

Міністерство освіти і науки України
Кам'янець-Подільський національний університет імені Івана Огієнка
Фізико-математичний факультет
Кафедра комп'ютерних наук

Кваліфікаційна робота
магістра
з теми «МЕТОДИ ТА ЗАСОБИ ОРГАНІЗАЦІЇ НЕЙРОМЕРЕЖ ДЛЯ
ГЕНЕРАЦІЇ ЗОБРАЖЕНЬ»

Виконав: здобувач вищої освіти
2 курсу, групи KN1-M23
спеціальності 122 Комп'ютерні науки
Кондрушенко Микола Володимирович

Керівник: Пилипюк Тетяна Михайлівна,
кандидат фізико-математичних наук,
доцент, доцент кафедри комп'ютерних наук

Рецензент: Газдюк Катерина Петрівна,
доктор філософії з інженерії програмного
забезпечення, доцент, в.о. зав. кафедри
програмного забезпечення комп'ютерних
систем Чернівецького національного
університету імені Юрія Федьковича

м. Кам'янець-Подільський, 2024 р.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	3
ВСТУП	4
1. ТЕОРЕТИЧНІ ЗАСАДИ ОРГАНІЗАЦІЇ НЕЙРОМЕРЕЖ	7
1.1. Математичні основи генеративних моделей	7
1.2. Особливості навчання нейромереж	8
1.3. Алгоритми оптимізації та функції втрат	15
1.4. Засоби створення нейронних мереж.....	20
Висновки до розділу 1	24
2. АНАЛІЗ ОСНОВНИХ АРХІТЕКТУР ГЕНЕРАТИВНИХ НЕЙРОМЕРЕЖ	25
2.1. Аналіз сучасних тенденцій у генеруванні зображень	25
2.2. Згорткові нейромережі.....	26
2.3. Трансформери	31
2.4. Автоенкодери	34
2.5. Генеративно-змагальні мережі.....	37
2.6. Дифузійні моделі	39
Висновки до розділу 2	42
3. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ	43
3.1. Сфери використання нейромереж для генерації зображень	43
3.2. Огляд використаних програмних засобів	44
3.3. Опис проведених експериментів.....	46
3.4. Аналіз отриманих результатів.....	48
3.5. Можливості інтеграції із іншими технологіями.....	53
Висновки до розділу 3	54
ВИСНОВКИ.....	56
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	59
Додатки.....	61

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

CNN (Convolutional Neural Network) – згорткова нейронна мережа

GAN (Generative Adversarial Network) – генеративно-змагальна мережа

VAE (Variational Autoencoder) – варіаційний автоенкодер

ViT (Vision Transformer) – трансформер для обробки зображень

SGD (Stochastic Gradient Descent) – стохастичний градієнтний спуск

MLP (Multi-Layer Perceptron) – багатошаровий перцептрон

SGD (Stochastic Gradient Descent) – стохастичний градієнтний спуск

ІНМ – штучні нейронні мережі

ВСТУП

У сучасному світі нейронні мережі відіграють важливу роль у багатьох сферах життя. Однією з найбільш інноваційних та перспективних технологій є генеративні нейронні мережі, здатні створювати реалістичні зображення та інші форми контенту. З одного боку це може сприяти розв'язанню численних наукових завдань, а з іншого – розширювати творчі та креативні можливості художників, дизайнерів та розробників. Генеративні моделі можуть використовуватися для створення нових візуальних елементів, архітектурних прототипів, анімацій тощо.

Актуальність. Генеративні нейромережі використовуються, у тому числі, для створення високоякісних зображень у таких сферах, як реклама, кіно, ігри, віртуальна та доповнена реальність. Компанії застосовують їх для залучення додаткових ресурсів у сфері дизайну, це дозволяє їм значно скоротити витрати на створення маркетингових матеріалів та різного роду візуальних продуктів. Це робить дані нейромережі особливо актуальними для малого бізнесу. Також застосування даного типу нейромереж дозволяє значно скоротити час роботи із візуальною складовою маркетингу. Не менш значною перевагою є ширші можливості створення персоналізованого контенту, що значно підвищує привабливість товарів і послуг, що пливає на розвиток компаній.

Ще один важливий напрямок, де потенційно можуть використовуватися генеративні нейромережі, – медична галузь. Тут за їх допомогою можна створювати нові зображення медичних знімків для навчання моделей діагностики. Це особливо важливо для тренування нейромереж в умовах обмежених даних, адже часто важко чи неможливо отримати велику кількість реальних медичних знімків для обчислень. Завдяки такій генерації можна підвищити точність діагностики та досягти вищої ефективності медичних процедур.

Швидкий розвиток сфери машинного навчання неодмінно призводить до значного зростання потреб в обчислювальних потужностях. У свою чергу це призводить до появи нових апаратних рішень, таких як GPU більшої потужності та тензорні процесори TPU, які були спеціально розроблені для машинного навчання нейронних мереж компанією Google. Як ніколи стали актуальними дослідження в напрямку розробки оптимізованих моделей нейромереж, які б використовували мінімум ресурсів, забезпечуючи при цьому високу продуктивність роботи.

Загалом, актуальність обумовлена інноваційністю самої технології, її широкими можливостями для застосування в науковій та комерційній сферах, а також необхідністю вирішення важливих методологічних питань, пов'язаних із підвищенням ефективності, точності та якості генерації зображень.

Метою дослідження є аналіз та порівняння методів та засобів організації нейронних мереж, що використовуються для генерації зображень, а також розробка рекомендацій щодо підвищення ефективності їхнього використання.

Завдання дослідження:

- здійснити аналіз досліджень та публікацій;
- ознайомитися та здійснити вибір відповідних бібліотек;
- здійснити дослідження ШНМ для генерації зображень.

Дослідження теоретичних засад організації нейронних мереж включає дослідження математичних основ функціонування нейромереж, алгоритмів навчання, функцій активації, функцій втрат.

Аналіз сучасних підходів до організації архітектури нейромереж, призначених для роботи із зображеннями, зосереджено на згорткових нейронних мережах (CNN), глибоких нейромережах із залишковими блоками, трансформерах, які також застосовують елементи залишкових нейронних мереж, що найкраще зарекомендували себе у задачах генерації зображень.

Об'єктом дослідження є нейронні мережі, що використовуються для генерації зображень.

Предметом дослідження є методи та засоби організації нейронних мереж, їх архітектури, а також підходи до покращення якості згенерованих зображень та ефективності генерації.

Методи дослідження. Застосовані теоретичні та емпіричні методи дослідження, які включають аналіз наукової літератури, систематизацію та узагальнення інформації для визначення стану досліджуваної проблематики, проведення експериментальних досліджень, аналіз результатів.

Апробація результатів дослідження. Матеріали досліджень були оприлюднені на звітній науковій конференції за підсумками науково-дослідної роботи здобувачів вищої освіти фізико-математичного факультету Кам'янець-Подільського національного університету імені Івана Огієнка у 2023-2024 н.р., 9-10 квітня 2024 року [3] та на II Міжнародній науково-практичній інтернет-конференції «Актуальні аспекти розвитку STEAM-освіти в умовах євроінтеграції» (м. Кропивницький, 26 квітня 2024 року) [5]. Статтю «Порівняльний аналіз нейромереж-трансформерів в задачах генерації зображень» опубліковано у Віснику Кам'янець-Подільського національного університету імені Івана Огієнка. Фізико-математичні науки. Випуск 17 [4].

Структура роботи. Кваліфікаційна робота складається зі вступу, трьох розділів, висновків, списку використаних джерел, додатків.

ВИСНОВКИ

Основною метою даної роботи було визначення і аналіз основних засобів та підходів до організації генеративних нейронних мереж. У ході виконання був здійснений комплексний аналіз сучасних підходів у сфері генерації зображень. У рамках теоретичного дослідження було охоплено ключові аспекти роботи генеративних моделей, пов'язані з математичною та архітектурною складовою. Що стосується математичної складової, то були розглянуті математичні основи функціонування генеративних нейромереж, методи навчання, оптимізації, підрахунку точності і коректності роботи моделей. Щодо архітектурних рішень, то розглянуто як ті, що застосовуються у генеративних моделях, так і ті, що часто використовуються як допоміжні, для виконання окремих функцій або використовуються як основа для роботи із зображеннями чи текстовими описами.

Одним з важливих аспектів роботи стало дослідження архітектур нейромереж. Розглянуто сучасні тенденції, з чого можна виділити, що на даний момент спостерігається чітка тенденція до використання генеративно-змагальних моделей (GAN) та дифузійних моделей. Інші ж моделі, такі як автоенкодери та трансформери є скоріше допоміжними. Окремо слід виділити згорткові нейромережі (CNN). Їхні згорткові шари використовуються у більшості моделей, пов'язаних із роботою над зображеннями, в тому числі й у генеративних моделях, у цих моделях також часто застосовуються дековолюційні шари. Постійно зростає глибина нейронних мереж, а це означає, що постійно зростають вимоги до апаратного забезпечення, через потребу в обробці все більшої кількості даних та створення зображень все кращої якості та деталізації.

Для створення нейромереж існує велика кількість програмних засобів та сервісів. Деякі із цих засобів були розглянуті і використані у даній роботі. Особливо слід відзначити бібліотеки Tensorflow та PyTorch.

Теоретичні дослідження супроводжувались практичними експериментами. У роботі здійснено порівняння роботи генеративних моделей та порівняння результатів із використанням різних функцій втрат. Крім того розглянуто також можливі сфери застосування даного типу нейромереж та сучасні тенденції їхнього розвитку.

Також у роботі здійснено оцінку можливостей інтеграції генеративних моделей із іншими сучасними технологіями. На мою думку найбільш перспективно виглядає інтеграція із системами VR/AR, де генеративні моделі дозволять забезпечити більш реалістичний ефект присутності завдяки високодеталізованому згенерованому контенту. Також це дозволить досягти високого рівня персоналізації, що важливо для покращення користувацького досвіду. На друге місце я постав інтеграцію із технологіями обробки значних обсягів даних. Тут генеративні нейромережі значно допоможуть зі створенням більш наочного візуального представлення даних. Також не слід забувати про сферу маркетингу, де дані нейромережі вже займають важливу нішу та інтеграцію із програмним забезпеченням для роботи із зображеннями (Photoshop, Canva).

Для кращого розуміння принципів та особливостей роботи генеративних нейромереж було проведено ряд експериментальних досліджень. Ці експерименти найкраще демонструють особливості роботи різних генеративних нейронних мереж та як може варіюватися результат, залежно від заданих гіперпараметрів чи вибору функції втрат. Така залежність також справедлива й для інших складових архітектури, таких як: функції активації, алгоритм навчання, функції оптимізації тощо. Детально опис проведених експериментів та аналіз отриманих результатів наведено у підрозділах 3.3 і 3.4.

Експериментальні дослідження показали важливість вивчати можливості застосування різних функцій втрат при розробці генеративних моделей. У такій самій мірі це стосується й інших складових генеративних нейромереж. Потрібно підбирати глибину архітектури, розміщення шарів, функції активації, різноманітні оптимізації, гіперпараметри тощо. Усі ці

елементи тісно взаємопов'язані і тому важко точно наперед передбачити результати навчання та роботи нейронної мережі. Крім основних елементів, потрібно також ретельно підбирати критерії оцінки якості роботи нейромережі, оскільки різні метрики будуть більш чи менш показовими для різних моделей. Існує кілька універсальних метрик, які і були використані у даній роботі (FID та IS). Також під час експериментів вдалося оцінити потреби у обчислювальних ресурсах, які є значними навіть для порівняно простих моделей, які були представлені у роботі. Не дивлячись на використання потужного «заліза» наданого Google Colab все одно навчання моделей потребувало значного часу, наприклад навчання моделі StyleGAN2 могло затягнутися до 2,5 год.

На основі результатів дослідження викладених у роботі можна виділити кілька рекомендацій для подальшого розвитку генеративних моделей. Доцільно зосередити увагу на оптимізації апаратної складової під час навчання, наприклад використання кластерів із декількох GPU або TPU для прискорення процесу навчання моделей. Для досягнення кращих результатів у генерації зображень рекомендується експериментувати із комбінуванням наявних архітектур, наприклад дифузійні моделі та GAN. Перспективною є інтеграція генеративних нейромереж із технологіями машинного аналізу текстів для створення мультимодальних моделей, які працюють одночасно з текстовими та візуальними даними. Окремо слід зазначити важливість розробки більш адаптивних функцій втрат, які могли б враховувати специфіку вхідних даних та потреби кінцевих користувачів. Не менш важливими є експерименти із використанням різних функцій активації та оптимізації, їх комбінування і підбір необхідних параметрів. Нарешті важливо розробляти інтерфейси, які полегшували б інтеграцію у вже існуючі системи, що дозволить розширити сферу їх використання.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Методи та технології обчислювального інтелекту: Навчальний посібник для студ. Спеціальності 122 «Комп'ютерні науки» / І. В. Федорін; КПІ ім. Ігоря Сікорського. Київ КПІ ім. Ігоря Сікорського, 2022. 73 с.
2. Методи та системи штучного інтелекту: Навчальний посібник для студентів напряму підготовки 6.050101 «Комп'ютерні науки» / Уклад. : А. С. Савченко, О. О. Синельніков. Київ: НАУ, 2017. 109 с.
3. Микола КОНДРУШЕНКО. Ефективність глибоких згорткових нейронних мереж у розпізнаванні зображень. Збірник матеріалів наукової конференції за підсумками науково-дослідної роботи здобувачів вищої освіти фізико-математичного факультету Кам'янець-Подільського національного університету імені Івана Огієнка у 2023-2024 н.р., 9-10 квітня 2024 року [Електронний ресурс]. Кам'янець-Подільський : Кам'янець-Подільський національний університет імені Івана Огієнка, фізико-математичний факультет, 2024. С. 53-57.
4. Микола КОНДРУШЕНКО. Порівняльний аналіз нейромереж-трансформерів в задачах генерації зображень. Вісник Кам'янець-Подільського національного університету імені Івана Огієнка. Фізико-математичні науки. Випуск 17. Кам'янець-Подільський : Кам'янець-Подільський національний університет імені Івана Огієнка, 2024. С. 56-61.
5. Пилипюк Т., Кондрушенко М. Застосування згорткових нейронних мереж. Актуальні аспекти розвитку STEAM-освіти в умовах євроінтеграції: збірник матеріалів II Міжнародної науково-практичної інтернет-конференції (м. Кропивницький, 26 квітня 2024 року). Кропивницький : ДонДУВС, 2024. С. 219-220.
https://dnuvs.ukr.education/wp-content/uploads/2024/06/zbirnyk_tez_konferencziya_steam_26_04_2024.pdf
6. Sahil Sidheekh, Sriraam Natarajan. Building expressive and tractable probabilistic generative models: A Review.
7. Diederik P. Kingma, Jimmy Lei Ba. Adam: A method for stochastic optimization.

8. Review and comparison of commonly used activations functions for deep neural networks.
9. Soham De, Anirbit Mukherjee, Enayat Ullah. Convergence of adaptive gradient methods
10. Yarin Gal, Zoubin Ghahramani. Dropout as a Bayesian Approximation: representing model uncertainty in deep learning.
11. Sergey Ioffe, Christian Szegedy. Batch Normalization: accelerating deep network training by reducing internal covariate shift.
12. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classifications with deep convolutional neural networks. In Advances in neural information processing systems.
13. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional Networks for large-scale image recognition.
14. Attention is all you need 12 Jun 2017 (перевидано 2 Aug 2023).
15. An image is worth 16x16 words: transformers for image recognition at scale 22 Oct 2020 (перевидано 3 Jun 2021).
16. Deep residual learning for image recognition 10 Dec 2015.
17. Swin Transformer: hierarchical vision transformer using shifted windows 17 Aug 2021.
18. Swin transformer 2: scaling up capacity and resolution. Apr 2022.
19. Diederik P. Kingma, Max Welling. Auto-Encoding Variational Bayes. 10 Dec 2022.
20. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Shafiq J. Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. 10 Jun 2014.
21. Jonathan Ho, Ajay Jain, Pieter Abbeel. Denoising Diffusion Probabilistic Models. 16 Dec 2020.
22. Alex Nichol, Prafulla Dhariwal. Improved Denoising Diffusion Probabilistic Models. 18 Feb 2021.