

Додаток А
Слайди презентації

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Атестаційна робота магістра

Дослідження нейромережевих методів для
розпізнавання людського обличчя

Науковий Керівник:

к.т.н. доц.

Чуприна А.С.

Виконав:

студент групи ІПЗм-18-1

Магда М.А.

2020

Мета дослідження:

- аналіз існуючих методів детекції обличчя на зображенні;
- аналіз проблем виявлення обличчя;
- аналіз застосування згорткових нейронних мереж;
- дослідження швидкості навчання різних архітектур мереж;
- дослідження ефективності різних архітектур мереж для складних випадків детекції.

Метод Віоли-Джонса



Wider Face (face detection benchmark):

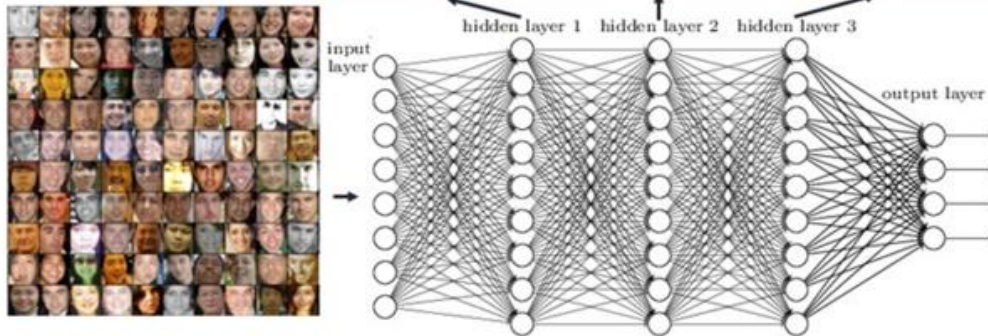
Easy: 41.2% Medium: 33.3% Hard: 13.7%

Проблеми при виявленні обличчя

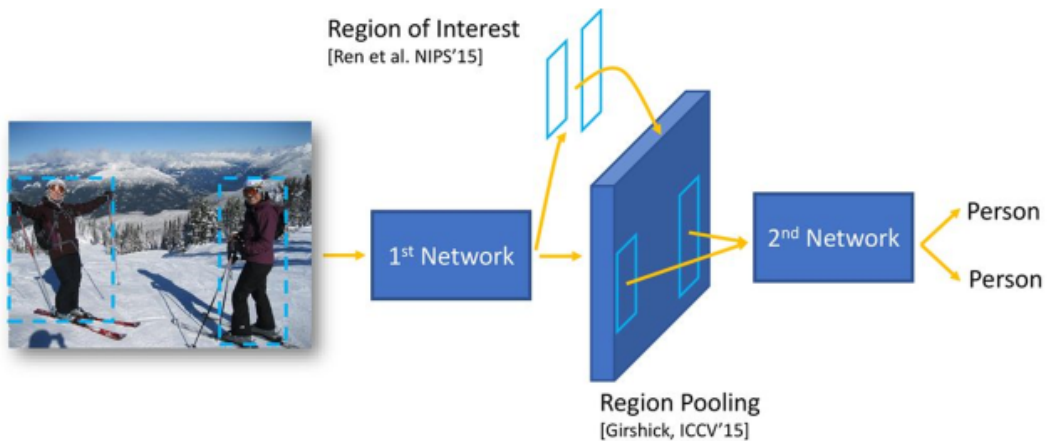


Згорткові нейронні мережі

Deep neural networks learn hierarchical feature representations

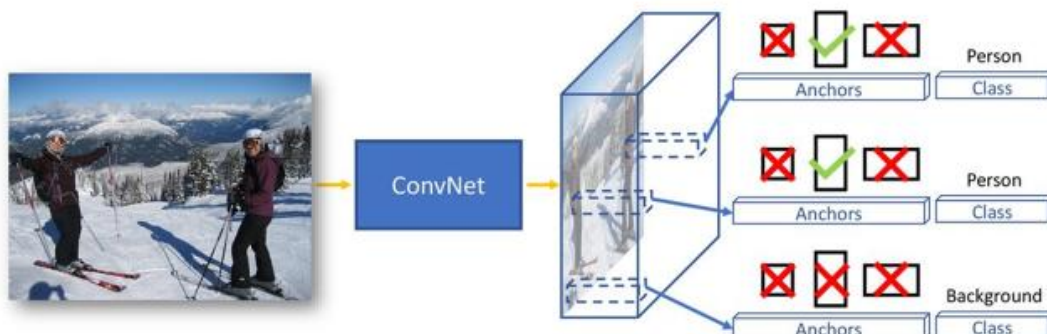


Двоступеневі детектори



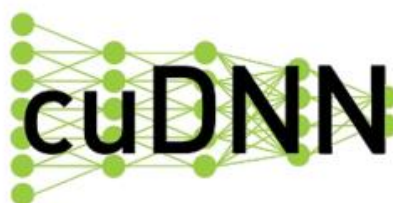
Приклади: Faster R-CNN (Inception v2), R-FCN (ResNet101)

Одноступеневі детектори



Приклади: RetinaNet (ResNet50), SSD (MobileNet v1)

Використані технології



Проведенні дослідження

- час навчання;
- розмір моделі;
- швидкість роботи;
- точність;

Аналіз результатів дослідження

Таблиця 1 – Результати дослідження швидкості навчання.

Модель	Розмір (MB)	Час навчання (хвилини)
Faster-RCNN	641 MB	40
R-FCN	623 MB	74
RetinaNet	203 MB	42
SSD	22 MB	53

Таблиця 2 – Результати оцінки якості роботи моделей.

Модель	Поза	Перекриття	Освітлення	Малий масштаб	Швидкість роботи
Faster-RCNN	0.56	0.7	0.58	0.6	416 ms
R-FCN	0.67	0.69	0.55	0.62	170 ms
RetinaNet	0.73	0.91	0.65	0.89	91 ms
SSD	0.72	0.68	0.6	0.67	22 ms

Аналіз результатів дослідження

Модель	Поза	Перекр.	Освіт.	Малі обличчя	Швидкість роботи	Час навчання	Розмір
Faster-RCNN	0.56	0.7	0.58	0.6	0	1	0
R-FCN	0.67	0.69	0.55	0.62	0.08	0	0.03
RetinaNet	0.73	0.91	0.65	0.89	0.19	0.94	0.71
SSD	0.72	0.68	0.6	0.67	1	0.62	1

Ранжування: (1, 2, 3, 4) > 5 > 7 > 6

Параметр	Поза	Перекр.	Освіт.	Масштаб	Швидкість	Час навчання	Розмір
Значення	0.2	0.2	0.2	0.2	0.1	0.07	0.03

11

Аналіз результатів дослідження

	$\alpha_1\beta_1a_{i1}$	$\alpha_2\beta_2a_{i2}$	$\alpha_3\beta_3a_{i3}$	$\alpha_4\beta_4a_{i4}$	$\alpha_5\beta_5a_{i5}$	$\alpha_6\beta_6a_{i6}$	$\alpha_7\beta_7a_{i7}$	Z^*
1	0.04179	0.05223	0.04328	0.04477	0	0.02611	0	0.20820
2	0.05	0.05149	0.04104	0.04626	0.00298	0	0.00033	0.19212
3	0.05447	0.06791	0.04850	0.06641	0.00709	0.02455	0.00794	0.27690
4	0.05373	0.05074	0.04477	0.05	0.03731	0.01619	0.01119	0.26395

$$Z^* = 0.26395 - \text{RetinaNet}$$

$$3 > 7 > 5 > (1, 2, 4, 6) \Rightarrow Z^* = 0.29182 - \text{SSD}$$

Висновки:

- проаналізовано існуючі нейромережеві методи для розпізнавання людського обличчя;
- проаналізовано проблеми виявлення обличчя;
- проаналізовано застосування згорткових нейронних мереж;
- досліджено швидкості навчання різних архітектур мереж;
- досліджено ефективність різних архітектур мереж для складних випадків детекції.

14

Подальші перспективи розвитку

- додавання нових архітектур мереж;
- додавання нових категорій зображень зі складностями;
- поповнення датасету зображень;
- створення інструменту для автоматичного розрахунку задачі прийняття рішення;

Дякую за увагу

—

Додаток Б

Тези доповіді

ОГЛЯД ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ У ЗАДАЧІ ВИЯВЛЕННЯ ОБ'ЄКТІВ

Магда Микита Андрійович

здобувач вищої освіти факультету програмної інженерії
Харківський національний університет радіоелектроніки

Науковий керівник: Чуприна Анастасія Сергіївна

канд. тех. наук, доцент, доцент кафедри програмної інженерії
Харківський національний університет радіоелектроніки

Україна

Згорткові нейронні мережи нещодавно показали неабияку ефективність у різних завданнях комп'ютерного зору, таких як виявлення та розпізнавання обличчя, класифікація об'єктів та виявлення об'єктів. Зокрема, тренована 8-шарова нейронна мережа під назвою AlexNet показала, що глибокі згорткові нейронні мережі можуть значно перевершити інші методи [1]. Мережа AlexNet 30 вересня 2012 року перемогла у змаганні по розпізнаванню образів ImageNet, досягнувши похибки top-5 у 15.3%, що майже на 11% нижче ніж попередники. Тому для вирішення завдання виявлення обличчя були обрані згорткові нейронні мережі, тому що вони забезпечують часткову стійкість до змін масштабу, зсувів, поворотам, перекриттям та інше.

В епоху глибокого навчання виявлення об'єктів можна поділити на 2 категорії: двоступеневі детектори та одноступеневі детектори, де перші роблять «грубий», «неточний» процес визначення, а другі – завершують в один крок.

Двоступеневі детектори. Спочатку набули популярності двоступеневі детектори. Першим та значним був R-CNN [2]. Ідея, що стоїть за R-CNN, проста: вона починається з вилучення набору об'єктних пропозицій (об'єктів-кандидатів) шляхом вибіркового пошуку. Потім кожен пропозицію змінюють у масштабі до фіксованого розміру і подають у модель CNN, підготовлену на ImageNet (скажімо, AlexNet) для отримання функцій. Нарешті, лінійні

класифікатори SVM використовуються для прогнозування наявності об'єкта в кожній області та для розпізнавання категорій об'єктів. Хоча R-CNN зробила великий прогрес, вона була дуже повільною через те, що потрібно було запустити CNN на великій кількості об'єктних пропозицій (близько 2000). Це приводить до дуже повільної роботи навіть на графічних процесорах (близько 14 секунд на одне зображення). Пізніше SPPNet [3] вирішив цю проблему обчислюючи карти характеристик тільки один раз для всього зображення і тоді представлення довільних областей фіксованої довжини можуть бути сформовані для навчання детекторів, що дозволяє уникнути багаторазового обчислення згорткових ознак.

Наступним кроком у розвитку була мережа Fast R-CNN, яка об'єднала найкраще від R-CNN та SPPNet [3]. Fast R-CNN дозволяє одночасно тренувати детектор та регресор обмежувальних контурів з тією ж конфігурацією мережі. Це дозволяє не втрачаючи точність (близько 65% на VOC07 датасеті) мати приріст швидкості у 200 разів порівняно з R-CNN.

У 2015 році була запропонована Faster R-CNN [4], яка використовувала Region Proposal Network – нейронну мережу для об'єктних пропозицій без втрати швидкості. Це стало причиною того, що Faster R-CNN стала першою нейронною мережею яка працює майже в реальному часі (17 кадрів на секунду).

Одноступеневі детектори. Наступним кроком розвитку була поява одноступеневих детекторів. Першим був запропонований YOLO (You only look once) в 2015 році [5]. Він надзвичайно швидкий: швидка версія спроможна обробляти близько 150 зображень на секунду, що в десятки разів перевищує двоступеневі алгоритми зберігаючи точність у 53% на VOC07 датасеті. Замість використання парадигми об'єктних пропозицій та їх верифікації, автори вирішили однією нейронною мережею розділяти зображення на регіони та прогнозувати граничні поля та ймовірності для кожного регіону одночасно. Хоча YOLO здається вирішенням проблеми детекції об'єктів, нажаль, вона дуже сильно страждає від падіння точності локалізації для невеликих об'єктів. Вирішити цю проблему намагалися наступні версії YOLO та SSD [6].

SSD (Single Shot Detector) [6] був запропонований у 2015 році та був другим одноступеневим детектором. Його головною перевагою була більша точність для маленьких об'єктів порівняно з іншими одноступеневими детекторами. Автори досягли точності у 77% на тому ж датасеті за допомогою детекції об'єктів на різних рівнях масштабу.

У 2017 році представлено RetinaNet [7]. Автори виявили причини, через які одноступеневі детектори були позаду двоступеневих за точністю. За їх словами, це через те, що при навчанні не приділялося уваги на великий контраст між об'єктами та фоном. Також вони представили нову функцію втрат «focal loss», що при навчанні приділяла більше уваги складним, не розпізнаним прикладам. Це все призвело до значного росту точності одноступеневих детекторів.

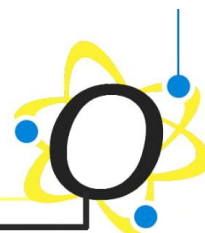
Список використаних джерел:

1. Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou & K. Q. Weinberger (ed.), *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc.
2. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *ArXiv:1311.2524 [Cs]*. Retrieved from <http://arxiv.org/abs/1311.2524>
3. K. He, X. Zhang, S. Ren, & J. Sun (2015). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9), 1904-1916.
4. S. Ren, K. He, R. Girshick, & J. Sun (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137-1149.
5. J. Redmon, S. Divvala, R. Girshick, & A. Farhadi (2016). You Only Look Once: Unified, Real-Time Object Detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779-788).

6. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. *ArXiv:1512.02325 [Cs]*, 9905, 21–37. doi: 10.1007/978-3-319-46448-0_2
7. T. Lin, P. Goyal, R. Girshick, K. He, & P. Dollár (2017). Focal Loss for Dense Object Detection. In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 2999-3007).

Додаток В

Апробація результатів роботи



Європейська наукова платформа
21037; Україна, Вінницька область,
м. Вінниця, вул. Зодчих, 18, офіс №81

ЄДРПОУ: 39965941

IBAN: UA6930268900026009055324296
МФО: 302689; Банк: «ПриватБанк»

ПІДТВЕРДЖЕННЯ ПРО УЧАСТЬ У МІЖНАРОДНІЙ НАУКОВО-ПРАКТИЧНІЙ КОНФЕРЕНЦІЇ ТА ПУБЛІКАЦІЮ НАУКОВОЇ РОБОТИ

Організаційний комітет міжнародної мультидисциплінарної науково-практичної конференції «Public communication in science: philosophical, cultural, political, economic and IT context», яка відбудеться 15 травня 2020 року у м. Х'юстон (USA), зареєстрував заяву на участь у заході.

Наукова робота, подана для публікації в рамках участі, відповідає усім встановленим вимогам, успішно пройшла оглядове рецензування та буде надрукована в збірнику наукових праць «ЛОГОΣ» за матеріалами конференції.

Відомості про роботу:

Автор(-и) роботи	Магда Микита Андрійович
Науковий керівник	Чуприна Анастасія Сергіївна
Секційний напрямок	Технічні науки та інформаційні технології
Назва роботи	ОГЛЯД ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ У ЗАДАЧІ ВИЯВЛЕННЯ ОБ'ЄКТІВ
Результати рецензування	успішно пройшла рецензування
Поточний статус	прийнята до публікації

Конференцію зареєстровано Державною науковою установою «УкрІНТЕІ» в базі даних науково-технічних заходів України та інформаційному бюлетені «План проведення наукових, науково-технічних заходів в Україні» (Посвідчення № 268 від 19.03.2020).

Захід включено до каталогу міжнародних наукових конференцій, схвалено на платформі ReserchBib та сертифіковано Euro Science Explorer™ за науковим стандартом SCC-2000. Збірнику будуть присвоєні унікальні коди ISBN та DOI, а також унікальний десятиковий класифікатор (УДК).

Матеріали конференції будуть розміщені у відкритому доступі на офіційному сайті з використанням програмного забезпечення Open Journal System на умовах ліцензії Creative Commons Attribution 4.0 International (CC BY 4.0). Бібліографічний опис матеріалів буде індексуватися в CrossRef, Google Scholar, ORCID, ResearchGate, OpenAIRE та OUCI (Open Ukrainian Citation Index).

12.05.2020

Голова Оргкомітету конференції
Голова Ради Європейської
наукової платформи
ГОЛДЕНБЛАТ МАРІЯ

