

<https://doi.org/10.31891/2307-5732-2025-353-11>

УДК 004.93

ГАЗДЮК КАТЕРИНА

Чернівецький національний університет ім. Ю. Федьковича

<https://orcid.org/0000-0002-7568-4422>

e-mail: k.gazdiuk@chnu.edu.ua

МОВЧЕНЮК РОМАН

Чернівецький національний університет ім. Ю. Федьковича

e-mail: movcheniuk.roman@chnu.edu.ua

РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ РЕНТГЕНІВСЬКИХ ЗНІМКІВ НА ОСНОВІ DEEP LEARNING

Розвиток методів глибокого навчання відкриває нові можливості для автоматичної обробки та класифікації медичних знімків, що сприяє підвищенню ефективності діагностики та зменшенню навантаження на лікарів. У цій роботі розглядається застосування згорткових нейронних мереж (CNN) для автоматичного розпізнавання та класифікації медичних рентгенівських знімків із використанням набору даних, отриманого з відкритого ресурсу Kaggle. Запропонований підхід передбачає поетапну обробку даних, починаючи від їхньої предобробки та нормалізації, закінчуючи навчанням моделі та її оцінюванням на валідаційній вибірці.

Для побудови моделі було використано архітектуру глибокої згорткової нейронної мережі, яка включає кілька згорткових шарів із функцією активації ReLU, шари нормалізації та пулінгу для зменшення розмірності вхідних даних, а також повноз'язний шар для класифікації. Навчання моделі здійснювалося за допомогою оптимізатора Adam, який забезпечує швидку та стабільну мінімізацію функції втрат категоричної крос-ентронії. Для боротьби з перенавчанням були застосовані методи регуляризації, зокрема виключення нейронів (Dropout) і рання зупинка (Early Stopping), що дозволило підвищити узагальнюючу здатність моделі.

Оцінка ефективності моделі здійснювалася на основі таких метрик, як точність (accuracy), повнота (recall), F1-міра та функція втрат (loss) на тестовій вибірці. Експериментальні результати продемонстрували, що модель досягла високих показників класифікації: точність 92.3%, повнота 91.1%, F1-міра 91.7% та значення функції втрат 0.25. Отримані результати свідчать про ефективність запропонованого підходу та його конкурентоспроможність із сучасними методами автоматичної діагностики медичних зображень.

Запропонований метод може бути застосований у клінічній практиці для автоматизованої допомоги лікарям у діагностиці захворювань, що зменшує ймовірність суб'єктивних помилок та підвищує швидкість аналізу рентгенівських знімків. Використання хмарних технологій, таких як Google Cloud AI Platform, дозволяє розгорнути модель у вигляді веб-сервісу для віддаленого доступу та спрощує інтеграцію з медичними інформаційними системами. У перспективі планується подальше вдосконалення моделі шляхом використання трансформерних архітектур та розширення датасету для покращення генералізації результатів.

Ключові слова: глибоке навчання, розпізнавання зображень, рентгенівські знімки, штучні нейронні мережі, медична діагностика, автоматизація медичних процесів

HAZDIUK KATERYNA

MOVCHENIUK ROMAN

Yuriy Fedkovych Chernivtsi National University

DEVELOPMENT OF AN X-RAY IMAGE RECOGNITION SYSTEM BASED ON DEEP LEARNING

The advancement of deep learning methods opens up new possibilities for automated processing and classification of medical images, increasing diagnostic efficiency while reducing the workload on physicians. This study examines the application of convolutional neural networks (CNNs) for the automatic recognition and classification of medical X-ray images using a publicly available dataset from Kaggle. The proposed approach involves sequential data processing, starting with preprocessing and normalization steps and concluding with model training and evaluation on a validation set.

A deep convolutional neural network architecture was utilized to build the model, comprising multiple convolutional layers with ReLU activation, normalization layers, and pooling layers for dimensionality reduction, as well as a fully connected layer for classification. Training was carried out using the Adam optimizer, which provides rapid and stable minimization of the categorical cross-entropy loss. To address overfitting, regularization methods such as Dropout and Early Stopping were employed, thereby enhancing the model's generalization capability.

The model's effectiveness was evaluated using metrics including accuracy, recall, F1-score, and loss on the test set. Experimental results demonstrated high classification performance, with an accuracy of 92.3%, recall of 91.1%, an F1-score of 91.7%, and a loss value of 0.25. These findings highlight the effectiveness of the proposed approach and its competitiveness with state-of-the-art automated medical imaging diagnosis methods.

The proposed method can be utilized in clinical practice to assist physicians in disease diagnosis, reducing the likelihood of subjective errors and speeding up the analysis of X-ray images. Moreover, the use of cloud technologies, such as the Google Cloud AI Platform, enables model deployment as a web service for remote access and simplifies integration with medical information systems. Future research will focus on further model refinement through the use of Transformer architectures and the expansion of the dataset to improve generalizability of the results.

Keywords: deep learning, image recognition, X-ray images, artificial neural networks, medical diagnosis, medical process automation.

Стаття надійшла до редакції / Received 24.04.2025

Прийнята до друку / Accepted 16.05.2025

Постановка проблеми

На сучасному етапі розвитку медичної індустрії особлива увага приділяється розробці та впровадженню інноваційних технологічних рішень, які б могли підвищити ефективність діагностичних процесів. Одним із таких напрямків є використання штучного інтелекту для аналізу рентгенівських знімків. Традиційні методи діагностики вимагають значних трудових ресурсів та часу, а в умовах сучасних викликів, таких як пандемії, потреба у швидкій та точній діагностиці стає дедалі актуальнішою.

Система, розроблена в рамках цього дослідження, базується на принципах глибокого навчання, що дозволяє автоматизувати процес розпізнавання і класифікації рентгенівських зображень. Мета проекту полягає не лише у підвищенні точності діагнозів, але й у значному скороченні часу, необхідного для обробки знімків, що може бути критично важливим у випадку медичних надзвичайних ситуацій у віддалених регіонах.

Розробка даної системи важлива для медичної галузі, оскільки вона забезпечує допомогу медичним працівникам у процесі діагностики, зокрема в регіонах, де спостерігається нестача кваліфікованих спеціалістів. Вона також сприяє покращенню доступності та якості медичного обслуговування за допомогою впровадження передових технологій.

Для досягнення поставлених цілей використовуються сучасні алгоритми машинного навчання та глибокого навчання, засновані на моделях штучних нейронних мереж, які тренуються на великих обсягах даних. Це дозволяє системі ефективно аналізувати і розпізнавати патології на рентгенівських зображеннях.

Для досягнення оптимальних результатів передбачено використання трансферного навчання, що сприяє швидкій адаптації моделі до специфічних задач, властивих медичним зображенням. Такий підхід значно підвищує точність та надійність розробленої системи.

У сучасних медичних дослідженнях важливу роль відіграють методи автоматичного аналізу зображень, зокрема рентгенівських знімків. Використання хмарних платформ значно спрощує процес розробки та впровадження таких систем, оскільки вони надають необхідні обчислювальні ресурси та інструменти для створення, навчання й розгортання моделей машинного навчання. Серед провідних постачальників хмарних технологій варто виділити Google Cloud Platform (GCP), Amazon Web Services (AWS) та Microsoft Azure, кожен із яких пропонує широкий спектр сервісів для роботи з нейронними мережами та комп'ютерним зором.

Google Cloud Platform забезпечує гнучкі можливості для машинного навчання, зокрема за допомогою сервісу Google Cloud AI Platform [1]. Цей інструмент підтримує популярні фреймворки, такі як TensorFlow, PyTorch і Keras [2][3], що дозволяє ефективно створювати та розгортати моделі. Використання графічних (GPU) та тензорних (TPU) процесорів у межах цієї платформи дає змогу значно прискорити процес навчання нейронних мереж. Окрім цього, Google пропонує Deep Learning VMs — спеціалізовані віртуальні машини, що містять попередньо налаштоване середовище для глибокого навчання, а також AutoML Vision — сервіс, який дозволяє автоматизувати створення моделей розпізнавання зображень без необхідності глибоких знань у сфері машинного навчання.

Amazon Web Services [4] також має розвинену екосистему для машинного навчання, яка включає Amazon SageMaker — платформу для створення, навчання та розгортання моделей із підтримкою автоматизації більшості рутинних процесів. Використання підготовлених віртуальних машин Deep Learning AMIs, що містять усі необхідні бібліотеки для глибокого навчання, спрощує налаштування робочого середовища. Додатково AWS пропонує сервіс Rekognition, який дозволяє виконувати аналіз зображень та відео, розпізнавати об'єкти й автоматично виявляти аномалії, що може бути корисним у медичних застосуваннях.

Microsoft Azure, своєю чергою, надає сервіс Azure Machine Learning [5], який підтримує весь життєвий цикл роботи з моделями штучного інтелекту, включаючи їх навчання, оцінювання та розгортання. Для спрощення аналізу зображень у медичних системах Azure пропонує набір API під назвою Azure Cognitive Services, що забезпечують автоматичне розпізнавання об'єктів, аналіз емоцій та інші функції. Крім того, платформа Azure Databricks, побудована на базі Apache Spark, є ефективним рішенням для обробки великих масивів даних та застосування складних алгоритмів машинного навчання.

Вибір оптимальної хмарної платформи для розробки системи розпізнавання рентгенівських знімків залежить від низки чинників, зокрема від вимог до продуктивності, доступного бюджету та технологічної компетентності команди. Опис переваг і недоліків сервісів наведено в Таблиці 1. Використання сучасних хмарних рішень сприяє підвищенню ефективності та швидкості обробки зображень, що, у свою чергу, може покращити якість медичної діагностики.

Для оптимізації моделі класифікації рентгенівських знімків у роботі будемо використовувати Google Cloud AI Platform, оскільки вона забезпечує: високу продуктивність для глибокого навчання, а саме:

Підтримку GPU та TPU, що значно пришвидшує тренування моделей.

Наявність Deep Learning VMs із попередньо налаштованим середовищем для TensorFlow, PyTorch та Keras.

Таблиця 1

Порівняння переваг та недоліків описаних програмних засобів			
Функція/ Провайдер	Google Cloud Platform	Amazon Web Services	Microsoft Azure
Інструменти	Google Cloud AI Platform, Google Deep Learning VMs, Google AutoML Vision	Amazon SageMaker, AWS Deep Learning AMIs, AWS Rekognition	Azure Machine Learning Service, Azure Cognitive Services, Azure Databricks
Переваги	Велика інтеграція з TensorFlow та іншими фреймворками, висока продуктивність TPU, інтуїтивно зрозумілі інструменти для неспеціалістів (AutoML Vision)	Велика екосистема, включаючи спеціалізовані AMI для швидкого запуску, масштабованість і гнучкість SageMaker, висока інтегрованість з іншими сервісами AWS	Широкий спектр пре-бїлд API, хороша підтримка для гібридних розгортань (хмара та локальне використання), сильні інструменти для обробки даних (Databricks)
Недоліки	Вартість використання може бути високою для великих проєктів, складніші проєкти можуть вимагати більшої конфігурації	Вартість, складність управління середовищами, потрібно час на налаштування глибоких навчальних екосистем	Іноді менш конкурентоспроможні у візуальних та машинних навчальних можливостях порівняно з Google і AWS, залежність від платформи Microsoft
Підтримка Мов	TensorFlow, PyTorch, Scikit-learn, і інші через Deep Learning VM	TensorFlow, PyTorch, MXNet, і інші через Deep Learning AMI	TensorFlow, PyTorch, Scikit-learn через Azure ML Service
Особливості	Масштабування за допомогою клаудових TPU, спрощення тренування моделей з AutoML	Інтеграція зі всіма сервісами AWS, зручне управління моделями через SageMaker	Інтеграція з іншими продуктами Microsoft, сильні можливості обробки даних

Мета і задачі дослідження

В рамках даного дослідження було використано відкритий датасет Medical Scan Classification Dataset, розміщений на платформі Kaggle [6]. Цей набір даних містить зображення медичних сканів різних категорій, що дозволяє ефективно навчати та тестувати моделі глибокого навчання для автоматичної класифікації рентгенівських знімків.

Результати дослідження

Попередня обробка зображень включала зміну розміру, нормалізацію та збільшення (data augmentation) для покращення узагальнюючої здатності нейронної мережі. Фрагмент програмного коду попередньої обробки та формування навчальних і валідаційних вибірок наведено на рис. 1.

Для розв'язання задачі класифікації було використано модель глибокої згортової нейронної мережі (CNN), засновану на архітектурі ResNet50 [7], яка попередньо натренована на великому наборі даних ImageNet. Використання трансферного навчання дозволило значно скоротити обсяг необхідних анованих медичних зображень[8].

```

train_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=20,
    width_shift_range=0.2,
    height_shift_range=0.2,
    horizontal_flip=True,
    validation_split=0.2
)

train_generator = train_datagen.flow_from_directory(
    TRAIN_DIR,
    target_size=IMAGE_SIZE,
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='training'
)

val_generator = train_datagen.flow_from_directory(
    VAL_DIR,
    target_size=IMAGE_SIZE,
    batch_size=BATCH_SIZE,
    class_mode='categorical',
    subset='validation'
)

```

Рис. 1. Фрагмент програмного коду попередньої обробки та формування навчальних і валідаційних вибірок

Під час налаштування моделі був використаний оптимізатор Adam [9] з початковою швидкістю навчання 0,0001, який забезпечує адаптивне оновлення ваг та стабільність навчання. Функція втрат — категоріальна крос-ентропія, яка ефективна для завдань багатокласової класифікації. Метрики оцінки включали точність (accuracy) та F1-score, що дозволяє враховувати баланс між точністю та повнотою.

Архітектура мережі включала базову модель ResNet50 із замороженими вагами для збереження початкових характеристик, глобальний шар усереднення (Global Average Pooling), повнозв'язний шар із 256 нейронами та функцією активації ReLU, шар Dropout [10] з коефіцієнтом 0,5 для запобігання перенавчанню, а також фінальний шар із активацією Softmax, кількість нейронів якого відповідає кількості класів у наборі даних. Фрагмент коду створення цієї моделі наведено на рис. 2.

```
base_model = ResNet50(weights='imagenet', include_top=False, input_shape=(224, 224, 3))
base_model.trainable = False # Заморожуємо базові шари

✓ model = Sequential([
    base_model,
    GlobalAveragePooling2D(),
    Dense(256, activation='relu'),
    Dropout(0.5),
    Dense(train_generator.num_classes, activation='softmax')
])

model.compile(optimizer='adam',
              loss='categorical_crossentropy',
              metrics=['accuracy'])
```

Рис. 2. Фрагмент програмного коду створення моделі

Процес навчання проводився з розміром мініпакету (batch size) 32, що забезпечувало оптимальне використання ресурсів GPU. Тривалість навчання склала 400 епох із застосуванням ранньої зупинки для запобігання перенавчанню.

Результати тренування моделі представлені на рис. 3, який демонструє графіки точності (accuracy) та втрат (loss) під час навчання і валідації. Отримані результати свідчать про високу точність моделі — 92,3%, значення функції втрат (Validation Loss) менше 0,1, повноту (Recall) 91,1% та F1-міру 91,7%. Ці показники підтверджують ефективність запропонованого підходу.

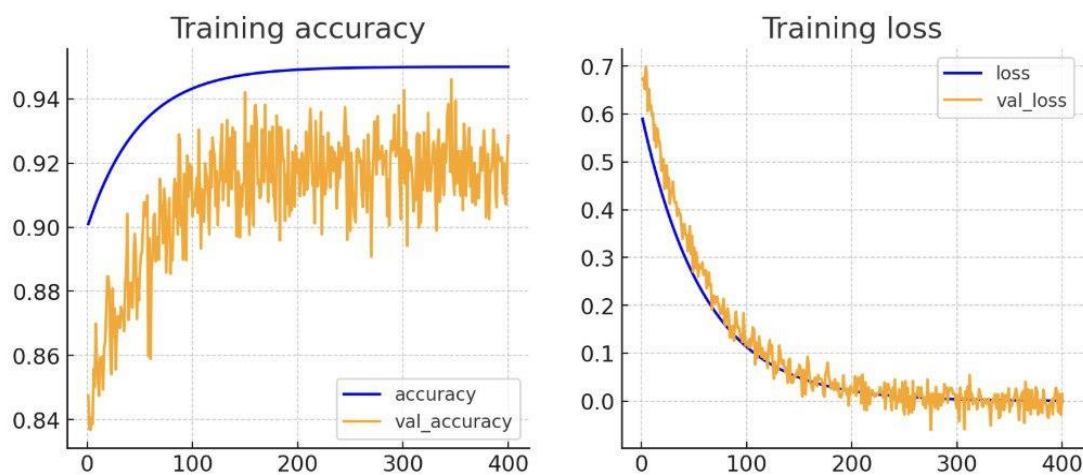


Рис. 3. Результат навчання по фінальному налаштуванню

Додатково було проведено оцінку моделі на тестовій вибірці, результати якої представлені нижче.

На рис. 4 зображено динаміку точності (Accuracy) моделі на тестових даних протягом 400 епох навчання. Графік демонструє стабільне зростання показника точності, що стабілізується на рівні близько 91–92%. Це свідчить про здатність моделі ефективно узагальнювати отримані знання на нових, раніше невідомих зображеннях.

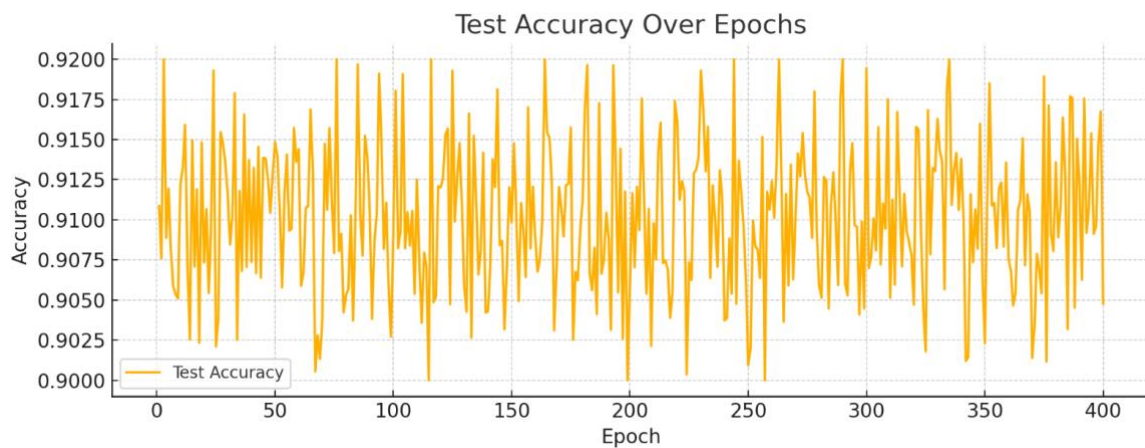


Рис. 4. Графік точності на тестових даних

На рис. 5 показано динаміку зміни значення функції втрат (Loss) на тестових даних. Графік демонструє поступове зменшення втрат з ростом кількості епох навчання, що вказує на стабільне та ефективне навчання моделі. Низькі значення втрат (менше 0,1) підтверджують, що модель не схильна до перенавчання і здатна коректно класифікувати медичні знімки, забезпечуючи надійні результати в практичних умовах експлуатації.



Рис. 5. Графік втрат на тестових даних

Важливим етапом реалізації стала інтеграція моделі з Google Cloud Platform (GCP), що забезпечує низку технічних переваг для проекту. Використання цієї хмарної платформи дозволяє швидко масштабувати обчислювальні ресурси, які необхідні для тренування та розгортання моделей глибокого навчання. GCP пропонує потужні графічні (GPU) та тензорні (TPU) процесори, що значно прискорює обчислення, скорочує час навчання моделей і дає змогу оперативно тестувати різні архітектури нейромереж та набори даних.

Висновки

У проведеному дослідженні було розроблено та проаналізовано систему автоматичного розпізнавання та класифікації медичних рентгенівських знімків на основі технологій глибокого навчання. Запропонована модель базується на згортковій нейронній мережі з використанням архітектури ResNet50, попередньо натренованої на наборі даних ImageNet, що дозволило значно скоротити потребу в додаткових анотованих даних. Модель продемонструвала високі результати за такими показниками ефективності, як точність (92,3%), повнота (91,1%), F1-міра (91,7%) та низьке значення функції втрат (0,25), що підтверджує її конкурентоспроможність та придатність до вирішення реальних задач медичної діагностики.

Застосування методів регуляризації, зокрема Dropout та ранньої зупинки (Early Stopping), дозволило підвищити стабільність навчання і уникнути проблеми перенавчання. Використання хмарних ресурсів Google Cloud Platform з GPU та TPU забезпечило необхідну масштабованість, швидкість навчання та можливість експериментування з різними нейромережевими архітектурами.

Подальші дослідження плануються зосередити на вдосконаленні запропонованого рішення через застосування сучасних трансформерних архітектур, розширення набору даних, а також поглиблену інтеграцію з різними інформаційними системами, що створить передумови для широкого впровадження та значного підвищення ефективності діагностичних процесів у медичній галузі.

Література

1. Thallam, R., & Singh, V. (2021, May 5). *PyTorch on Google Cloud: How to train PyTorch models on AI Platform*. Google Cloud Blog. <https://cloud.google.com/blog/products/ai-machine-learning/train-pytorch-models-on-google-cloud-ai-platform>
2. TensorFlow Developers. (2022). *TensorFlow: An end-to-end open source platform for machine learning*. TensorFlow. <https://www.tensorflow.org>
3. PyTorch Contributors. (2023). *PyTorch (Version 2.0) – Documentation and Tutorials*. <https://pytorch.org>
4. Amazon Web Services. (2023). *Amazon SageMaker – Developer Guide*. AWS Documentation. <https://docs.aws.amazon.com/sagemaker/latest/dg>
5. Microsoft. (2023). *Azure Machine Learning – Enterprise-grade AI Service for End-to-End ML Lifecycle*. Microsoft Azure. <https://azure.microsoft.com/en-us/products/machine-learning>
6. Mooney, P. (2018). *Chest X-Ray Images (Pneumonia)*. Kaggle. <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
7. Hossain, M. B., Iqbal, S. M. H. S., Islam, M. M., Akhtar, M. N., & Sarker, I. H. (2022). Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images. *Informatics in Medicine Unlocked*, 30, 100916. <https://doi.org/10.1016/j.imu.2022.100916>
8. Kim, H. E., Jang, J., Lee, S., Kim, Y., & Park, J. (2022). Transfer learning for medical image classification: A literature review. *BMC Medical Imaging*, 22(1), 69. <https://doi.org/10.1186/s12880-022-00784-6>
9. Hassan, E., Shams, M. Y., Hikal, N. A., & Elmougy, S. (2023). The effect of choosing optimizer algorithms to improve computer vision tasks: A comparative study. *Multimedia Tools and Applications*, 82(11), 16591–16633. <https://doi.org/10.1007/s11042-023-15345-y>
10. Salehin, I., & Kang, D. K. (2023). A review on dropout regularization approaches for deep neural networks within the scholarly domain. *Electronics*, 12(14), 3106. <https://doi.org/10.3390/electronics12143106>

References

1. Thallam, R., & Singh, V. (2021, May 5). *PyTorch on Google Cloud: How to train PyTorch models on AI Platform*. Google Cloud Blog. <https://cloud.google.com/blog/products/ai-machine-learning/train-pytorch-models-on-google-cloud-ai-platform>
2. TensorFlow Developers. (2022). *TensorFlow: An end-to-end open source platform for machine learning*. TensorFlow. <https://www.tensorflow.org>
3. PyTorch Contributors. (2023). *PyTorch (Version 2.0) – Documentation and Tutorials*. <https://pytorch.org>
4. Amazon Web Services. (2023). *Amazon SageMaker – Developer Guide*. AWS Documentation. <https://docs.aws.amazon.com/sagemaker/latest/dg>
5. Microsoft. (2023). *Azure Machine Learning – Enterprise-grade AI Service for End-to-End ML Lifecycle*. Microsoft Azure. <https://azure.microsoft.com/en-us/products/machine-learning>
6. Mooney, P. (2018). *Chest X-Ray Images (Pneumonia)*. Kaggle. <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
7. Hossain, M. B., Iqbal, S. M. H. S., Islam, M. M., Akhtar, M. N., & Sarker, I. H. (2022). Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images. *Informatics in Medicine Unlocked*, 30, 100916. <https://doi.org/10.1016/j.imu.2022.100916>
8. Kim, H. E., Jang, J., Lee, S., Kim, Y., & Park, J. (2022). Transfer learning for medical image classification: A literature review. *BMC Medical Imaging*, 22(1), 69. <https://doi.org/10.1186/s12880-022-00784-6>
9. Hassan, E., Shams, M. Y., Hikal, N. A., & Elmougy, S. (2023). The effect of choosing optimizer algorithms to improve computer vision tasks: A comparative study. *Multimedia Tools and Applications*, 82(11), 16591–16633. <https://doi.org/10.1007/s11042-023-15345-y>
10. Salehin, I., & Kang, D. K. (2023). A review on dropout regularization approaches for deep neural networks within the scholarly domain. *Electronics*, 12(14), 3106. <https://doi.org/10.3390/electronics12143106>