Mustafa, S., Dauda, A. B., & Dauda, M. (2017). Image processing and SVM classification for melanoma detection. 2017 International Conference on Computing Networking and Informatics (ICCNi), 1–5. <https://doi.org/10.1109/iccn.2017.8123777>
- [11] Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S., Jafari, M., Ward, K., & Najarian, K. (2016). Melanoma detection by analysis of clinical images using convolutional neural network. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). <https://doi.org/10.1109/embc.2016.7590963>
- [12] Ottom, M. A. (2019). Convolutional neural network for diagnosing skin cancer. *International Journal of Advanced Computer Science and Applications*, 10(7). <https://doi.org/10.14569/ijacsa.2019.0100746>
- [13] Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63. <https://doi.org/10.9735/2229-3981>
- [14] Raut, N., Shah, A., Vira, S., & Sampat, H. (2018). A study on different techniques for skin cancer detection. *International Research Journal of Engineering and Technology (IRJET)*, 5(3), 613–617.
- [15] Refianti, R., Benny, A., & Poetri, R. (2019). Classification of Melanoma Skin Cancer using Convolutional Neural Network. *International Journal of Advanced Computer Science and Applications*, 10(3), 409–417. <https://doi.org/10.14569/ijacsa.2019.0100353>
- [16] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [17] Simonyan, K., Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv. <https://arxiv.org/abs/1409.1556>

- [18] Sultana F., Sufian A., & Dutta P. (2018). *Advancements in Image Classification using Convolutional Neural Network*. 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). <https://doi.org/10.1109/icrcicn.2018.8718718>
- [19] TensorFlow Lite guide. (2021). TensorFlow. <https://www.tensorflow.org/lite/guide>
- [20] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data, 5(1). <https://doi.org/10.1038/sdata.2018.161>
- [21] Yu, C., Yang, S., Kim, W., Jung, J., Chung, K.-Y., Lee, S. W., & Oh, B. (2018). Acral melanoma detection using a convolutional neural network for dermoscopy images. *PLOS ONE*, 13(3). <https://doi.org/10.1371/journal.pone.0193321>
- [22] Zoph, B., Vasudevan, V., Shlens, J., and Le, Q.V. (2018). Learning Transferable Architectures for Scalable Image Recognition. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8697–8710. <https://doi.org/10.1109/CVPR.2018.00907>

T. V. Demianets, D. V. Fedasyuk

Lviv Polytechnic National University, Lviv, Ukraine

APPLICATION OF CONVOLUTIONAL NEURAL NETWORK FOR DETECTION OF MELANOMA USING SKIN LESION IMAGE ON MOBILE DEVICE

A melanoma is the deadliest skin cancer, so early diagnosis can provide a positive prognosis for treatment. Modern methods for early detecting melanoma on the image of the tumor are considered, and their advantages and disadvantages are analyzed. The article demonstrates a prototype of a mobile application for the detection of melanoma on the image of a mole based on a convolutional neural network, which is developed for the Android operating system. The mobile application contains melanoma detection functions, history of the previous examinations and a gallery with images of the previous examinations grouped by the location of the lesion. The HAM10000-based training dataset has been supplemented with the images of melanoma from the archive of The International Skin Imaging Collaboration to eliminate class imbalances and improve network accuracy. The search for existing neural networks that provide high accuracy was conducted, and VGG16, MobileNet, and NASNetMobile neural networks have been selected for research. Transfer learning and fine-tuning has been applied to the given neural networks to adapt the networks for the task of skin lesion classification. It is established that the use of these techniques allows to obtain high accuracy of the neural network for this task. The process of converting a convolutional neural network to an optimized Flatbuffer format using TensorFlow Lite for placement and use on a mobile device is described. The performance characteristics of the selected neural networks on the mobile device are evaluated according to the classification time on the CPU and GPU and the amount of memory occupied by the file of a single network is compared. The neural network file size was compared before and after conversion. It has been shown that the use of the TensorFlow Lite converter significantly reduces the file size of the neural network without affecting its accuracy by using an optimized format. The results of the study indicate a high speed of application and compactness of networks on the device, and the use of graphical acceleration can significantly decrease the image classification time of the tumor. According to the analyzed parameters, NASNetMobile was selected as the optimal neural network to be used in the mobile application of melanoma detection.

Keywords: Android application; tumor images; artificial intelligence; machine learning; transfer learning; fine tuning.

Інформація про авторів:

Дем'янець Тарас Валентинович, магістр, кафедра програмного забезпечення. Email: tarasdemyanets1803@gmail.com

Федасюк Дмитро Васильович, д-р техн. наук, професор, завідувач кафедри програмного забезпечення.

Email: dmytro.v.fedasyuk@lpnu.ua; <https://orcid.org/0000-0003-3552-7454>; ResearcherID: P-2810-2016

Цитування за ДСТУ: Дем'янець Т. В., Федасюк Д. В. Застосування згорткової нейронної мережі для виявлення меланоми за зображенням новоутворення на мобільному пристрої. Український журнал інформаційних технологій. 2021, т. 3, № 1. С. 08–14.

Citation APA: Demianets, T. V., & Fedasyuk, D. V. (2021). Application of convolutional neural network for detection of melanoma using skin lesion image on mobile device. *Ukrainian Journal of Information Technology*, 3(1), 08–14.

<https://doi.org/10.23939/ujit2021.03.008>