

ОГЛЯД СУЧАСНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ДЛЯ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ ЗА МАТЕРІАЛАМИ КОНКУРСІВ KAGGLE 2019-2020 РОКІВ

Вінницький національний технічний університет, Україна

Анотація

Комп'ютерний зір ось уже кілька десятиліть є предметом зростаючого інтересу і ретельних досліджень. Фактично, комп'ютерний зір стає більш майстерним у визначенні закономірностей по зображеннях, ніж зорова когнітивна система людини. З такими ж вражаючими досягненнями штучного інтелекту, коли технології комп'ютерного зору стають все більш поширеними в різних галузях, майбутнє комп'ютерного зору здається повним багатообіцяючих і неймовірних результатів. Проаналізовано сучасні задачі із розпізнавання зображень та їх рішення за матеріалами конкурсів 2019-2020 років на базі платформи Kaggle.

Ключові слова: комп'ютерний зір, розпізнавання зображень, змагання Kaggle, технології штучного інтелекту, передбачення.

Abstract

Computer vision has been the subject of growing interest and careful research for decades. In fact, computer vision is becoming more adept at determining patterns in images than the human cognitive system. With the same impressive advances in artificial intelligence as computer vision technology becomes more common in various fields, the future of computer vision seems to be full of promising and incredible results. Modern problems of image recognition and their solutions based on the materials of the 2019-2020 competitions based on the Kaggle platform are analyzed.

Keywords: computer vision, image recognition, Kaggle competition, artificial intelligence technologies, prediction.

Вступ

Комп'ютерний зір, або здатність систем зі штучним інтелектом «бачити», як люди, ось уже кілька десятиліть є предметом зростаючого інтересу і ретельних досліджень. Як спосіб імітації зорової системи людини, дослідження в області комп'ютерного зору спрямовані на розробку моделей, які можуть автоматизувати завдання, що вимагають візуального пізнання. Однак процес розшифрування зображень через велику кількість багатовимірних даних, які потребують аналізу, набагато складніший, ніж розуміння інших форм двійкової інформації. Це ускладнює розробку систем штучного інтелекту, здатного розпізнавати візуальні дані.

Але використання глибокого навчання і штучних нейронних мереж робить комп'ютерний зір більш здатним відтворювати людський зір. У ряді задач, комп'ютерний зір вже стає більш майстерним у визначенні закономірностей по зображеннях, ніж зорова когнітивна система людини.

З такими ж вражаючими досягненнями штучного інтелекту, коли технології комп'ютерного зору стають все більш поширеними в різних галузях, майбутнє комп'ютерного зору здається повним багатообіцяючих і неймовірних результатів.

Дана стаття присвячена аналізу сучасних задач із класифікації зображень та їх рішень, за матеріалами конкурсів 2019-2020 років на базі платформи Kaggle (<https://www.kaggle.com/>).

Результати дослідження

Сучасні технології комп'ютерного зору засновані на алгоритмах глибокого навчання, які використовують особливий вид нейронних мереж, які називаються згортковими нейронними мережами (рис. 1).

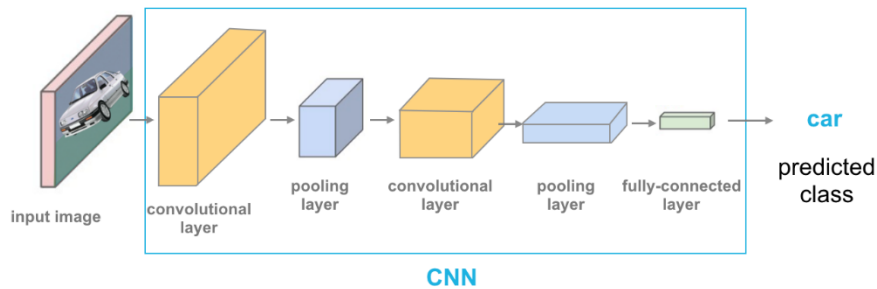


Рисунок 1 – Приклад згорткової нейронної мережі класифікації зображень [1]

Ці нейронні мережі навчаються з використанням тисяч зразків зображень, які допомагають алгоритму зрозуміти і розпізнати (класифікувати) все ключове, що міститься у зображенні. Ці нейронні мережі сканують зображення піксель за пікселем, щоб ідентифікувати шаблони і «запам'ятовувати» їх. Вони, також, запам'ятовують ідеальний результат, який повинен мати місце для кожного вхідного зображення (в разі навчання з учителем), або класифікують компоненти зображень за характеристиками сканування, такими як контури і кольори. Ця пам'ять потім використовується системою в якості довідкової під час аналізу більшої кількості зображень. І з кожною ітерацією система штучного інтелекту стає кращою в забезпеченні правильного результату.

Нейронні мережі комп'ютерного зору сьогодні застосовуються в різних галузях та сферах. За матеріалами конкурсів Kaggle за 2019-2020 роки було відібрано наступні сфери застосування комп'ютерного зору для класифікації зображень та визначено найкращі їх рішення за матеріалами, розміщеними переможцями цих конкурсів у платформі Kaggle:

- Класифікація хвороб листя маніоку: визначення типу захворювання, наявного на зображенні листя маніоку. Популярним є використання моделі класифікатора CropNet від Tensorflow Hub [2].
- Виявлення легеневої емболії: класифікація випадків легеневої емболії при комп'ютерній томографії грудної клітки. Популярним є використання архітектури нейронної мережі EfficientNet в якості основи [3].
- Виявлення прогресії легеневого фіброзу: прогнозування зниження функції легень. Переможне рішення принесло використання моделі квантильної регресії Quantile Regression та архітектури нейронної мережі EfficientNet [4].
- Класифікація меланоми: визначення меланоми на знімках. Переможним рішенням було використання ансамблю з EfficientNet, архітектури на основі стиснення se_resnetx101, та архітектури нейронної мережі resnet_101 [5]. Також популярним є використання архітектури SENet [6].
- Глобальне виявлення пшениці: класифікація видів пшениці. Популярним було використання детектора EfficientDet та архітектури Faster RCNN, реалізованої на функціональній пірамідальній мережі FPN [7], також високі результати були при використанні сімейства моделей виявлення об'єктів зі складним масштабуванням (5-тої версії) - YOLO5, але умови конкурсу не допускали його використання, тому дуже ефективні рішення були дискваліфіковані, однак, це не робить їх неефективними в широкому сенсі.
- Патологія рослин: визначення категорії позакоренових хвороб яблунь. Переможним є використання магістралі seresnextnet50 [8], також популярним у використанні є залишкова мережа для класифікації зображень ResNet та архітектура нейронної мережі EfficientNet [9].
- Рукописна класифікація бенгальської мови: класифікування компонентів рукописного бенгальського. Ефективними у використанні виявилися EfficientNet [10], xresnet, Densenet (Densely Connected Convolutional Network) [11].
- Виявлення внутрішньочерепного крововиливу: визначення гострих внутрішньочерепних крововиливів та їх підтипів. Найкращим у застосуванні виявилася нейронна мережа з тривалою короткочасною пам'яттю LSTM [12].

Крім зазначених видів архітектур велике значення має, звичайно, і вдале передоброблення даних і післяоброблення і велика кількість різноманітних прийомів, які дозволяють суттєво поліпшити ефективність існуючих інформаційних технологій, але кожен їх набір є специфічним для кожної предметної області.

Висновки

Розглянуто сучасні рішення для класифікації зображень за матеріалами конкурсів Kaggle 2019-2020 років. Перспективою розвитку розглянутих методів та моделей може бути створення аналогічних методів та моделей для зображень, зокрема із використанням архітектур EfficientNet, Faster RCNN, Densenet, різних варіантів ResNet та SENet, детектора EfficientDet, LSTM, а також багаточисленних методів перед- та післяоброблення даних.

Варто зазначити, що машинне навчання змогло вирішити багато питань вилучення ознак та обробки даних з використанням технологій комп'ютерного зору. Проте, нейронні мережі можуть знаходити різне застосування, причому не тільки для розпізнавання зображень, а й у багатьох інших сферах. Нейронні мережі здатні до навчання, завдяки чому їх можна оптимізувати і максимально збільшувати функціональність. Зазначимо, що дослідження можливостей комп'ютерного зору – це одна з найперспективніших областей в даний час, оскільки в майбутньому вони будуть застосовуватися практично повсюди, в різних областях науки і техніки, так як вони здатні значно полегшити працю, а іноді і убезпечити людей.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Cezanne Camacho. Convolutional Neural Networks. – Режим доступу: https://cezannec.github.io/Convolutional_Neural_Networks/
2. Jannis. Cassava Leaf Disease Classification. 1st place solution. – Режим доступу: <https://www.kaggle.com/c/cassava-leaf-disease-classification/discussion/221957>
3. Guanshuo Xu. RSNA STR Pulmonary Embolism Detection. 1st place solution. – Режим доступу: <https://www.kaggle.com/c/rsna-str-pulmonary-embolism-detection/discussion/194145>
4. Art. OSIC Pulmonary Fibrosis Progression. 1st place solution. – Режим доступу: <https://www.kaggle.com/c/osic-pulmonary-fibrosis-progression/discussion/189346>
5. Bo. SIIM-ISIC Melanoma Classification. 1st place solution. – Режим доступу: <https://www.kaggle.com/c/siim-isic-melanoma-classification/discussion/175412>
6. Alexandr Kalinin. SIIM-ISIC Melanoma Classification. 1st place solution in ISIC 2019 challenge. – Режим доступу: <https://www.kaggle.com/c/siim-isic-melanoma-classification/discussion/154683>
7. Dung Nb. Global Wheat Detection. 1st place solution [MIT-Compliant]. – Режим доступу: <https://www.kaggle.com/c/global-wheat-detection/discussion/172418>
8. Yelan. Plant Pathology 2020 – FGVC7. 1st place solution (single model). – Режим доступу: <https://www.kaggle.com/c/plant-pathology-2020-fgvc7/discussion/154056>
9. Leon Shanguan. Plant Pathology 2020 – FGVC7. Best single model. – Режим доступу: <https://www.kaggle.com/c/plant-pathology-2020-fgvc7/discussion/140014>
10. Deoxy. Bengali.AI Handwritten Grapheme Classification. 1st place solution. – Режим доступу: <https://www.kaggle.com/c/bengaliai-cv19/discussion/135984>
11. Dr Hb. Bengali.AI Handwritten Grapheme Classification. Best single model. – Режим доступу: <https://www.kaggle.com/c/bengaliai-cv19/discussion/123198>
12. RSNA Intracranial Hemorrhage Detection. 2nd place solution – sequential model. – Режим доступу: <https://www.kaggle.com/c/rsna-intracranial-hemorrhage-detection/discussion/117228>

Мокін Віталій Борисович – д-р техн. наук, професор, завідувач кафедри системного аналізу та інформаційних технологій, e-mail: ybmokin@gmail.com;

Коменчук Олег Вікторович – аспірант кафедри системного аналізу та інформаційних технологій, e-mail: komenchuk77@gmail.com.

Mokin Vitalii B. – Dr. Sc. (Eng.), Professor, Head of the Chair of Systems Analysis and Information Technology, e-mail: ybmokin@gmail.com;

Komenchuk Oleh V. – Post-Graduate Student of the Chair of Systems Analysis and Information Technology, e-mail: komenchuk77@gmail.com.