

УДК 621.382

ДЕНІСОВ Р. В., ПОПОВИЧ П. В.

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського" Київ, Україна

ОСОБЛИВОСТІ ПОПЕРЕДНЬОЇ ОБРОБКИ ТА ГРУПУВАННЯ ТРЕНУВАЛЬНИХ ДАНИХ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ПІДВИЩЕННЯ ТОЧНОСТІ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ОСНОВІ MOBILENETV2

Мета. Дослідження можливостей підвищення точності та варіативності розпізнавання різних груп об'єктів зі схожою надлишковою інформацією нейронною мережею, після попередньої обробки та групування тренувальних зображень для подальшого використання на мікроконтролерах. Перевірено вплив видалення зайвої інформації у тренувальних даних на практичні значення точності розпізнавання різних категорій об'єктів на основі архітектури MobileNet V2.

Методика. Створення різних груп тренувальних зображень з використанням програмної платформи Edge Impulse. Обробка зображень методом видалення зайвих об'єктів у графічному редакторі Adobe Photoshop. Перевірка точності розпізнавання необробленої, обробленої та змішаної груп тренувальних зображень.

Результати. На базі програмної платформи Edge Impulse створено декілька груп тренувальних зображень. У частині зображень за допомогою графічного редактора Adobe Photoshop видалено усі зайві елементи. Встановлено, що при наявності схожої зайвої інформації на тренувальних даних, у різних категоріях об'єктів, у результаті розпізнавання виникають нерозпізнані та помилково розпізнані тестові зображення. Експериментально отримано, що метод видалення зайвої інформації з тренувальних зображень дає більш чіткий розподіл ознак, а поєднання необроблених та оброблених тренувальних даних дають середній приріст у точності розпізнавання більше ніж 10% для кожної категорії, а також суттєве зменшення нерозпізнаних та невірно розпізнаних тестових зображень, при тій самій кількості тренувальних даних.

Наукова новизна. Отримано комбінований метод обробки та групування тренувальних даних, що підвищує точність розпізнавання об'єктів без збільшення кількості тренувальних даних. Досліджено вплив схожої надлишкової інформації у різних категоріях об'єктів на точність розпізнавання.

Практична значимість. Отримані результати дозволяють здійснити підвищення точності розпізнавання різних груп об'єктів зі схожою надлишковою інформацією однією нейронною мережею без підвищення кількості тренувальних зображень.

Ключові слова: попередня обробка зображень; мікроконтролери; розпізнавання зображень; згорткові нейронні мережі; Edge Impulse; MobileNetV2.

Вступ. В наші дні розробка та впровадження нейронних мереж стали актуальними завданнями у багатьох сферах життя. Однак точність розпізнавання об'єктів за допомогою згорткових нейронних мереж залишається однією з ключових проблем, особливо коли мова йде про роботу з обмеженими ресурсами, такими як мікроконтролери. Особливо сильно точність розпізнавання втрачається у нейронних мережах, де різні категорії мають схожу надлишкову інформацію на тренувальних зображеннях, наприклад, небо. Існують різні методики підвищення точності розпізнавання, такі як розширення даних (Data Augmentation), балансування класів (Class Balancing), видалення аномалій (Outlier Removal), а також метод підвищення кількості даних. Кожна методика має різні значення результативності та вплив на обсяг необхідної пам'яті. Так було досліджено вплив поступового розширення кількості тренувальних даних для відслідковування приросту точності [1, 2]. Отримано, що зі збільшенням тренувальних даних є великий приріст у точності. При цьому, існує граничне значення для кожної мережі окремо, коли кількість тренувальних даних перестає давати

приріст точності розпізнавання. Встановлено, що навіть при зменшені набору тренувальних даних, можна отримати вірогідність розпізнавання конкретної моделі 90%. У роботі з дослідження впливу розміру набору даних на ефективність класифікації було встановлено, що загальна продуктивність класифікаторів більше залежить від того, наскільки набір тренувальних даних відповідає кінцевій меті розпізнавання, а не від кількості тренувальних зображень [3].

Постановка завдання. Метою даної роботи є аналіз впливу попередньої обробки тренувальних зображень методом видалення зайвої інформації у Adobe Photoshop та змішування оброблених та необроблених тренувальних даних на точність розпізнавання різних груп об'єктів (зі схожою надлишковою інформацією) однією нейронною мережею при невеликій кількості тренувальних даних.

Для здійснення аналізу необхідно:

- На базі програмної платформи Edge Impulse створити нейронну мережу, що буде розпізнавати такі категорії як "windows", "doors", "trees", "traffic lights", "crosswalk" та оцінити точність розпізнавання обраних категорій зображення на основі архітектури MobileNet V2.
- Виконати обробку частини тренувальних зображень методом видалення зайвих об'єктів. Змішати оброблені зображення з необробленими зображеннями та проаналізувати приріст точності розпізнавання.

Результати попередніх досліджень. Однією з головних тенденцій розвитку архітектур згорткових нейронних мереж є те, що мережі стають дедалі глибшими, а точність розпізнавання вищою. Однак, збільшення глибини нейронної мережі вимагає більших обчислювальних ресурсів для тренування та подальшого розгортання мережі, а також великої кількості якісних тренувальних даних. Враховуючи попит використання нейронних мереж на приладах з обмеженими ресурсами, такими як мікроконтролер, виникає потреба у зменшенні розміру мережі та набору даних.

Попередня обробка тренувальних даних – один з основних методів підвищення точності нейронних мереж для пристроїв з обмеженими ресурсами. У роботі з дослідження впливу попередньої обробки даних на класифікацію пухлин головного мозку за допомогою capsulnet було встановлено, що при використанні необроблених даних точність класифікації становила 87% [4]. Після виконання попередньої обробки даних, а саме повороту та масштабування зображень, точність класифікації пухлин становила 92,6%.

У роботі з порівняння методів доповнення та методу попередньої обробки тренувальних даних для глибокого навчання та хемометричної класифікації інфрачервоних спектрів було встановлено, що поєднання методів попередньої обробки та методу доповнення можуть покращити точність розпізнавання на 10% при невеликому наборі даних, що складається з 80 зображень [5].

Дослідження доводять, що поєднання різних методів попередньої обробки даних та розширення бази даних дає більший приріст точності розпізнавання у порівнянні з використанням кожного методу окремо [6].

Надлишкові дані і неточності в тренувальних зображеннях можуть призвести як до підвищення точності розпізнавання, так і до зниження [7]. Некоректні тренувальні дані можуть суттєво знижувати точність розпізнавання об'єктів. У дослідженні впливу балансу білого на точність розпізнавання об'єктів встановлено, що при використанні попередньо нормалізованих за балансом білого тренувальних зображень підвищуються точність створення класифікаторів, і як наслідок, підвищується точність розпізнавання об'єктів [8].

Враховуючи, що використання нейронних мереж набуває з кожним роком все більшого поширення, компанією Google розроблено клас високоефективних нейронних мереж MobileNets [9]. Засновані на спрощеній архітектурі мережі MobileNets використовують

згортку за глибиною, а також гіперпараметри для стиснення глибокої нейронної мережі до меншого розміру, що при цьому дозволяє отримувати досить високі результати розпізнавання [10, 11].

Таким чином, поєднання різних методів попередньої обробки тренувальних даних разом з методом розширення даних дає можливість отримати максимальний приріст точності розпізнавання навіть у випадку невеликого набору тренувальних даних. А використання архітектури MobileNetV2 дає можливість розгортати отримані нейронні мережі на пристроях з обмеженими ресурсами, такими як мікроконтролер.

Далі наведено опис, умови проведення експерименту та власне його результати.

Основні методи попередньої обробки тренувальних даних. Розширення даних (Data Augmentation). Метод розширення даних, такий як обертання, обрізка, зміна масштабу, додавання шуму та інші операції, допомагає моделям стати більш стійкими до різних варіацій вхідних даних. Наприклад, обртаючи та віддзеркалюючи зображення, можна покращити роботу моделі для об'єктів з різних ракурсів [12].

Проте, як вказано у роботі з дослідження надлишковості вибірки тренувальних даних, враховуючи наявні дані та методи, надлишкова вибірка в її поточних формах і методологіях є ненадійною для навчання на незбалансованих класах даних і її слід уникати в реальних програмах [13].

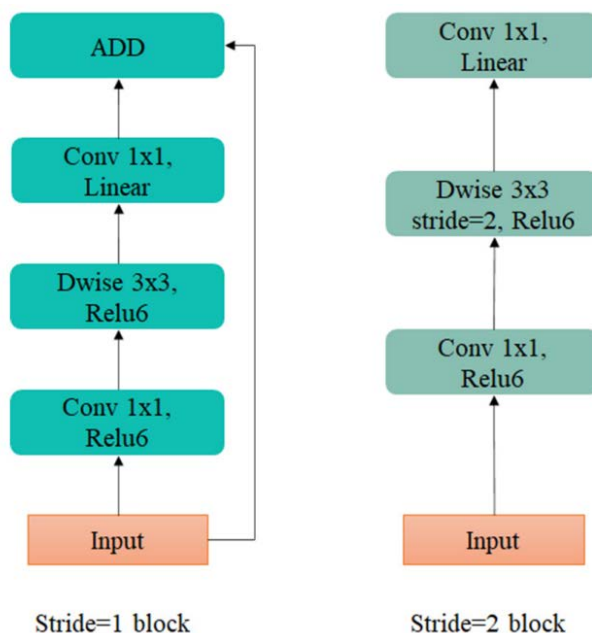
Видалення аномалій (Outlier Removal). Видалення аномальних або невідповідних даних з тренувального набору може покращити стабільність та точність моделі. Наявність невідповідних елементів у даних може значно знизити точність моделі або, ще гірше, призвести до упередженості моделі, що призведе до неточної класифікації. Таким чином, виявлення зайвих елементів і їх усунення має вирішальне значення для створення якісних навчальних наборів даних [14].

Отже, комбінуючи різні методи попередньої обробки тренувальних даних, можна отримати суттєвий приріст точності розпізнавання об'єктів.

Особливості архітектури MobileNet V2. Модель MobileNetV2 базується на попередній версії MobileNetV1, головна відмінність MobileNetV2 полягає у додаванні нового шару – інвертованого залишкового блоку із лінійними вузькими місцями (bottleneck layers), які зменшують розмір вхідних даних. Згортки, що розділяються за глибиною, які використовують в MobileNetV1, також застосовують в MobileNetV2. У поєднанні з поглибленою згорткою та поточною згорткою, загальну кількість параметрів і обчислювальну вартість моделі можна зменшити приблизно до 18 у порівнянні зі звичайною згорткою. Модель спочатку розширює низьковимірне стиснене представлення вхідних даних до високорозмірного та фільтрує його за допомогою згортки за глибиною. Згодом об'єкти проєктують назад до низьковимірною представлення з лінійною згорткою. Така структура в моделі може зберігати інформацію, вирішувати проблему негнучкої кількості фільтрів у MobileNetV1 і одночасно підтримувати легкість блоку [15]. На рис. 1 показано згортковий блок MobileNetV2.

До основних переваг MobileNetV2 можна віднести високу точність розпізнавання під час стиснення мережі для застосування на пристроях з обмеженими ресурсами, такими як мікроконтролери, а також можливість реалізації за допомогою стандартних операцій у сучасних фреймворках.

Експериментальна оцінка точності розпізнавання попередньо необроблених та змішаних тренувальних даних. Для оцінки впливу попередньої обробки тренувальних даних методом видалення зайвих об'єктів у Adobe Photoshop на точність розпізнавання об'єктів проведено ряд експериментів на основі програмної платформи Edge Impulse [16].



Джерело: [15].

Рис. 1. Згортковий блок MobileNetV2

Edge Impulse – це інструмент, призначений для інженерів та дослідників, які працюють з нейронними мережами для розпізнавання різних видів даних, включаючи зображення, звукові сигнали, дані датчиків та інші електричні сигнали. Він спрямований на створення моделей для використання у вбудованих системах та спеціалізованих пристроях. Edge Impulse використовує фреймворк Keras, який є високорівневим інтерфейсом програмування (API) для платформи машинного навчання TensorFlow, розробленою компанією Google. Програмна бібліотека з відкритим кодом та набором функцій для машинного навчання на основі нейронних мереж глибокого навчання TensorFlow містить бібліотеку моделей MobileNet. Edge Impulse надає можливість вибрати одну з декількох наявних моделей MobileNet з подальшим налаштуванням гіперпараметрів, таких як розмір зображення та коефіцієнт стиснення мережі. Після тренування мережі є можливість вибрати необхідну платформу розробки і завантажити отриману мережу у пам'ять мікроконтролера. Edge Impulse також надає можливість отримувати тестові чи тренувальні дані за допомогою сторонніх пристроїв, наприклад, телефону чи мікроконтролера з камерою, підключених до проекту.

Adobe Photoshop – це потужний графічний редактор, який використовують для обробки та редагування зображень і графіки. Це програмне забезпечення має широкий спектр функцій та можливостей, серед найважливіших можна виділити такі:

1. Редагування та обробка зображень: Adobe Photoshop надає інструменти для виконання різних операцій із зображеннями, таких як обрізання, корекція кольорів, ретушування, видалення об'єктів, зміна розміру і форми, застосування фільтрів тощо.

2. Можливість роботи з шарами та редагувати їх окремо, що дозволяє створювати складні композиції, а також має різноманітні інструменти для роботи з текстом.

3. Photoshop підтримує роботу з RAW-файлами, що дозволяє обробляти зображення в їх природному форматі з більшою глибиною кольору і контролем над експозицією.

Для проведення експериментів створено 5 груп тренувальних даних із назвами "windows", "door", "trees", "traffic lights", "crosswalk" для 5 категорій об'єктів відповідно. Набір категорій обрано таким чином, щоб у майбутньому отриману нейронну мережу можна було

застосовувати для орієнтації у просторі людей з частковою або повною втратою зору. Кожна група містить по 30 тренувальних необроблених зображень.

Після створення груп обрано колірний простір RGB, сформовано характерні ознаки для кожної категорії та виконано навчання нейронної мережі.

Для проведення експериментів обрано модель нейронної мережі MobileNetV2 з розміром вхідного зображення 96*96 пікселів і коефіцієнтом стиснення мережі 0,35. Дана модель має високі показники точності розпізнавання, а коефіцієнт стиснення 0,35 забезпечує оптимальні параметри розміру мережі для подальшого застосування на мікроконтролерах.

Результатом розпізнавання є імовірність визначення приналежності об'єкта на зображенні до груп, які було використано для тренування моделі, де сумарна імовірність приналежності до обох груп дорівнює 1 і розподіляється між ними пропорційно до результатів розпізнавання.

Метою першого експерименту є перевірка точності розпізнавання кожної категорії об'єктів на 30 тестових зображеннях. Половина тестових зображень було взято з відкритих ресурсів, 15 тестових зображень було отримано за допомогою фотографування реальних об'єктів за допомогою смартфона.

У другому експерименті досліджено вплив попередньої обробки тренувальних даних, а саме видалення сторонніх об'єктів у редакторі Adobe Photoshop, на точність розпізнавання об'єктів. Після видалення у половині тренувальних зображень зайвих елементів було змішано один до одного необроблені і оброблені зображення, виконано перенавчання нейронної мережі і перевірено точність розпізнавання об'єктів на тих самих тестових зображеннях.

Аналіз результатів експерименту. В даному розділі наведено результати проведених експериментів. Точність розпізнавання визначалася на основі імовірності приналежності зображення групам "windows", "doors", "trees", "traffic lights", "crosswalk". Тобто, якщо імовірність того, що зображення належить до однієї з категорій перевищувала 0,5, то результат вважався позитивним. Середнє значення ймовірності обчислювалось на основі 30 тестувань для кожної категорії зображень.

Після створення груп тренувальних зображень на базі програмної платформи Edge Impulse отримано розподіл ознак для кожної категорії об'єктів.

Ознаки зображення в нейронних мережах – це абстрактні представлення ключових характеристик або властивостей зображення, які використовуються для його аналізу та розпізнавання. Ознаки виникають під час проходження зображення через різні шари нейронної мережі і є числовими або векторними значеннями, які відображають важливі деталі та структуру зображення. Ознаки зображення можуть містити такі характеристики: контур та форма, кольори і текстури, локальні особливості об'єктів, глибини і рельєфи а також абстрактні ознаки.

Графічне відображення отриманих ознак – це візуалізація виділення ознак з кожного зображення кожної групи тренувальних даних. Різні групи об'єктів зображень мають свої характерні ознаки, що надалі будуть використані для навчання мережі. Якщо після генерації ознак візуальний розподіл ознак різних груп не має чіткого поділу, це означає, що у тренувальних даних різних категорій зображень є велика кількість схожої інформації.

Проаналізувавши графічне відображення отриманого розподілу, можна зробити висновок, що надлишкова інформація на тренувальних зображеннях, яка має схожий характер, наприклад, небо, може створювати помилки у розрізненні ознак категорій. Розподіл ознак зображено на рис. 2.

Після навчання нейронної мережі перевірено точність розпізнавання об'єктів на 30 тестових зображеннях для кожної категорії. Середнє значення точності розпізнавання, кількість нерозпізнаних зображень у кожній категорії ($N_{\text{нерозпізнане}}$) та кількість помилково розпізнаних

зображень ($N_{\text{помилково розпізнане}}$) зазначено у таблиці 1. Зображення вважається нерозпізнаним, якщо не отримало приналежність до однієї з категорій більше ніж у 50%. Помилково розпізнаним зображенням вважається зображення, що отримало приналежність до невідповідної йому категорії більше ніж у 50%.

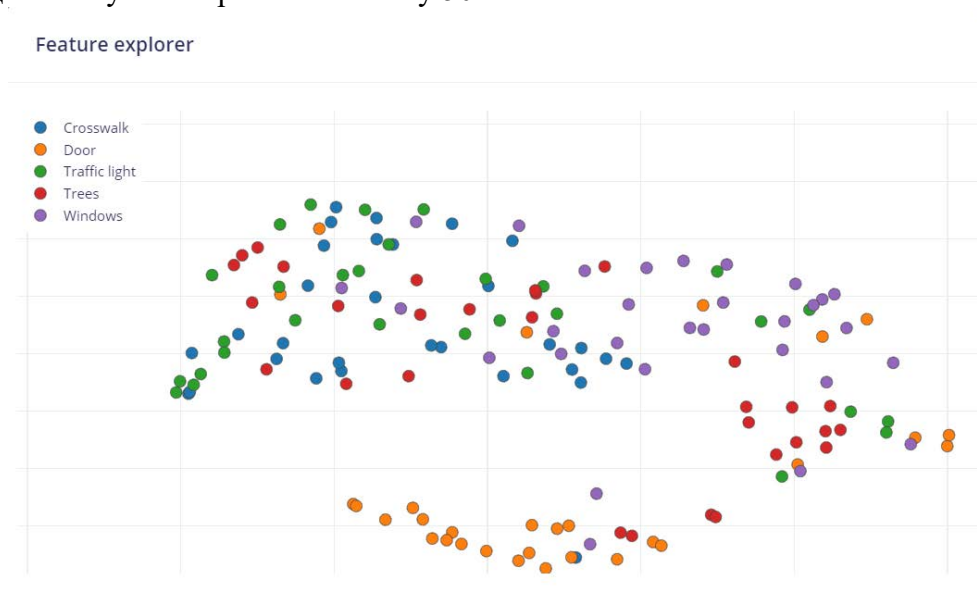


Рис. 2. Графічне відображення розподілу ознак необроблених тренувальних даних

Категорія "trees" має найбільшу кількість нерозпізнаних та помилково розпізнаних тестових зображень через те, що дерева можуть бути присутні на тренувальних зображеннях різних категорій, наприклад, відбиватися або просвічуватися через вікна, знаходитися поруч із пішохідним переходом. У такому випадку результат розпізнавання розподіляється між категоріями, що мають схожі ознаки. На рис. 3 зображено приклад розподілу вірогідності нерозпізнаного тестового зображення дерева.

CROSSWALK	DOORS	TRAFFIC_LIGHT	TREES	WINDOWS
0.04	0.01	0.20	0.30	0.45

Рис. 3. Нерозпізнане зображення дерева

З отриманих результатів можна зробити висновок, що зображення, які мають схожу сторонню інформацію у різних категоріях об'єктів, можуть бути нерозпізнані або можуть отримати невірну класифікацію.

Таблиця 1

Результати точності розпізнавання мережі з необробленими тренувальними даними

Назва групи об'єктів	Точність	$N_{\text{нерозпізнане}}$	$N_{\text{помилково розпізнане}}$
Trees	0,57	5	4
Windows	0,63	4	3
Doors	0,68	2	3
Traffic lights	0,73	2	2
Crosswalk	0,74	3	2

Для перевірки впливу попередньої обробки тестових даних у програмі Adobe Photoshop оброблено по 15 тренувальних зображень для кожної категорії методом видалення зайвих об'єктів з зображення. Приклад обробки тренувальних даних зображено на рис.4.

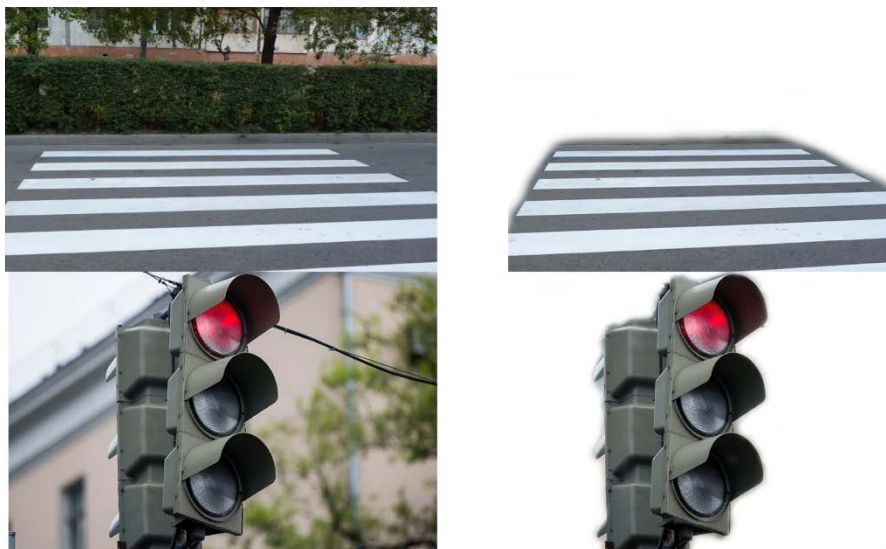


Рис. 4. Приклад необроблених і оброблених тренувальних даних

Після обробки половини тренувальних даних у кожній категорії виконано нову генерацію ознак з відредагованих тренувальних даних (рис. 5). З отриманого розподілу, можна зробити висновок, що видалення зайвих елементів з даних дало більш рівномірний розподіл ознак у випадку використання лише половини від стартових тренувальних зображень. Після чого було виконано тренування нейронної мережі та перевірено точність розпізнавання на тих самих тестових зображеннях. Результати визначення точності розпізнавання після обробки тренувальних даних наведено у таблиці 2.

Таблиця 2

Результати точності розпізнавання мережі з обробленими тренувальними даними

Назва групи об'єктів	Точність	N _{нерозпізнане}	N _{помилково розпізнане}
Trees	0,61	2	2
Windows	0,66	1	2
Doors	0,72	2	1
Traffic lights	0,79	1	1
Crosswalk	0,82	1	1

З отриманих даних можна побачити, що видалення зайвих елементів призводить до більш чіткого формування ознак, і як наслідок, до меншої кількості помилкових і нерозпізнаних зображень. Також отримано приріст середньої точності розпізнавання категорій об'єктів та зменшення нерозпізнаних і помилково розпізнаних тестових зображень.

Враховуючи те, що у реальному житті під час використання нейронних мереж для розпізнавання різних категорій об'єктів зайва інформація буде постійно знаходитись біля об'єктів розпізнавання, для фінальної перевірки точності розпізнавання було змішано у пропорції один до одного необроблені і оброблені тренувальні данні. Кінцевий розподіл ознак зображено на рис. 6.

Feature explorer

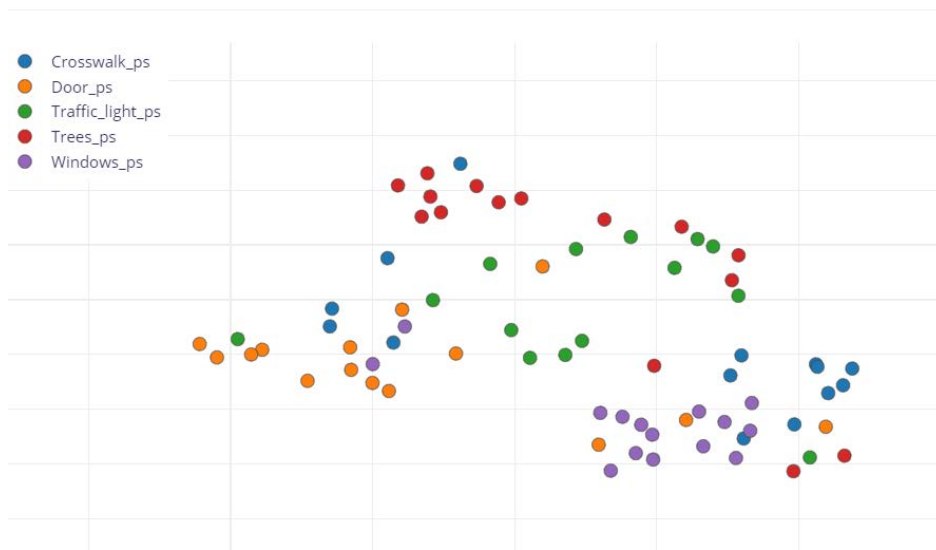


Рис. 5. Графічне відображення розподілу ознак оброблених тренувальних даних

Feature explorer

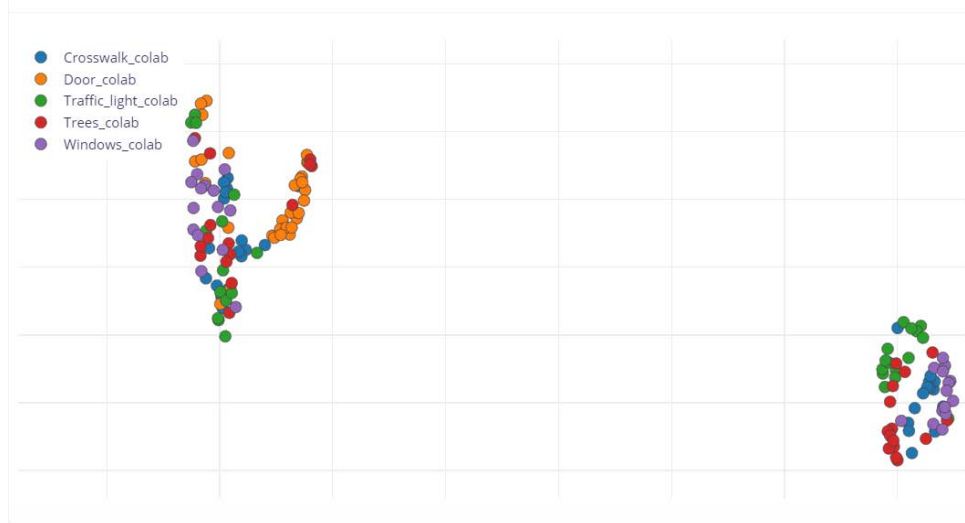


Рис. 6. Графічне відображення розподілу ознак оброблених та необроблених тренувальних даних

Проаналізувавши отриманий розподіл ознак, можна зробити такі висновки: для кожної категорії об'єктів сформовано різні ознаки на базі оброблених і необроблених тренувальних даних, таким чином розширено спектр можливих варіацій вхідних даних для розпізнавання. Отриманий розподіл більш рівномірний, ніж два попередніх розподіли, що в свою чергу означає потенційне підвищення точності розпізнавання і зменшення кількості помилкових та нерозпізнаних зображень.

Також можна зробити висновки щодо важливості вірного групування тренувальних даних, якщо необхідно розпізнавати різні категорії об'єктів зі схожою надлишковою інформацією однією нейронною мережею.

Після генерації ознак, на основі змішаних тренувальних даних виконано навчання нейронної мережі та перевірку на тих самих тестових зображеннях. Результати розпізнавання відображено у таблиці 3.

Таблиця 3

Результати точності розпізнавання мережі зі змішаними тренувальними даними.

Назва групи об'єктів	Точність	N _{нерозпізнане}	N _{помилково розпізнане}
Trees	0,73	1	1
Windows	0,77	1	0
Doors	0,80	0	1
Traffic lights	0,85	0	0
Crosswalk	0,87	0	0

Використання змішаного методу тренування дає приріст точності, зменшує кількість нерозпізнаних та помилково розпізнаних об'єктів. Слід зауважити, що у категоріях "traffic lights" та "crosswalk" усі тестові зображення було розпізнано вірно, а у категоріях "windows", "doors", "trees" ще залишилися нерозпізнані або помилково розпізнанні тестові зображення. Це може свідчити про те, що для розпізнавання схожих категорій за візуальними ознаками необхідно збільшувати кількість тренувальних даних у порівнянні з категоріями, де відмінності у об'єктах більш очевидні на рівні класифікації ознак.

Отже, метод попередньої обробки тренувальних даних у поєднанні з необробленими тренувальними даними дає приріст точності розпізнавання різних категорій об'єктів, зменшує кількість помилкових та нерозпізнаних зображень. Середній приріст точності для кожної окремої категорії становить більше 10%, що є досить суттєвим, враховуючи те, що кількість тренувальних даних лишилася незмінною. Такий підхід до обробки тренувальних даних може використовуватися для тренування нейронних мереж, які можуть розгортатися на пристроях з обмеженими ресурсами, для об'єктів з обмеженою кількістю тренувальних даних, а також за необхідності навчати нейронну мережу розпізнавати об'єкти зі схожою надлишковою інформацією.

Висновки. Проведено експериментальне дослідження впливу попередньої обробки тренувальних даних методом видалення зайвих об'єктів, а саме "windows", "doors", "trees", "traffic lights", "crosswalk" у графічному редакторі Adobe Photoshop на точність розпізнавання різних категорій об'єктів нейронною мережею на базі архітектури MobileNetV2 з розміром зображення 96*96 пікселів та коефіцієнтом стиснення мережі 0.35. Встановлено, що у випадку наявності схожої надлишкової інформації на тренувальних даних у різних категоріях об'єктів розподіл характерних ознак для різних категорій може бути не чіткий, і як наслідок, виникають нерозпізнані та невірно розпізнанні тестові зображення.

Також експериментально отримано, що метод видалення зайвої інформації з тренувальних зображень дає більш чітке виділення ознак, а поєднання необроблених та оброблених тренувальних даних дає середній приріст точності розпізнавання більше ніж 10% для кожної категорії, а також суттєве зменшення нерозпізнаних та невірно розпізнаних тестових зображень.

Також встановлено, що метод видалення зайвої інформації буде максимально ефективним у мережах, де різні категорії мають схожу колірну інформацію, але суттєво відрізняються за своєю формою. При цьому кількість тренувальних зображень не була збільшена, що дозволяє використання такого методу для обробки зображень, у різних ситуаціях, особливо коли мережа буде використовуватися у приладах з обмеженими ресурсами, такими як мікроконтролери.

References

1. Thian, Y. L., Ng, D. W., Hallinan, J. T. P. D. et al. (2022). Effect of Training Data Volume on Performance of Convolutional Neural Network Pneumothorax Classifiers. *J Digit Imaging*, 35, 881–892. <https://doi.org/10.1007/s10278-022-00594-y>.
2. Khan, I. U., Raiaan, M. A. K., Fatema, K., Azam, S., Rashid, R. U., Mukta, S. H., Jonkman, M., De Boer, F. (2023). A Computer-Aided Diagnostic System to Identify Diabetic Retinopathy, Utilizing a Modified Compact Convolutional Transformer and Low-Resolution Images to Reduce Computation Time. *Biomedicines*, 11, 1566. <https://doi.org/10.3390/biomedicines11061566>.
3. Kiran Maharana, Surajit Mondal, Bhushankumar Nemade. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, Vol. 3, Iss. 1, P. 91–99. <https://doi.org/10.1016/j.gltp.2022.04.020>.
4. Vimal Kurup, R., Sowmya, V., Soman, K. P. (2020). Effect of Data Pre-processing on Brain Tumor Classification Using Capsulenet. In: *ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Management (ICICCT 2019)*. Gunjan, V., Garcia Diaz, V., Cardona, M., Solanki, V., Sunitha, K. (eds). Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_13.
5. Blazhko, U., Shapaval, V., Kovalev, V., Kohler, A. (2021). Comparison of augmentation and pre-processing for deep learning and chemometric classification of infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, Vol. 215, 104367. <https://doi.org/10.1016/j.chemolab.2021.104367>.
6. Obaid, H. S., Dheyab, S. A., Sabry, S. S. (2019). The Impact of Data Pre-Processing Techniques and Dimensionality Reduction on the Accuracy of Machine Learning. *2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)*, Jaipur, India, 2019, pp. 279–283, DOI:10.1109/IEMECONX.2019.8877011.
7. Mohammed, R., Rawashdeh, J., Abdullah, M. (2020). Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. *2020 11th International Conference on Information and Communication Systems (ICICS)*, Irbid, Jordan, 2020, pp. 243–248, DOI:10.1109/ICICS49469.2020.239556.

Література

1. Thian Y. L., Ng D. W., Hallinan J. T. P. D. et al. Effect of Training Data Volume on Performance of Convolutional Neural Network Pneumothorax Classifiers. *J Digit Imaging*. 2022. No. 35. P. 881–892. <https://doi.org/10.1007/s10278-022-00594-y>.
2. Khan I. U., Raiaan M. A. K., Fatema K., Azam S., Rashid R. U., Mukta S. H., Jonkman M., De Boer F. A Computer-Aided Diagnostic System to Identify Diabetic Retinopathy, Utilizing a Modified Compact Convolutional Transformer and Low-Resolution Images to Reduce Computation Time. *Biomedicines*. 2023. No. 11. Art. 1566. <https://doi.org/10.3390/biomedicines11061566>.
3. Kiran Maharana, Surajit Mondal, Bhushankumar Nemade. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*. 2022. Vol. 3, Iss. 1. P. 91–99. <https://doi.org/10.1016/j.gltp.2022.04.020>.
4. Vimal Kurup R., Sowmya V., Soman K. P. Effect of Data Pre-processing on Brain Tumor Classification Using Capsulenet. In: *ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Management (ICICCT 2019)*. Gunjan V., Garcia Diaz V., Cardona M., Solanki V., Sunitha K. (eds). Springer, Singapore, 2020. https://doi.org/10.1007/978-981-13-8461-5_13.
5. Blazhko U., Shapaval V., Kovalev V., Kohler A. Comparison of augmentation and pre-processing for deep learning and chemometric classification of infrared spectra. *Chemometrics and Intelligent Laboratory Systems*. 2021. Vol. 215. Art. 104367. <https://doi.org/10.1016/j.chemolab.2021.104367>.
6. Obaid H. S., Dheyab S. A., Sabry S. S. The Impact of Data Pre-Processing Techniques and Dimensionality Reduction on the Accuracy of Machine Learning. *2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)*. Jaipur, India, 2019. P. 279–283. DOI:10.1109/IEMECONX.2019.8877011.
7. Mohammed R., Rawashdeh J., Abdullah M. Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. *2020 11th International Conference on Information and Communication Systems (ICICS)*. Irbid, Jordan, 2020.

8. Afifi, M., Brown, M. (2019). What Else Can Fool Deep Learning? Addressing Color Constancy Errors on Deep Neural Network Performance. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South), pp. 243–252, DOI:10.1109/ICCV.2019.00033.
9. Howard, G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T. et al. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv*, 17 Apr 2017, P. 1–9, DOI: 10.48550/arXiv.1704.04861.
10. Mao, W.-L., Chen, S.-H., Huang, Y.-T., Yang, Y.-T., Chou, P.-H. (2023). Indoor Scene Recognition Using ARM-based MobileNets Architectures. *2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, Paris, France, pp. 225–230, DOI: 10.1109/ICUFN57995.2023.10199386.
11. Ma H., Ye Y., Dong, J., Bo, Y. (2022). An Intelligent Garbage Classification System Using a Lightweight Network MobileNetV2. *2022 7th International Conference on Signal and Image Processing (ICSIP)*, Suzhou, China, pp. 531–535, DOI:10.1109/ICSIP55141.2022.9886985.
12. Cap, Q. H., Uga, H., Kagiwada, S., Iyatomi, H. (2022). LeafGAN: An Effective Data Augmentation Method for Practical Plant Disease Diagnosis. *IEEE Transactions on Automation Science and Engineering*, Vol. 19, No. 2, P. 1258–1267, DOI:10.1109/TASE.2020.3041499.
13. Tarawneh, A. S., Hassanat, A. B., Altarawneh, G. A., Almuhaimeed, A. (2022). Stop Oversampling for Class Imbalance Learning: A Review. *IEEE Access*, Vol. 10, P. 47643–47660, DOI: 10.1109/ACCESS.2022.3169512.
14. Perez, H., Tah, J. H. M. (2020). Improving the Accuracy of Convolutional Neural Networks by Identifying and Removing Outlier Images in Datasets Using t-SNE. *Mathematics*, 8, 662. <https://doi.org/10.3390/math8050662>.
15. Dong, K., Zhou, C., Ruan, Y., Li, Y. (2020). MobileNetV2 Model for Image Classification. *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, Guangzhou, China, 2020, pp. 476–480, DOI:10.1109/ITCA52113.2020.00106.
2020. P. 243–248. DOI:10.1109/ICICS49469.2020.239556.
8. Afifi M., Brown M. What Else Can Fool Deep Learning? Addressing Color Constancy Errors on Deep Neural Network Performance. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South), 2019. P. 243–252. DOI:10.1109/ICCV.2019.00033.
9. Howard G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv*. 17 Apr 2017. P. 1–9. DOI:10.48550/arXiv.1704.04861.
10. Mao W.-L., Chen S.-H., Huang Y.-T., Yang Y.-T., Chou P.-H. Indoor Scene Recognition Using ARM-based MobileNets Architectures. *2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*. Paris, France, 2023. P. 225–230. DOI:10.1109/ICUFN57995.2023.10199386.
11. Ma H., Ye Y., Dong J., Bo Y. An Intelligent Garbage Classification System Using a Lightweight Network MobileNetV2. *2022 7th International Conference on Signal and Image Processing (ICSIP)*. Suzhou, China, 2022. P. 531–535. DOI:10.1109/ICSIP55141.2022.9886985.
12. Cap Q. H., Uga H., Kagiwada S., Iyatomi H. "LeafGAN: An Effective Data Augmentation Method for Practical Plant Disease Diagnosis. *IEEE Transactions on Automation Science and Engineering*. 2022. Vol. 19, No. 2. P. 1258–1267. DOI:10.1109/TASE.2020.3041499.
13. Tarawneh A. S., Hassanat A. B., Altarawneh G. A., Almuhaimeed A. Stop Oversampling for Class Imbalance Learning: A Review. *IEEE Access*. 2022. Vol. 10. P. 47643–47660. DOI: 10.1109/ACCESS.2022.3169512.
14. Perez H., Tah J. H. M. Improving the Accuracy of Convolutional Neural Networks by Identifying and Removing Outlier Images in Datasets Using t-SNE. *Mathematics*. 2020. 8, 662. <https://doi.org/10.3390/math8050662>.
15. Dong K., Zhou C., Ruan Y., Li Y. MobileNetV2 Model for Image Classification. *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*. Guangzhou, China, 2020. P. 476–480. DOI:10.1109/ITCA52113.2020.00106.

16. Edge impulse. [edgeimpulse.com](https://www.edgeimpulse.com/).
<https://www.edgeimpulse.com/>

URL: 16. Edge impulse. [edgeimpulse.com](https://www.edgeimpulse.com/). URL:
<https://www.edgeimpulse.com/>

DENISOV ROSTYSLAV

Post graduate student,
Department of Acoustic and
Multimedia Electronic Systems,
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
<https://orcid.org/0000-0003-1146-9114>
E-mail: rostikdenisov@gmail.com

POPOVYCH P.V.

PhD, Associate Professor,
Department of Acoustic and
Multimedia Electronic Systems,
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
<http://orcid.org/0000-0002-1572-3127>
Scopus Author ID: 55225965700
Researcher ID: J-6574-2017
E-mail: p.popovich80@gmail.com

DENISOV R. V., POPOVYCH P. V.

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

FEATURES OF NEURAL NETWORK PRE-PROCESSING AND GROUPING OF TRAINING DATA TO IMPROVE ACCURACY OF OBJECT RECOGNITION BASED ON MOBILENETV2

Purpose. Investigation of the possibilities of increasing the accuracy and variability of recognition of different groups of objects with similar redundant information by a neural network, after preprocessing and grouping of training images for further use on microcontrollers. The effect of removing redundant information in the training data on the practical values of recognition accuracy of different categories of objects based on the MobileNet V2 architecture is verified.

Methodology. Creating different groups of training images using the Edge Impulse software platform. Image processing by the method of removing unnecessary objects in the graphic editor Adobe Photoshop. Testing the recognition accuracy of raw, processed and mixed groups of training images.

Findings. Several groups of training images have been created on the basis of the Edge Impulse software platform. In part of the images, all unnecessary elements were removed using the Adobe Photoshop graphic editor. It was established that in the presence of similar redundant information on the training data, in different categories of objects, unrecognized and falsely recognized test images appear as a result of recognition. It was experimentally found that the method of removing redundant information from training images gives a clearer distribution of features, and the combination of raw and processed training data gives an average increase in recognition accuracy of more than 10% for each category, as well as a significant reduction of unrecognized and incorrectly recognized test images, while the same amount of training data.

Originality. A combined method of training data processing and grouping was obtained, which increases the accuracy of object recognition without increasing the amount of training data. The effect of similar redundant information in different categories of objects on recognition accuracy was investigated.

Practical value. The obtained results make it possible to increase the accuracy of recognition of different groups of objects with similar redundant information by one neural network without increasing the number of training images.

Keywords: pre-processing of images; microcontrollers; image recognition; convolutional neural networks; Edge Impulse; MobileNetV2.