

ОБГРУНТУВАННЯ НАБОРУ ДАНИХ ДЛЯ НЕЙРОМЕРЕЖЕВОЇ ЛОКАЛІЗАЦІЇ ЧАСТИН ТІЛА ЛЮДИНИ НА ФОТОЗНІМКАХ

Вінницький національний технічний університет

Анотація

Описано основну ідею та мету нейромережевої локалізації частин тіла людини на фотознімках. Проаналізовано та оцінено різні набори даних, для майбутнього навчання та тестування нейронної мережі. Обрано датасет з зображеннями, які найбільше підходять для вирішення даної задачі.

Ключові слова: згорткова нейронна мережа, датасет, вибірка даних, зображення, частини тіла.

Abstract

The main idea and purpose of neural network localization of human body parts in photographs are described. Various data sets were analyzed and evaluated for future neural network training and testing. A dataset with images that are most suitable for this task is selected.

Keywords: convolutional neural network, dataset, data sampling, images, body parts.

Вступ

Сьогодні існує безліч задач у різних галузях, які потребують альтернативних рішень. Однією з цих задач є локалізація частин тіла людини. Ця задача може вирішити багато проблем пов'язаних з людським фактором або необачністю. Наприклад, допомога оцінки спортсменів, для суддей на спортивних змаганнях, адже техніка та майстерність відіграють важливу роль.

Найкращим вирішенням даної проблеми є побудова нейронної мережі. Для того, щоб побудувати нейронну мережу, яка буде якісно та швидко виконувати поставлену задачу, потрібно дотримуватися ряд правил та порад, які стосуються підготовки до створення та безпосередньо її реалізації.

Найголовнішою задачею перед програмною реалізацією нейронної мережі є підбір датасету, адже від того яким якісним він буде, залежить подальша розробка нейронної мережі, її навчання, і як наслідок якість та точність роботи побудованої моделі.

Результати дослідження

Зараз нейронні мережі виконують найрізноманітніші задачі, які стосуються різних галузей, від промисловості до повсякденного життя. Навіть у спорті знайшлося застосування для них. Наприклад моделі навчені грати у різноманітні настільні ігри, такі як шашки та шахи, давно перевершили майстрів свого діла серед людей. Але нейронних мереж, які допомагали б оцінювати спортсменів на різних змаганнях майже не існує. Саме тому було прийнято рішення створити модель нейронної мережі, яка визначала б частини тіла спортсмена і оцінювала її позу, тобто техніку спортсмена і правильність виконання вправ. Дана мережа допомагала б суддям оцінювати спортсменів у різних спортивних дисциплінах, наприклад гімнастиці або фехтуванні. Оскільки нейронні мережі вимагають великих ресурсів та багато часу, для початку буде створена модель розпізнавання частин тіла людини на фотознімках, яку потім можна буде модифікувати.

Для такого роду задач використовуються згорткові нейронні мережі. Для них є дуже важливим, обрати правильну вибірку, оскільки вона складається із зображень. Головною задачею даних нейронних мереж є розпізнавання образів, тому важливо підібрати датасет зображень з гарною якістю та деталізацією [1].

Для вирішення предствленої задачі було знайдено 3 датасети на яких зображені люди у різних позах, для того щоб нейронна мережа навчилася максимально якісно розпізнавати частини тіла людини.

Перший датасет - «Martial Arts, Dancing and Sports Dataset» представляє собою набір зображень спеціально створених для схожої задачі з оцінки поз людини у танцях(хіп-хоп та джаз), бойових мистецтвах(тай-чи та карате) та спортивних дисциплінах(баскетбол, волейбол, футбол, регбі, теніс та бадмінтон). Приклад зображень даного датасету зображений на рисунку 1 [2].



Рисунок 1 – Приклад зображень із датасету MADS

Як можемо бачити даний датасет містить зображення, які мають високу якість та чіткість, що є великою перевагою для майбутнього навчання нейронної мережі. Різні пози людей також відіграють важливу роль в якості роботи майбутньої моделі, адже це допоможе якісніше її навчити розпізнавати частини тіла людини.

Із недоліків даного датасету можна виділити одноманітний фон та малий розмір вибірки зображень. Фон відіграє важливу роль у навчанні, так як при наявності зображень із різним фоном, модель навчається краще і тим самим росте якість розпізнавання. Розмір вибірки зображень також впливає на якість розпізнавання створеної моделі. Чим більше зображень людей із різними позами та різним фоном, тим краще модель навчиться розпізнавати частини тіла.

Наступний датасет був запозичений із набору датасетів «СUHK». Він представляє собою набір зображень людей зроблених за допомогою камер зовнішнього відео нагляду. Приклад зображень даного датасету показаний на рисунку 2 [3].



Рисунок 2 - Приклад зображень із датасету CUHK

Вибірка складається із 13167-ми зображень, що є безумовною перевагою. Даний датасет має хорошу якість та чіткість зображень, але фон майже завжди однаковий, що є безумовно недоліком даної вибірки. Також люди на зображеннях приймають лише кілька поз, що не дозволить навчити нейронну мережу до належного рівня.

Останній датасет, який був відібраний для подальшого аналізу, це датасет «Leeds Sports Pose». Він включає в себе набір зображень людей, як із повсякденного життя так і з спортивних дисциплін. Вибірка містить 8128 зображень, як високої якості так і низької. Зображення найрізноманітніші, що є безумовною перевагою в подальшому навчанні моделі [4]. Також великим

плюсом є уже позначенні частини тіла на зображеннях, що може значно пришвидшити створення навчальної вибірки та може бути використано при навчанні моделі. Приклад зображень даного датасету показаний на рисунку 3.



Рисунок 3 - Приклад зображень із датасету Leeds Sports Pose

Єдиним недоліком даного датасету є не найкраща якість деяких зображень, але це компенсується великою кількістю наявних зображень у вибірці.

Проаналізувавши представлені датасети, вважаємо, що набір даних Leeds Sports Pose є найбільш підходящим для вирішення задачі із визначенням частин тіла людини. Саме тому навчання моделі буде здійснюватися за допомогою даного датасету. Також для побудови моделі може бути використаний набір даних MADS, але модель буде мати гіршу точність.

Висновки

У даній роботі було описано основну ідею та мету побудови нейронної мережі для визначення частин тіла людини. Було представлено кілька наборів даних відібраних для навчання моделі. Обґрунтовано всі недоліки та переваги кожного із них та відібрано один, який є найбільш підходящим та дасть можливість побудувати нейронну мережу із найкращим результатом розпізнавання частин тіла.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Модифікований метод виявлення частин тіла людини на зображеннях [Електронний ресурс] / Поліщук М. О.// Магістерська дисертація —2018. — С. 65. — Режим доступу : https://ela.kpi.ua/bitstream/123456789/26685/1/PolishchukMO_magistr.pdf. (дата звернення : 21.02.2021).
2. Martial Arts, Dancing and Sports Dataset. *Visal* : веб-сайт URL: <http://visal.cs.cityu.edu.hk/research/mads/> (дата звернення : 21.02.2021).
3. CUHK Person Re-identification Datasets. *Multimedia Laboratory* : веб-сайт URL: http://www.ee.cuhk.edu.hk/~xgwang/CUHK_identification.html (дата звернення : 21.02.2021).
4. Unite the People – Closing the Loop Between 3D and 2D Human Representations. *UP-3D* : веб-сайт URL: <http://files.is.tuebingen.mpg.de/classner/up/> (дата звернення : 22.02.2021).

Ліщук Андрій Романович – студент групи 2AKIT-176, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail : a.r.lichchuk@gmail.com

Науковий керівник: **Штовба Сергій Дмитрович** - д-р. техн. наук, професор кафедри КСУ, Вінницький національний технічний університет, м. Вінниця.

Lishchuk Andrii R. – student of 2AKIT-17b group, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail : a.r.lichchuk@gmail.com

Supervisor: **Shtovba Serhiy.** – Dr. Sc. (Eng.), Professor of Computer Control Systems Department, Vinnytsia National Technical University, Vinnytsia, shtovba@vntu.edu.ua