

Міністерство освіти і науки України
Криворізький національний університет
Факультет інформаційних технологій
Кафедра автоматизації, комп'ютерних наук і технологій

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття ступеня вищої освіти – бакалавр

за освітньо-професійною програмою

«Комп'ютерні науки»

зі спеціальності

122 – Комп'ютерні науки

тема роботи:

«Розпізнавання та генерація графіки на основі штучного інтелекту»

Виконав студент гр. КН-20 _____ Нужда А. Р.

Керівник _____ Кумченко Ю. О.

Нормоконтроль _____ Маринич І. А.

Завідувач кафедри _____ Рубан С. А.

Кривий Ріг – 2024

КРИВОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет: інформаційних технологій

Кафедра: автоматизації, комп'ютерних наук і технологій

Ступінь вищої освіти: Бакалавр

Спеціальність: 122 – Комп'ютерні науки

ЗАТВЕРДЖУЮ

Зав. кафедрою: к.т.н. Рубан С.А.

« 27 » березня 2024 р.

ЗАВДАННЯ

на кваліфікаційну роботу бакалавра

студентові групи КН-20 Нужді Артему Романовичу

1. Тема кваліфікаційної роботи: «Розпізнавання та генерація графіки на основі штучного інтелекту»

затверджено наказом по університету № 235с від 27.03.2024 р.

2. Термін здачі кваліфікаційної роботи: 05.06.2024 р.

3. Склад кваліфікаційної роботи: Пояснювальна записка обсягом 50с., додатки, презентація у Microsoft PowerPoint (8 слайдів) в електронному та друкованому вигляді

4. Консультанти кваліфікаційної роботи:

Розділ 1-2

доц. Кумченко Ю. О.

Нормоконтроль

доц. Маринич І. А.

6. Консультанти розділів роботи

5. Календарний план:

| № | Етапи роботи | Термін виконання |
|---|---|------------------|
| 1 | <i>Вступ</i> | <i>01.04.24</i> |
| 2 | <i>Розділ 1</i> | <i>05.04.24</i> |
| 3 | <i>Розділ 2</i> | <i>01.05.24</i> |
| 4 | <i>Висновки</i> | <i>25.05.24</i> |
| 5 | <i>Оформлення кваліфікаційної роботи</i> | <i>28.05.24</i> |
| 6 | <i>Підготовка презентації та графічного матеріалу</i> | <i>20.05.24</i> |
| 7 | <i>Підготовка доповіді до захисту</i> | <i>05.06.24</i> |

6. Дата видачі завдання: 29.01.2024р.

Керівник _____ / Кумченко Ю. О./

7. Запевнення: Я, Нужда Артем Романовий, запевняю, що ця кваліфікаційна робота виконана самостійно, не містить академічного плагіату, фабрикації, фальсифікації. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

Із чинним Положенням про академічну доброчесність Криворізького національного університету ознайомлений.

Чітко усвідомлюю, що в разі виявлення у кваліфікаційній роботі умисних порушень робота не допускається до захисту або оцінюється незадовільно.

Студент _____ / Нужда А. Р./

АНОТАЦІЯ

Нужда А. Р./ Розпізнавання та генерація графіки на основі штучного інтелекту Кваліфікаційна робота на здобуття ступеня вищої освіти – бакалавр, за спеціальністю 122 Комп'ютерні науки. Криворізький національний університет, Кривий Ріг, 2024.

Робота складається зі вступу, двох розділів, висновків, переліку використаної літератури з 20 позицій та 1 додатки. Загальний обсяг роботи становить 61 сторінок, з яких основний зміст роботи викладено на 51 сторінках, включає 4 таблиць і 32 рисунки.

У даній кваліфікаційній роботі бакалавра було проведено всебічне дослідження можливостей ШІ в області розпізнавання та генерації графіки. Зокрема, проаналізовано як позитивні, так і негативні сторони використання таких технологій. Серед позитивних аспектів можна виділити підвищення ефективності процесів, скорочення часу на створення графічних матеріалів та можливість генерувати високоякісні зображення без потреби в людському втручанні. З іншого боку, не менш важливим є вивчення негативних сторін, таких як можливість зловживання технологіями, етичні питання та ризики втрати робочих місць у сфері творчих професій.

Крім теоретичного аналізу, в роботі було розроблено програмні рішення для розпізнавання та генерації зображень. Ці програми демонструють практичне застосування алгоритмів ШІ та їхню здатність вирішувати конкретні задачі, пов'язані з графікою. Розроблені програми дозволяють автоматизувати процеси обробки та створення зображень, що відкриває нові можливості для розвитку різних галузей.

Таким чином, ця дипломна робота має на меті не лише розкрити потенціал ШІ у сфері графіки, але й підкреслити важливість обережного та відповідального підходу до впровадження новітніх технологій.

СИСТЕМА КОМП'ЮТЕРНОГО ЗОРУ, ШТУЧНИЙ ІНТЕЛЕКТ,
НЕЙРОННА МЕРЕЖА, НАВЧАННЯ З ВЧИТЕЛЕМ, DATASET IMAGENET,
БІБЛІОТЕКА TENSORFLOW.

ЗМІСТ

| | |
|---|----|
| ВСТУП..... | 7 |
| РОЗДІЛ 1 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ ШТУЧНОГО ІНТЕЛЕКТУ | 8 |
| 1.1 Проблематика штучного інтелекту..... | 8 |
| 1.2 Визначення штучного інтелекту | 9 |
| 1.3 Чи можна вважати текст, створений ChatGPT, книжкою | 12 |
| 1.4 Проблеми з фейками та книгами, створеними ШІ..... | 13 |
| 1.5 Штучний інтелект і робота художників: небезпека чи ні | 15 |
| 1.6 Використання ШІ у кіновиробництві..... | 19 |
| 1.7 Етичні питання та захист авторських прав..... | 20 |
| 1.8 Історія розвитку комп'ютерного зору | 21 |
| 1.9 Нейронні мережі для мобільних пристроїв | 24 |
| Висновки до розділу..... | 27 |
| РОЗДІЛ 2 РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ТА ГЕНЕРАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ | 28 |
| 2.1 Архітектура згорткових нейронних мереж | 28 |
| 2.2 Середовище навчання нейронної мережі..... | 29 |
| 2.2.1 TensorFlow..... | 30 |
| 2.2.2 Keras..... | 30 |
| 2.2.3 TensorFlow Lite | 31 |
| 2.2.4 TensorFlow API Detection..... | 31 |
| 2.2.5 Google Colaboratory | 31 |
| 2.3 Виявлення об'єкту | 32 |
| 2.3.1 Оцінка достовірності..... | 34 |
| 2.3.2 Розташування об'єктів на зображенні | 36 |
| 2.4 Навчання мережі..... | 36 |
| 2.5 Донавчання мережі..... | 37 |

| | |
|---|----|
| 2.6 Встановлення моделі на Android смартфон..... | 37 |
| 2.7 Тестування моделі розпізнавання об'єктів | 39 |
| 2.8 Нейромережі для генерації зображень..... | 41 |
| 2.8.1 Leonardo AI | 43 |
| 2.8.2 BING IMAGE CREATOR | 45 |
| 2.8.3 DREAM ВІД WOMBO | 46 |
| 2.8.4 CRAIYON..... | 47 |
| 2.8.5 Ideogram..... | 47 |
| 2.9 Розробка додатку генерації зображень за текстовим запитом за допомогою ШІ48 | |
| Висновки до розділу..... | 55 |
| ВИСНОВКИ..... | 57 |
| СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ | 58 |
| Додаток А. Лістинг Кода | 60 |

ВСТУП

В сучасному світі технології штучного інтелекту (ШІ) стрімко розвиваються і все більше інтегруються в різні сфери людської діяльності. Одним з найбільш перспективних напрямків є використання ШІ для розпізнавання та генерації графічних зображень. Ця тема набуває все більшого значення завдяки широкому спектру можливих застосувань, від медицини до розваг, від дизайну до безпеки.

Штучний інтелект має здатність аналізувати велику кількість даних і виявляти в них закономірності, що робить його незамінним інструментом для розпізнавання зображень. Завдяки методам глибокого навчання, нейронні мережі можуть досягати високої точності у класифікації та ідентифікації об'єктів на зображеннях. Це відкриває можливості для створення систем, здатних автоматично розпізнавати обличчя, визначати діагнози за медичними знімками, аналізувати сцени відеоспостереження та багато іншого.

Крім розпізнавання, ШІ також успішно застосовується для генерації нових графічних зображень. Зокрема, генеративно-змагальні мережі (Generative Adversarial Networks, GANs) демонструють вражаючі результати у створенні фотореалістичних зображень, генерації нових дизайнерських рішень, створенні анімацій та навіть в мистецтві. GANs навчаються створювати зображення, що виглядають як реальні, шляхом змагання двох нейронних мереж — генератора та дискримінатора, що дозволяє досягати високої якості синтезованих зображень.

РОЗДІЛ 1. ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ ШТУЧНОГО ІНТЕЛЕКТУ

1.1 Проблематика штучного інтелекту

Впродовж 2023 року розпочалися суперечки щодо впливу штучного інтелекту на мистецтво. Деякі українські митці вже висловлюють занепокоєння, вважаючи, що такі інструменти, як Midjourney, знецінюють саме поняття мистецтва. Наприклад Amazon почав відслідковувати книги які, можливо, були створені за допомогою ШІ та видаляти їх з платформи.



Рисунок 1.1 – Колаж Вікторія Железна, згенеровано штучним інтелектом

Протягом цього року найбільші технокомпанії представили свої версії чатботів зі штучним інтелектом. Першим був ChatGPT від OpenAI, який став доступний в Україні у лютому. Потім Google влітку показав світу свій чатбот Bard а взимку запустив найдосконалішу модель ШІ ім'я якої Gemini Microsoft випустив чатбот Copilot, здатний генерувати пісні.

Стрімкий розвиток і розповсюдження ШІ впливають на нашу мову. Впливові словники з США та Великобританії визначили свої слова року, і більшість з них пов'язані з цією технологією. Наприклад, Collins назвав "штучний інтелект" словом 2023 року, а Кембриджський словник виділив слово

"галюцинувати" (hallucinate), що також пов'язане з ШІ. Переліки популярних слів року від Оксфордського словника та Merriam-Webster також включають терміни, пов'язані з ШІ [1].

1.2 Визначення штучного інтелекту

Хоча штучний інтелект поступово проникає в наше повсякденне життя, багато хто все ще асоціює його з науково-фантастичною літературою або чимось надто далеким і невідомим. Проте ми щодня користуємося цими інструментами, часто не усвідомлюючи цього.

Наприклад, Google Maps та його аналоги використовують ШІ для аналізу швидкості руху транспорту, для попередження користувачів про аварії та ремонтні роботи. Spotify аналізує наші музичні уподобання і пропонує треки, які можуть сподобатися.

Під штучним інтелектом зазвичай мають на увазі комплекс технологій, які дозволяють виконувати "інтелектуальні" завдання без людського втручання. Нещодавні досягнення в галузі ШІ вивели його на новий рівень: тепер він здатен генерувати зв'язні тексти, відповідати на запитання та створювати ілюстрації.

Найвідоміші інструменти

- MidJourney, DALL-E, Stable Diffusion – для створення цифрових зображень;
- ChatGPT, Bard, Gemini – для написання текстів та пошуку відповідей на запитання.

Хоча розробники стверджують, що данні інструменти полегшують життя, Художники глибоко стурбовані стрімким прогресом технологій.

Один із страхів музикантів полягає в тому, що штучний інтелект може позбавити їх роботи. Штучний інтелект вже активно використовується в музичній індустрії через кілька методів:

- інструменти (схожі на автотюн), що покращують музику;
- інструменти, які генерують відео з зовнішнім виглядом артистів;
- інструменти, які генерують музику, схожу працю інших музикантів.

Кожен із цих інструментів може становити загрозу для митців і технічних працівників індустрії, вважає професор музичних технологій з Університету Індіани (США) Джейсон Паламара.

Штучний інтелект навчається на величезних базах даних існуючої музики, це дозволяє йому досліджувати музичні структури, динаміку, гармонії, мелодії, ритми, тембри та форми, а також генерувати новий матеріал. Наприклад, початківець може скористатися програмою на зразок GarageBand або придбати набір бітів та семплів у мережі й записати вокал через смартфон.



Рисунок 1.2 – Інтерфейс програми GarageBand [18]

Данні інструменти є надзвичайно корисними для новачків, які часто стикаються з нестачею досвіду, навичок або ресурсів для звернення до професіоналів у галузі зведення та мастерингу. Наприклад, користувачі програми Boomtу можуть вибрати з кількох варіантів, таких як Global Groove або Rap Beats, для створення інструментальних партій. Після цього їх можна переаранжувати або додати вокальні треки. Застосунок SongStarter від BandLab дозволяє створювати інструментальні композиції на основі введеного тексту та смайликів [1].

Ці технології спрощують процес створення музики, це дозволяє музикантам і продюсерам зосередитися на творчості, доручивши технічні завдання штучному

інтелекту. Видання Time розповідає про новий інструмент YouTube від Google DeepMind під назвою Lygia. Користувачі можуть ввести, наприклад, "Балада про те, як протилежності притягуються, оптимістична акустика", і миттєво отримати уривок пісні у стилі Чарлі Пута.

"Якщо штучний інтелект може миттєво створити 'пісню Чарлі Пута', що це означає для самого Чарлі Пута або для інших музикантів-початківців, які бояться, що їх замінять?", — міркує видання. Музичні автори отримують роялті від стримінгових платформ. Наприклад, Spotify виплачує 70% заробітку від прослуховувань авторам композицій. Якщо компанія зможе генерувати музику за допомогою ШІ, це може означати більше прибутків для неї та менше для музикантів, зменшуючи людський вклад у створення музики.

Деніел Ек - генеральний директор Spotify повідомив, що не має наміру повністю блокувати контент, створений штучним інтелектом. Він вважає, що ця технологія має місце в музичній індустрії, однак її використання для підробки справжніх виконавців без їхнього дозволу є неприпустимим. ШІ відтворює голоси зі старих записів

Штучний інтелект допомагає відтворювати голоси або музичні партії зі старих записів. Так для останньої пісні The Beatles — Now and Then, голос Джона Леннона було відтворено. Демозапис було зроблено наприкінці 1970-х років у його домі в Нью-Йорку. Для завершення пісні Пол Маккартні скористався штучним інтелектом. Продюсер Родні Джеркінс для витягання аудіозапису Ol' Dirty Bastard з касети VHS для треку SZA використав ШІ.

Тепер до "Греммі" допускаються пісні з елементами, створеними ШІ. Основна вимога полягає в тому, що авторство більшої частини музикального твору або тексту має належати людині. Наприклад, сонграйтер Ghostwriter номінував трек "Heart on My Sleeve", який був створений за допомогою штучного інтелекту та включав голоси Дрейка і The Weeknd.

Директор Warner Music Group Роберт Кінкл зауважив: "Ви повинні прийняти технологію, тому що не можете зупинити її розвиток". Музика, згенерована ШІ, залишається в "сірій" зоні через складнощі з правами на пісні, які імітують голоси виконавців без використання захищених текстів або аудіозаписів.

Universal Music, ABKCO і Concord Publishing звернулися до суду штату

Теннессі з позовом проти компанії Anthropic, звинувачуючи її в порушенні прав видавців. Причиною стало використання текстів більше 500 пісень для тренування свого чатбота Claude.

1.3 Чи можна вважати текст, створений ChatGPT, книжкою

У лютому 2023 року Reuters розповіли історію Бретта Шиклера, який за допомогою ChatGPT створив 30-сторінкову ілюстровану дитячу книжку за кілька годин і виставив її на продаж на Amazon. Хоча його прибуток становив менше 100 доларів, це надихнуло його на написання ще книг за допомогою ШІ.

"Я бачив, як люди роблять на цьому кар'єру", — сказав Шиклер, який використовував підказки в ChatGPT, такі як "напишіть історію про батька, який навчає свого сина фінансової грамотності".

Це лише один приклад того, як ШІ впливає на літературу. Люди, які раніше і не мріяли написати книгу, можуть швидко створювати тексти та публікувати їх онлайн.



Рисунок 1.3 – Створена ШІ обкладинка книги "Хочу на Марс" [19].

В основному це стосується обкладинок і ілюстрацій, але видавництво "Ранок" вже опублікувало книгу, повністю створену нейромережею під наглядом редакторки Мар'яни Горянської та дизайнера Олександра Ковалєвського. Ця книга має назву "Хочу на Марс".

Видавництва, які вже використовують ШІ для створення обкладинок:

1."Темпора": "Краще не читай" Катерини Орловської, "Правік та інші часи" Ольги Токарчук.

2."Видавництво Старого Лева": малюнки до книжки Марини Пономаренко "Книжка любові та люті".

."Бородатий Тамарин": Анне Свердруп-Тайгесон "Жаль, дзижчи, кусай. Навіщо нам комахи".

Kotibook також розглядає можливість використання ШІ для обкладинок.

У коментарі для видання "Читомо" засновник видавництва "Прометей" Андрій Носач зазначив, що ШІ може бути корисним для створення прототипів обкладинок, але творчий потенціал професійних ілюстраторів часто пропонує більш оригінальні рішення.

1.4 Проблеми з фейками та книгами, створеними ШІ

Amazon видалив п'ять книг, підписаних ім'ям Джейн Фрідман, хоча вона не мала жодного стосунку до їх створення. Також постраждав колишній кореспондент BBC Рорі Селлан-Джонс, біографію якого написав невідомий автор і виклав на Amazon без його відома.

На платформі з'явилися книги про "містичну смерть" Вікторії Амеліної, в яких в анотації не згадується війна. В описі зазначено, що "це глибоке дослідження життя, творчості та загадкової смерті видатної української письменниці", і також згадана країна, яка "витримує шторми як внутрішніх, так і зовнішніх конфліктів".

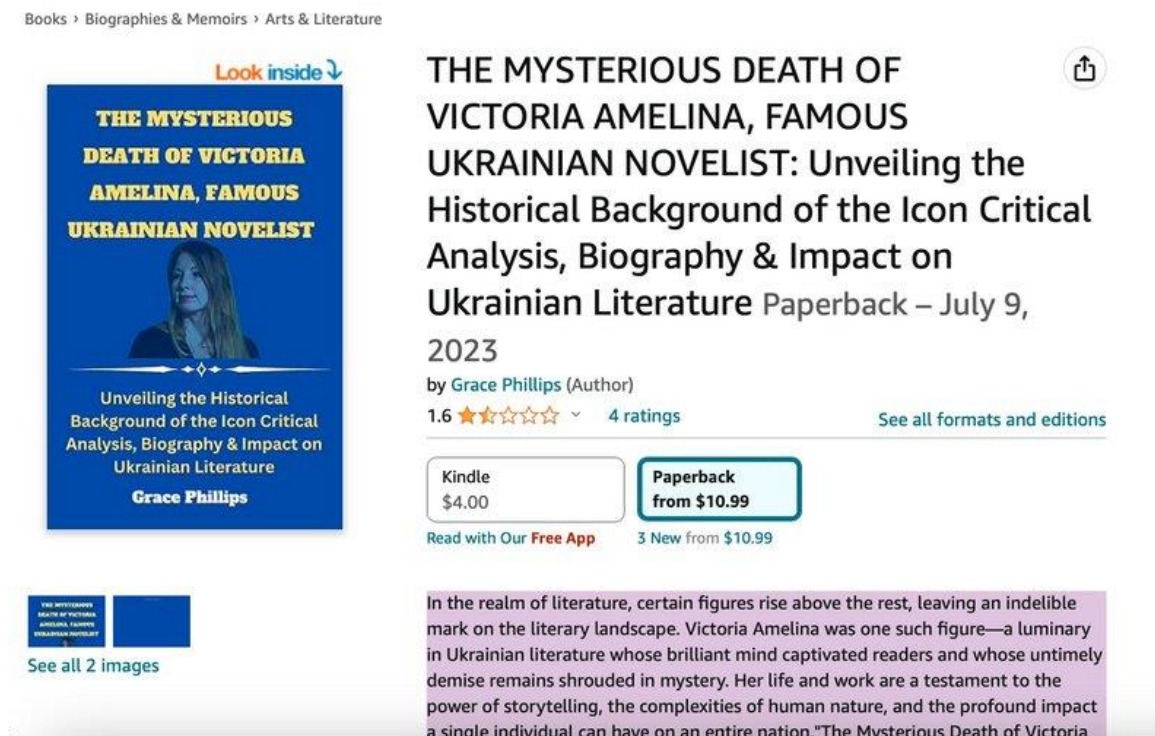


Рисунок 1.4 – Книга про смерть Вікторії Амеліної, створена ШІ. Скриншот зі сторінки Amazon [20]

Після низки скандалів, пов'язаних з книгами, створеними штучним інтелектом, Amazon запровадив нове обмеження: автори можуть публікувати не більше трьох книг на день. Однак, поки що не ясно, чи це допоможе уникнути появи текстів, згенерованих ChatGPT.

Інструменти штучного інтелекту здатні створювати відверті фальсифікації. Наприклад, до дня народження українського письменника Олеся Гончара "Новий канал" опублікував його біографію, згенеровану ChatGPT, яка містила неправдиві факти. Це не перший випадок: раніше в мережі поширювалися фальшиві біографії Тараса Шевченка, стверджуючи, що він був одним з перших українців, хто займався фотографією[1]. Гільдія авторів США звернулася з відкритим листом до генеральних директорів компаній, які займаються розробкою штучного інтелекту. Письменники закликають отримувати згоду авторів перед використанням їхніх творів для "навчання" ШІ. Лист підтримали понад 8000 авторів, включаючи Маргарет Етвуд, Нору Робертс, Дена Брауна, Дженніфер Іган, Луїз Ердріх, Джеймса Паттерсона, Сюзанну Коллінз, Селесте Інг, Лі Мін-Джин, Ендрю Соломона та інших.

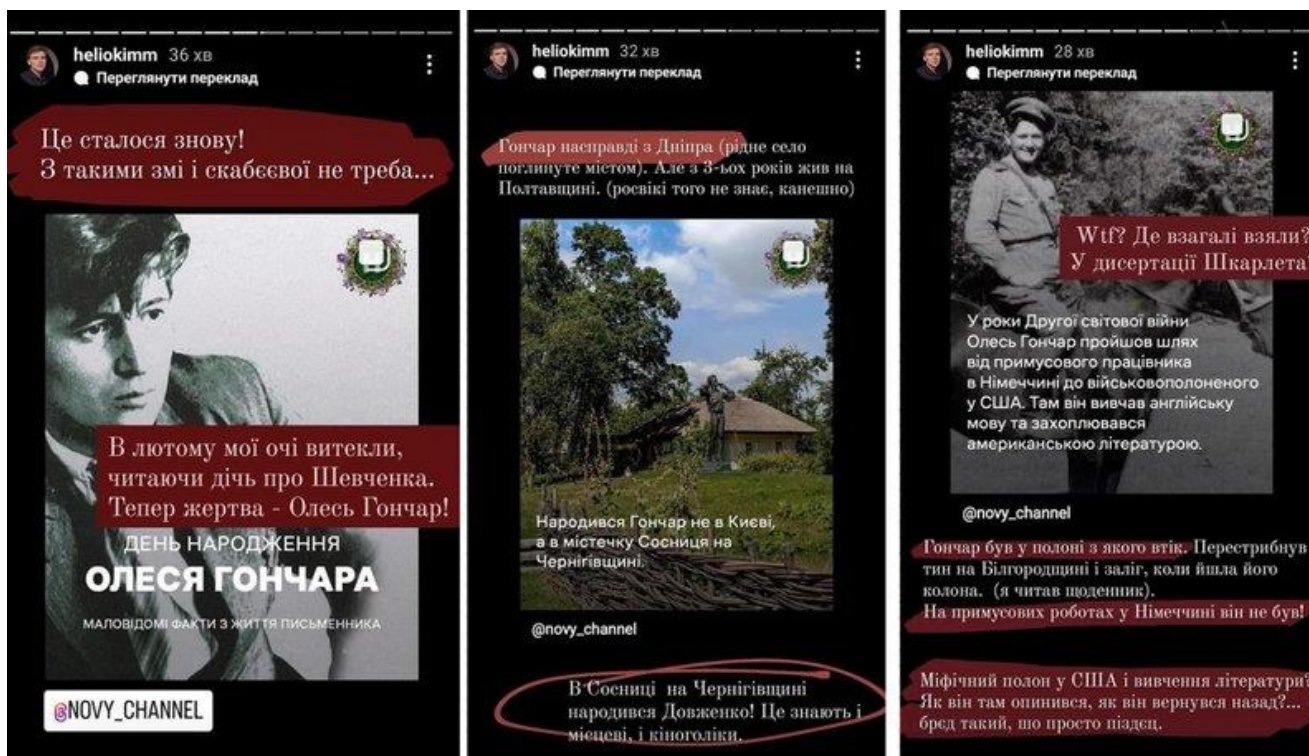


Рисунок 1.5 – Фейкові відомості про Гончара, створені ШІ Скріншот сторіз Олександра Єскименка в Instagram

Заклик до регулювання ШІ

У вересні американські письменники Джонатан Франзен, Джон Грішем, Джордж Мартін та інші подали позов проти компанії OpenAI, стверджуючи, що їхні книги використовувалися для "навчання" ChatGPT без дозволу.

Думка Оксани Забужко про ШІ в літературі

Письменниця Оксана Забужко заявила, що вона впевнена ШІ не зможе повністю замінити письменників ні зараз, ні в майбутньому. Вона вважає, що програми не здатні створювати тексти так, як це роблять люди: "Боятися є чого, але не того. Таким чином збільшуються можливості генерування брехні. А саме створення штучної реальності. Створення крипто-пам'яті, підсовування людям, скажімо, того, що вони насправді не зазнали, як їхніх власних переживань"[1].

1.5 Штучний інтелект і робота художників: небезпека чи ні

Художники, мабуть, найбільше стурбовані впливом інструментів штучного інтелекту на свою професію. Цього року в X (колишній Twitter) розгорілася дискусія про те, чи можна вважати згенеровані ШІ ілюстрації мистецтвом, чи

етично їх використовувати, і чи може людина називатися художником, якщо вона просто натискає кілька кнопок, щоб отримати готове зображення. Багато хто також критикував недосконалість цих інструментів: зображення людей з додатковими пальцями чи відсутніми руками, а також схожість на роботи інших митців.

Проте письменник і співзасновник видавництва "Бородатий Тамарин" Макс Кідрук захищає використання ШІ для створення обкладинок. "MidJourney не забере роботу у таких митців, як Катерина Кошелева, яка малювала ілюстрації для «Колонії», бо ніколи — ніколи! — навіть не наблизиться до її рівня. Він також не забере роботу у Федора Гонци, який створював для нас тривимірні карти," — зазначив Кідрук у своєму пості на Facebook [1].

Він підкреслює, що MidJourney не розуміє, що саме робить, і хоча навчається на роботах інших митців, автори-початківці роблять те саме. "На нейронному рівні ми є не меншими «плагіаторами», ніж ШІ," — додає він. Водночас Кідрук визнає, що проблема з ліцензуванням та авторським правом існує.

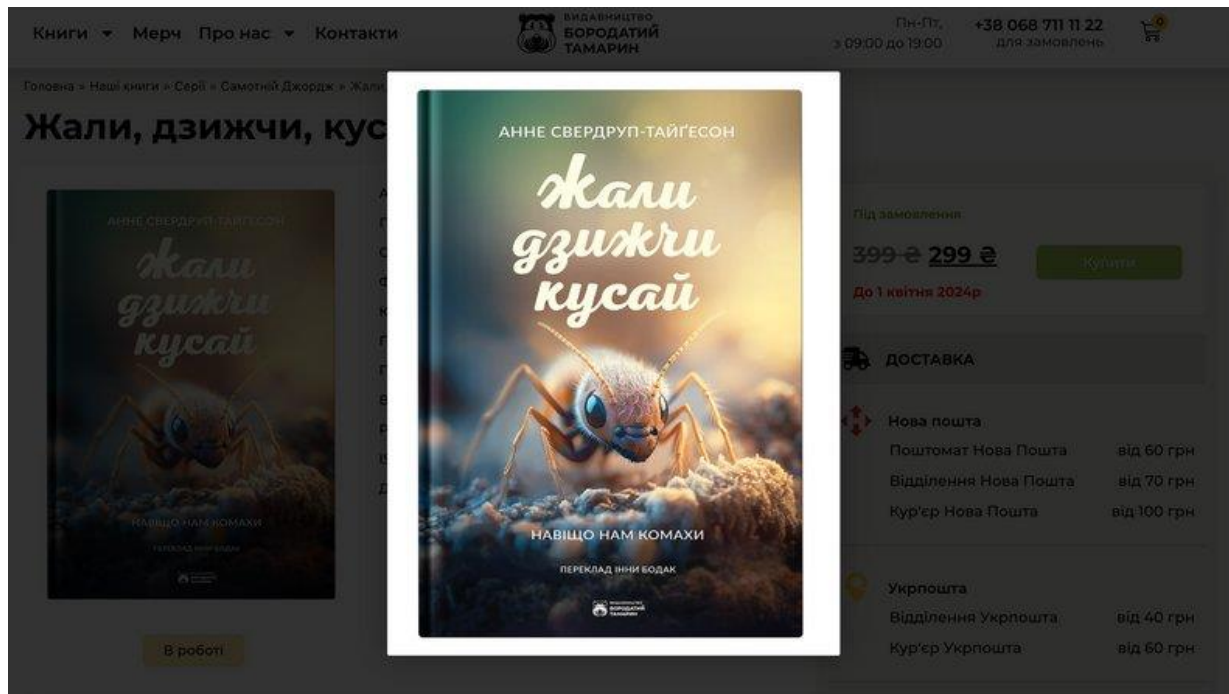


Рисунок 1.6 – Зображення попередньої обкладинки, створене ШІ та допрацьоване художницею Бородатий тамарин

Ілля Стронговський - співзасновник "Видавництво" має більш скептичне ставлення до використання таких інструментів, особливо через питання

авторського права. Він зауважує, що ШІ використовують дані з глобальної мережі, не враховуючи, чи захищені ці зображення авторським правом.

Саша Биченко - Дизайнер візуальних комунікацій, співзасновник та креативний директор студії Kultura зазначає, що не вбачає у цьому "проблеми або зради". На його думку, це скоріше цікавий експеримент для видавців, авторів та всієї індустрії, і не варто робити передчасні висновки щодо нової технології.

Він також додає, що згенеровані ШІ ілюстрації часто позбавлені метафор та розуміння контексту, оскільки неймережі імітують, а не створюють.

Критики ШІ стверджують, що використання генераторів робить мистецтво надто доступним і простим, що призводить до зниження його цінності. Вони вказують на відсутність емоційної глибини та автентичності в таких зображеннях, які є характерними для людської творчості. ШІ не здатний передати складність людських емоцій або індивідуальність людського самовираження. Крім того, створені ним ілюстрації не мають унікального авторського стилю, оскільки генератори аналізують мільйони зображень і видають щось загальне та середньостатистичне..

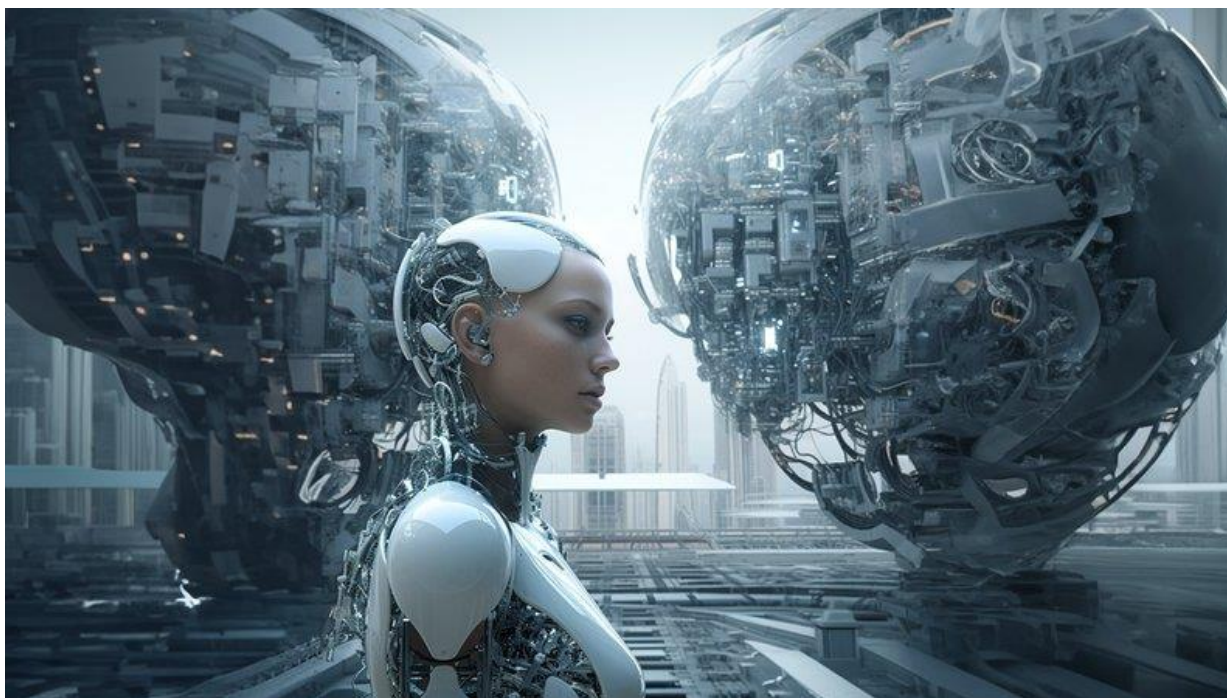


Рисунок 1.7 – Образ штучного інтелекту: як він бачить себе

Зображення створене за запитом штучного інтелекту показати себе самого. У цьому випадку ШІ уявив себе у вигляді жінки. плив ШІ на традиційний артринок

Збільшення кількості ілюстрацій, створених за допомогою ШІ, може дестабілізувати традиційний артринок. Оскільки ці інструменти стають більш доступними, це може призвести до зниження цінності традиційного мистецтва в очах суспільства [1].

Упередженість ШІ та соціальні контексти Однією з небезпек є те, що ШІ часто не враховує соціальні контексти. Наприклад, тестування Midjourney компанією Craft Clarity виявило упередженість: на промт "фото CEO" програма видавала виключно зображення білих чоловіків, а на промт "щаслива пара" більшість фото видавало гетеросексуальні пари однакової раси. Якщо не враховувати ці аспекти, можна підкрипити гендерні та культурні стереотипи.

Сара Андерсен, Келлі Маккернан та Карла Ортіс у 2023 році художниці подали позов проти Stability AI, DeviantArt та Midjourney, звинувачуючи їх у використанні захищених авторським правом матеріалів для навчання генераторів зображень. Позов також стверджує, що ШІ-інструменти порушують права на публічність та закони Каліфорнії про недобросовісну конкуренцію. Юристка Ольга Петрів зазначила, що хоча Сара Андерсен створила унікального персонажа, виграти цю справу буде складно, оскільки ШІ навчався на величезній кількості інформації з інтернету.

Використання ШІ у кіновиробництві Штучний інтелект вже активно використовується у створенні фільмів, зокрема для написання сценаріїв, генерування ідей та планування зйомок. Наприклад, індійський режисер Шекхар Капур використав ChatGPT для написання сценарію до свого фільму. Warner Bros. і 20th Century Fox використовують ШІ для прогнозування успіху фільмів та підбору контенту для певної аудиторії.

Протести акторів та сценаристів проти ШІ У 2023 році сценаристи та актори Голлівуду вийшли на страйк, вимагаючи регулювання використання ШІ у творчих проектах. Вони прагнули гарантій, що ШІ не буде використовуватися для написання або переписування сценаріїв без дозволу авторів, а також що їхні обличчя та голоси не будуть незаконно використовуватися для створення контенту [1].

Приклади зловживань ШІ у кіноіндустрії Британський актор Стівен Фрай виявив, що його голос використали для озвучування фільму без його відома. Актор

Том Генкс став жертвою незаконного використання його обличчя у відеорекламі стоматологічного страхування.

Загрози та можливості ШІ Незважаючи на протести, деякі експерти бачать у ШІ потенціал для революції у кіновиробництві, демократизації процесу створення фільмів та надання можливостей талановитим людям з обмеженими ресурсами.

Том Генкс застеріг, що не має жодного відношення до реклами стоматологічного страхування. Скриншот з його сторінки в Instagram



Рисунок 1.8 – Том Генкс та попередження про шахрайську рекламу

Позови акторів та митців проти ШІ

Актриса Скарлетт Йоганссон пішла до суду проти додатку ШІ Lisa AI: 90s Yearbook & Avatar, який використав її ім'я та зовнішній вигляд у 22-секундній рекламі без її дозволу.

1.6 Використання ШІ у кіновиробництві

У фільмі "Таємне вторгнення" (Secret Invasion) для створення початкових титрів використовували ШІ, що викликало жваве обговорення в соціальних

мережах. Джефф Сімпсон, який працював над візуальною частиною, вважає ШІ неетичним, небезпечним та "створеним для того, щоб знищити кар'єри художників".

Індійський актор Аніл Капур проти ШІ

Актор з Індії Аніл Капур надав позов до Верховного суду країни через численні спотворені відеоролики, гіфки та емодзі з його зображенням. Він заборонив ШІ використовувати його обличчя та голос, зазначивши: "Це не тільки для мене. Сьогодні я захищаю себе, але коли мене не буде, моя сім'я повинна мати право захищати мою особистість і отримувати з цього вигоду в майбутньому".

Виклики ШІ у мистецтві: чи є причина для страху?

У чому полягають основні проблеми з використанням ШІ у творчості? Митці та ілюстратори в соцмережах висловлюють побоювання, що все може стати занадто одноманітним. Оскільки ШІ навчається на величезній кількості зображень, книг і пісень з інтернету, він видає щось середньостатистичне, яке, хоч і приваблює для широкої аудиторії, може знецінити унікальність людського мистецтва.

1.7 Етичні питання та захист авторських прав

Серед нерозв'язаних проблем – етичні аспекти використання ШІ. Як захистити авторів від копіювання їхніх робіт? Як боротися з крадіжкою окремих елементів? Поки що жодна країна не знайшла ефективного рішення.

Позитивні аспекти ШІ у творчості

Попри негативні сценарії, інструменти ШІ можуть сприяти творчому процесу. Вони можуть використовуватися як прототипи для проєктів, допомагати у мозковому штурмі ідей для блогів, планувати фотосесії, шукати референси та надихати ілюстраторів.

Використання ШІ для натхнення та створення

ШІ може також надихати на створення нових проєктів. Наприклад, Ілля Стронговський пропонує створювати приватні датасети з ілюстрацій самого автора або з зображень, очищених від прав. Це дозволить генерувати прототипи малюнків у власному стилі, які можна творчо доопрацювати.

Визнання мистецтва, створеного ШІ

Деякі мистецтвознавці стверджують, що розвиток мистецтва, створеного ШІ, підвищує цінність людського мистецтва, оскільки воно є більш унікальним.

Конкурси та фестивалі для ШІ-мистецтва

Незважаючи на всі виклики, появляються змагання та виставках для відео та фотографій, створених за допомогою ШІ. Наприклад, на Ballarat International Foto Biennale були вручені нагороди за найкращі зображення, згенеровані ШІ. Головну нагороду отримала художниця із Швеції Анніці Норденшельд за її фото "Закохані сестри-близнючки" [1].



Рисунок 1.9 – Фото переможця премії. Фото: ballaratfoto/Instagram

1.8 Історія розвитку комп'ютерного зору

Загалом, історія розвитку комп'ютерного зору відображає історію розвитку штучного інтелекту. Обидва напрямки спираються на біологічні процеси, що базуються на вивченні роботи мозку. Поява штучних нейронних мереж, вдосконалення обчислювальних ресурсів та збільшення обсягу даних сприяли просуванню обох напрямків.

Багато з наших знань про зорове сприйняття сьогодні має свої корені у нейрофізіологічних дослідженнях, проведених на кішках у 1950-х і 1960-х роках.

Вчені помітили ієрархічність в роботі нейронів: від реакції на прості ознаки, такі як краї, до більш складних, таких як форми, і нарешті до складних візуальних уявлень.

В основу систем комп'ютерного зору було покладено реконструкцію неврологічних структур людини в цифровій формі. Ці системи також використовують ієрархічний підхід до обробки та аналізу візуальних даних, подібно до біологічних аналогів.

Штучний інтелект виник як область наукових досліджень під час літнього семінару, що відбувся в Дартмутському коледжі в 1956 році. Вчені об'єднали різні напрямки з метою розвитку ідей про мислячі машини. Перші дослідники були вкрай оптимістичними стосовно можливостей цих областей та вбачали у них потенціал змінити світ. Однак несправедливо високі очікування із боку громадськості та фінансові труднощі призвели до різкого спаду в інтересі до штучного інтелекту.

Початкові обчислювальні ресурси були недостатні для вирішення складних завдань, які ставили перед собою дослідники. Більшість роботи велися в ізольованих групах без належної наукової підтримки.

Згодом, уряди та корпорації втратили інтерес до штучного інтелекту, що спричинило "Зиму штучного інтелекту". Хоча в 1980-90-х роках спостерігався певний прогрес, багато серйозних дослідників уникали терміну "штучний інтелект".

З появою Інтернету та зростанням обчислювальних можливостей збільшилася доступність даних. Разом з тим, розвиток нейронних мереж та алгоритмів удосконалив можливості штучного інтелекту. У 2012 році проривом стало вирішення задачі візуального розпізнавання на конкурсі ImageNet Large Scale Visual Recognition Challenge (ILSVRC) (рис. 2.1) .

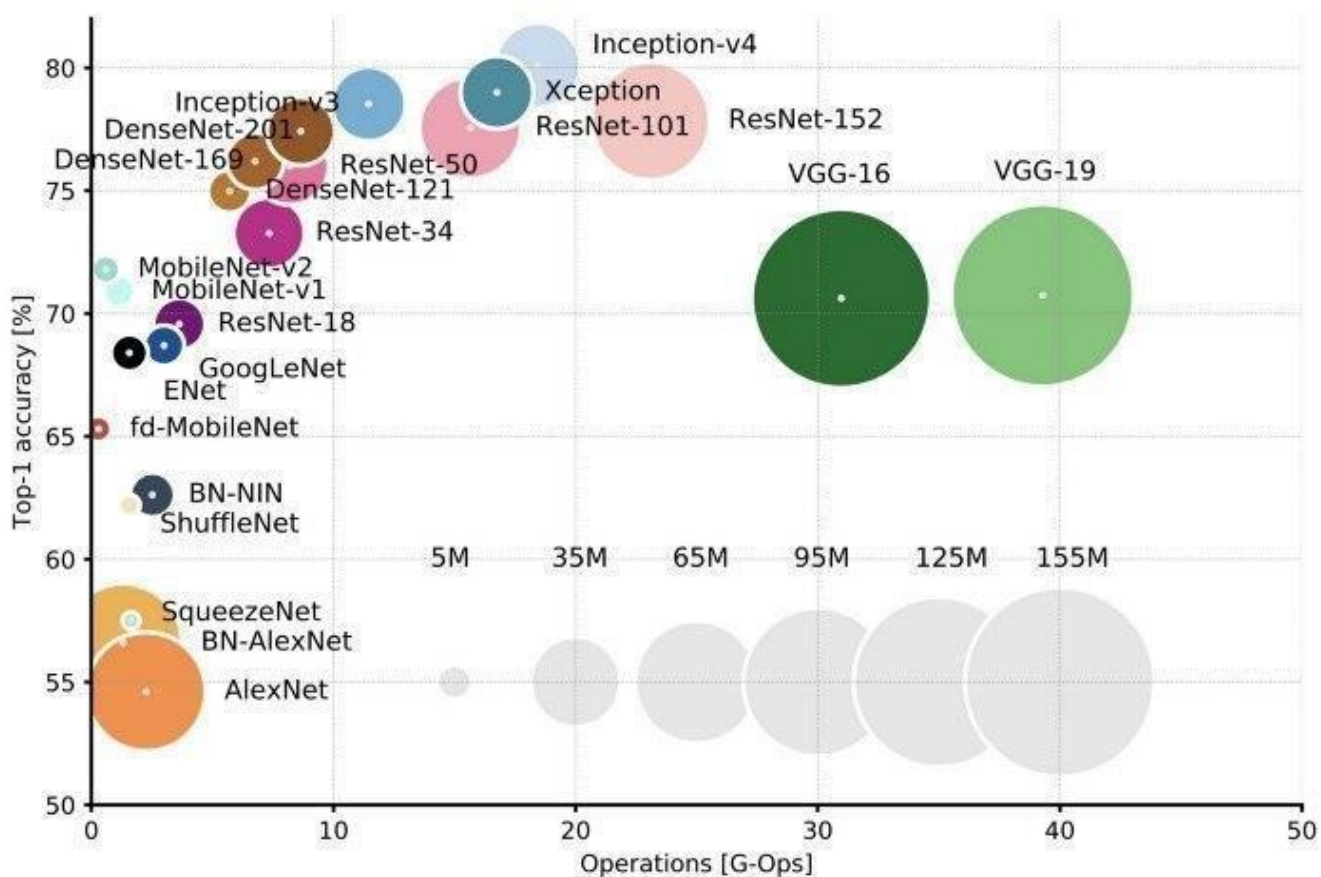


Рисунок 1.10 – Графік відображає залежність точності найвпливовіших нейронних мереж від кількості операцій, які вони виконують

Ця нейронна мережа, відома як AlexNet, була розроблена для класифікації зображень і складається з п'яти загорткових шарів і трьох повністю зв'язаних шарів. Завдяки глибоким нейронним мережам, як AlexNet, відбулася революція у сфері штучного інтелекту. AlexNet досягла помилок на рівні 16,4%, а подальші розробки знизили рівень помилок у конкурсі ILSRVC до кількох відсотків, встановивши глибокі нейронні мережі як стандарт для розпізнавання зображень. Цей успіх відкрив шлях для поширення штучного інтелекту в Кремнієвій долині.

Штучний інтелект впроваджується в різноманітні сфери нашого життя як невід'ємний елемент. Наприклад, системи штучного інтелекту, такі як AlphaGo, використовуються в ігровій індустрії для розвитку нових стратегій. У сфері охорони здоров'я алгоритми глибокого навчання застосовуються для фільтрації шуму в слухових апаратах, а також у медичній діагностиці для класифікації зображень та розпізнавання образів. Штучний інтелект також застосовується в різних галузях, таких як роздрібна торгівля, де системи комп'ютерного зору використовуються для виявлення дефектів у виробництві або для автоматичного

розрахунку покупок без касира.

Хоча ми досягли значних успіхів у цих галузях, потрібно розуміти, що ще багато часу пройде, перш ніж комп'ютери зможуть інтерпретувати зображення так само ефективно, як люди. Проте у найближчому майбутньому штучний інтелект ймовірно буде використовуватися для розширення можливостей людей, як у медицині, так і в інших галузях.

1.9 Нейронні мережі для мобільних пристроїв

Нейронні мережі, спеціально розроблені для мобільних пристроїв, представляють наступне покоління еволюції штучного інтелекту. Завдяки подальшому розвитку виконання логічних операцій, тепер можна використовувати їх на пристроях із низьким енергоспоживанням та обмеженою пам'яттю. Представлено методи підвищення продуктивності виконання, зберігаючи при цьому точність результатів.

SqueezeNet була випущена в 2016 році і використовує новий будівельний блок під назвою "пожежний модуль". Цей модуль складається з 1×1 згорткового шару, щоб зменшити кількість каналів з мінімальною кількістю параметрів, та наступного поєднання 1×1 і 3×3 згорткових шарів, що утворює розширений блок, який знову збільшує глибину мережі. Крім цього, цей модуль використовує концепцію відкладеної дискретизації, що призначена для зниження роздільної здатності карт функцій у пізніших етапах мережі, щоб підвищити точність.

SqueezeNext - це вдосконалена версія SqueezeNet з рядом змін. По-перше, вона використовує більш ефективний метод стискання каналу за допомогою двоетапного модуля стискання, що суттєво зменшує загальну кількість параметрів для згорток 3×3 . Крім того, вона застосовує окремі згортки розміром 3×3 , що сприяє зменшенню обсягу моделі, а також усуває додатковий 1×1 модуль після етапу стиснення. До того ж, мережа включає елементне з'єднання з пропуском додавання, схоже на архітектуру ResNet.

ShuffleNet використовує дві нові операції, такі як поелементна групова згортка та перетасування каналів, для зменшення обчислювальних витрат,

зберігаючи при цьому точність.

MobileNet, що з'явився у квітні 2017 року, побудований на концепції роздільних згорток по глибині, яка вперше була представлена мережею Xception. Цей обчислювальний блок складається з першого рівня ядер, які опрацьовують кожен вхідний канал окремо, разом з глибинною згорткою.

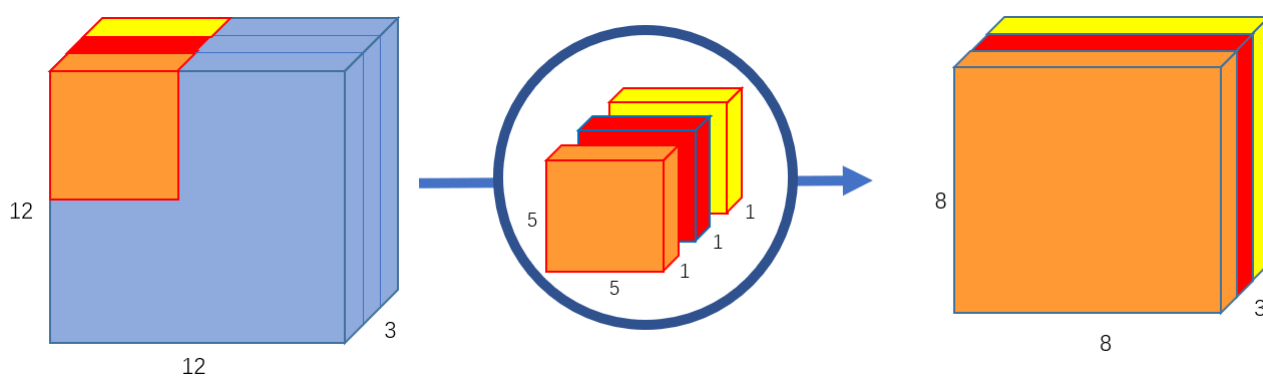


Рисунок 1.11 Операція згортки у глибину,

За процесом на(Рисунок 1.11) на якою слідує інший, точковий (Рисунок 1.12), де використовується ядро розміром 1x1: ядро, що застосовується до кожної точки. Це ядро має глибину, яка відповідає кількості каналів у вхідному зображенні.

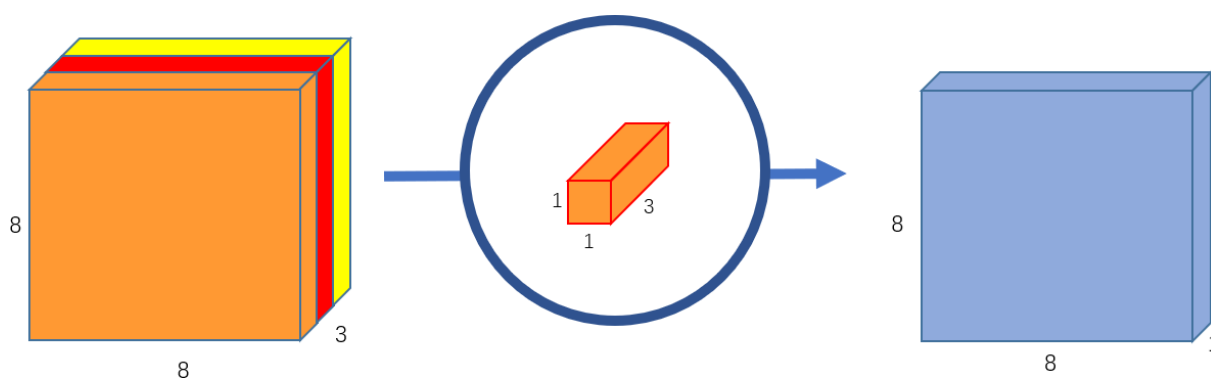
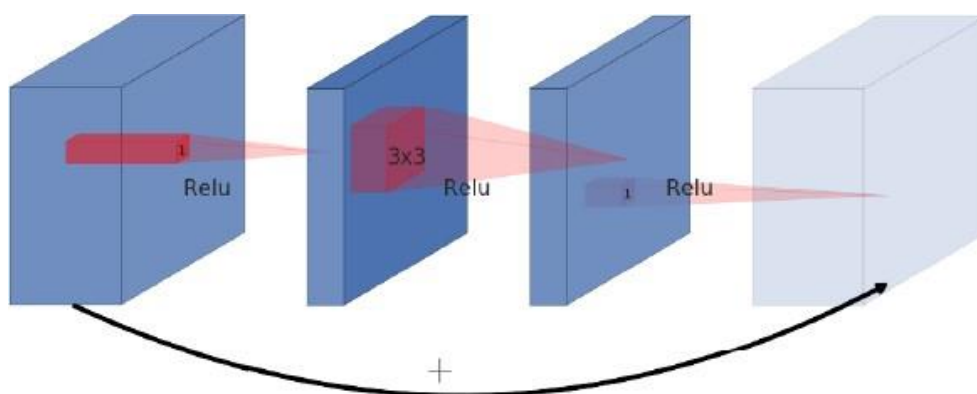


Рисунок 1.12 – Операція точкової згортки

MobileNet також має гіперпараметр, відомий як множник ширини, який регулює кількість вихідних каналів наприкінці кожного блоку, а також загальну кількість параметрів у моделі.

Через рік з'явилася MobileNetV2. У версії V2 було додано лінійні шари вузьких місць та інвертовані залишкові блоки. Ці блоки об'єднують старт і фініш згорткового блоку завдяки пропуску з'єднань. Додавання цих елементів дає

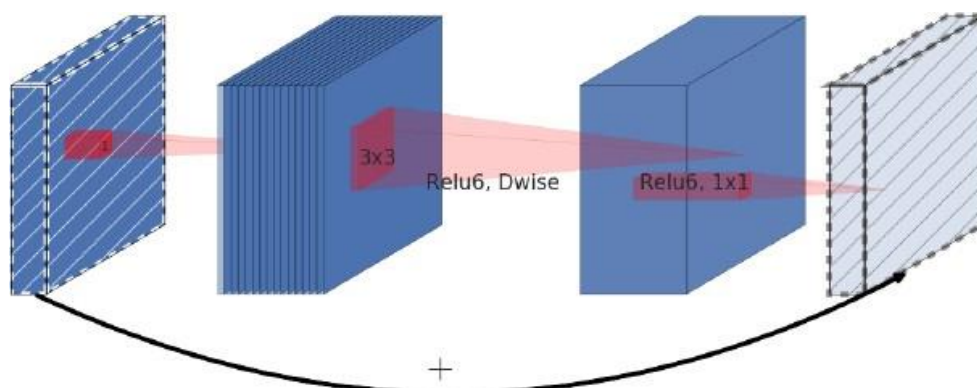
мережі можливість отримувати доступ до попередніх активацій, які не зазнали змін у згортковому блоці. Ця стратегія стала ключовою у розробці глибоких мереж. Вихідний залишковий блок дотримується схеми широкий->вузький->широкий або вузький->широкий->вузький (рис. 1.13) якщо звернути увагу на кількість каналів. Автори називають цей підхід перевернутим залишковим блоком, оскільки між вузькими частинами мережі існують пропускні з'єднання, що є протилежним до роботи початкових залишкових з'єднань. З перевернутими залишковими блоками ми стискаємо шари, де є пропускні з'єднання, що погіршує продуктивність мережі. Автори висунули концепцію лінійного стиснутого каналу, де остання згортка залишкового блоку має прямий вихід перед додаванням до



вихідних активацій.

Рисунок 1.13 – Розвиток архітектур блоків роздільної згортки.

Пошук нейронної архітектури (NAS) – це сучасна та перспективна галузь, яка має на меті використання алгоритмів пошуку, зазвичай методів підкріпленого навчання або еволюційної оптимізації, для знаходження оптимальної комбінації модулів, що забезпечує найкращий компроміс між затримкою та точністю.



Наприклад, для визначення метрики оптимізації можна використовувати багатоцільову функцію Парето, як це демонструється в архітектурі MnasNet. Інші

приклади важливих досягнень у пошуку архітектури мобільних мереж включають FBNet, ChamNet і MobileNetV3. Для детального аналізу цієї техніки можна звернутися до відповідних джерел.

EfficientNet значно покращує можливості тонкого налаштування моделі, надаючи користувачам змогу вибирати оптимальний баланс між ефективністю та точністю. Автори впроваджують інтегрований метод масштабування гіперпараметрів, який одночасно коригує глибину мережі (кількість шарів або блоків), ширину (кількість каналів) та розмір вхідного зображення, впливаючи на роздільну здатність внутрішніх карт функцій. Цей підхід до масштабування базується на тому, що незалежні зміни в розмірах мережі не впливають на баланс між затримкою і точністю. Нарешті, цей новий метод масштабування застосовується до базової архітектури мережі, знайденої шляхом багатоцільової оптимізації та визначеного простору пошуку, як в MnasNet.

Висновки до розділу

Штучний інтелект вже є невід'ємною частиною нашої реальності. Нейромережі продовжуватимуть свій розвиток, а людство навчиться співіснувати з ними, приймаючи нові виклики і намагаючись їх контролювати.

Колись персональні комп'ютери та смартфони здавалися дивними новинками, а художники боялися, що фотоапарати заберуть у них роботу. Однак усі ці технології стали невід'ємною частиною нашого життя та продовжують еволюціонувати.

РОЗДІЛ 2.

РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ТА ГЕНЕРАЦІЇ
ЗОБРАЖЕНЬ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ

2.1 Архітектура згорткових нейронних мереж

Протягом останнього десятиліття згорткові нейронні мережі (CNN) стали провідними моделями штучного інтелекту для надточного розпізнавання об'єктів. Принцип роботи CNN нагадує функціонування людського ока (рис.3.1). Згорткові нейронні мережі (ЗНМ) здебільшого складаються з двох ключових компонентів: елементів для вилучення ознак та класифікатора. Елементи для вилучення ознак містять згорткові шари, де фільтр ковзає по вхідній матриці, і шари максимального об'єднання, які зменшують розмір матриць, створених попередніми згортковими шарами, до більш компактних форм. Об'єднання, як правило, здійснюється шляхом вибору максимального або середнього значення в заданій області. Операція об'єднання, як і операція згортки, розбиває матрицю на фрагменти, а потім застосовує об'єднання поетапно. Наприклад, операція об'єднання може розбити згортку на фрагменти розміром 2x2 із кроком 1x1.

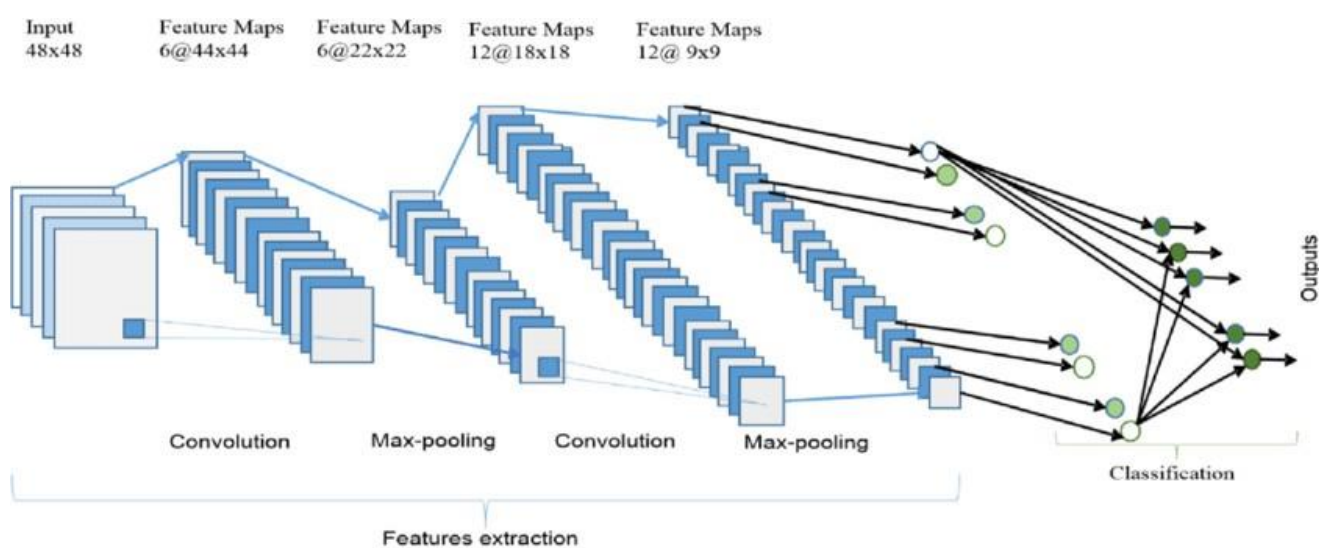


Рисунок 2.1 – Архітектура згорткових нейронних мереж (CNN)

На рисунку 2.2, показується як відбуваються чотири операції об'єднання. Останній рівень згорткової нейронної мережі (CNN) генерує вихідні дані, які надходять на вхід до повністю підключеної мережі, що виконує функцію класифікації.

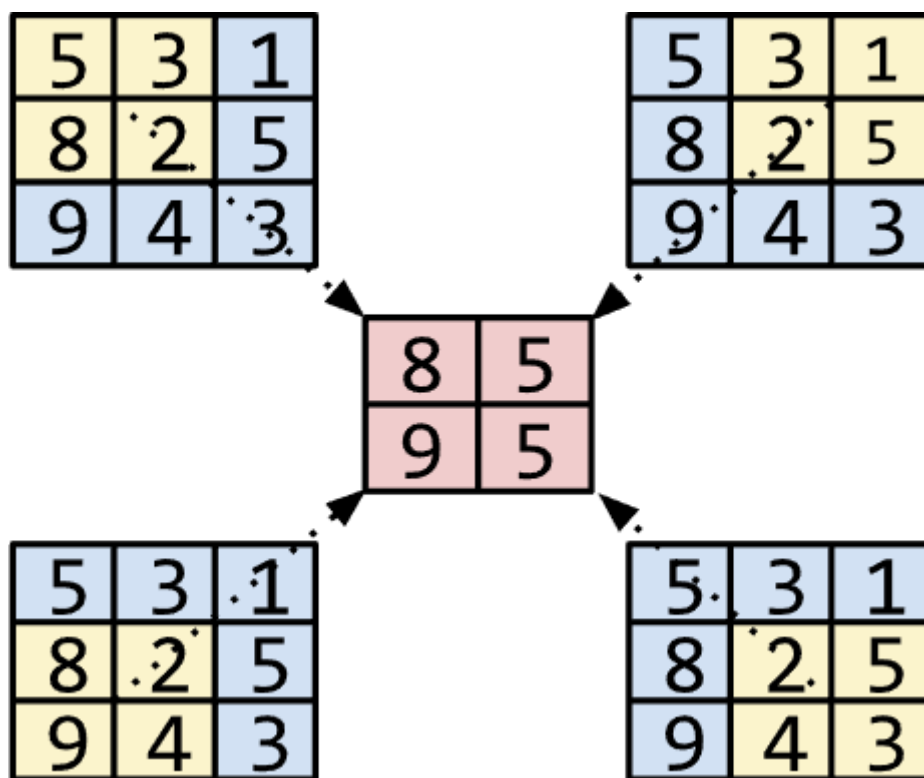


Рисунок 2.2 – Процедура об'єднання

Об'єднання та його роль в CNN

Об'єднання сприяє досягненню трансляційної інваріантності в початковій матриці, що означає здатність алгоритму ефективно класифікувати зображення, навіть якщо позиція об'єктів на зображенні змінюється. Наприклад, алгоритм може незмінно впізнати собаку незалежно від того, чи розташована вона у центрі кадру чи на його лівому краю.

2.2 Середовище навчання нейронної мережі

Для впорядкування процесу навчання використовувалася операційна система Windows 10 x64, хоча цей процес можна виконувати в різних

середовищах. Середовище від Microsoft було вибрано через сумісність із деякими програмами та фреймворками, які будуть описані у наступних розділах.

2.2.1 TensorFlow

Фреймворк TensorFlow, розроблений командою Google Brain Team, використовує синтез кількох методів обчислювальної алгебри разом з оптимізацією обчислення нейронних мереж. Це спрощує виконання різноманітних завдань, що зазвичай потребують значного часу для моделей у інших умовах [7].

Основні можливості TensorFlow включають:

- Визначення, оптимізація та ефективне виконання математичних операцій з багатовимірними масивами (тензорами).
- Підтримка більшості методів навчання нейронних мереж, таких як регресійні функції, шарові моделі, оптимізатори, відсіювання тощо.
- Автономне виконання операцій обчислення, тобто TensorFlow вирішує, які обчислення виконуються на центральному процесорі, а які на графічному процесорі.

Ці функції роблять TensorFlow винятковою бібліотекою для розробки нейронних мереж, зосереджуючись на створенні, модифікації та навчанні обчислювальних моделей. Іншим важливим фактором є безкоштовний доступ до великої кількості протестованих моделей, підтримуваних командою TensorFlow і спільнотою, які доступні в репозиторії Github.

2.2.2 Keras

Keras – це API, який використовує функції TensorFlow (а також інших бібліотек, таких як Theano). Він був розроблений з урахуванням зручності та модульності. Keras дозволяє легко реалізовувати потужні, але складні функції TensorFlow і інтегрується з Python без значних змін або налаштувань.

2.2.3 TensorFlow Lite

TensorFlow Lite - набір ресурсів для реалізації машинного навчання на пристроях, що дозволяє розробникам виконувати моделі на мобільних, вбудованих та периферійних пристроях. Основні особливості TensorFlow Lite:

- Оптимізація для машинного навчання на пристрої за рахунок усунення ключових обмежень: затримка, конфіденційність, можливість підключення, розмір моделі та енергоспоживання.

- Сумісність із численними платформами, включаючи Android, iOS, вбудований Linux, а також мікроконтролери.

- Можливість використання різних мов програмування, таких як Java, Swift, Objective-C, C++ і Python.- Висока продуктивність завдяки апаратному прискоренню та оптимізації моделей.

- Постачання наскрізних прикладів для різних завдань машинного навчання, зокрема класифікації зображень, розпізнавання об'єктів, оцінки пози, відповідей на запитання, класифікації тексту та інших.

2.2.4 TensorFlow API Detection

Для розробки програми було використано TensorFlow Detection API. Це було найкращим варіантом через тип необхідної програми, оскільки він працює на базі TensorFlow і надає кілька прикладів і попередньо створених моделей для проєктів виявлення об'єктів.

2.2.5 Google Colaboratory

Google Colaboratory (або Google Colab) – це хмарний сервіс від Google Research, що дозволяє користувачам писати та виконувати код безпосередньо в Він сумісний з мовою програмування Python 3.6.9 і призначений для виконання завдань у сфері машинного навчання, аналізу даних та освітніх проєктів. Основні переваги Google Colab:

- Інтеграція з TensorFlow і підтримка TensorBoard без необхідності

налаштування.

- Безкоштовний доступ через обліковий запис Gmail, що не потребує встановлення Jupyter.
- Можливість збереження блокнотів у Google Drive або локально у форматі Jupyter (.ipynb).
- Використання віртуальної машини для запуску коду ізольовано від інших користувачів і ресурсів.

2.3 Виявлення об'єкту

За умови наявності зображення або відеопотоку, модель для виявлення об'єктів здатна ідентифікувати об'єкти з попередньо визначеного набору і надати дані про їхнє розташування на зображенні. На рисунку 2.3 показано приклад розпізнавання двох об'єктів та анотації їхнього положення.

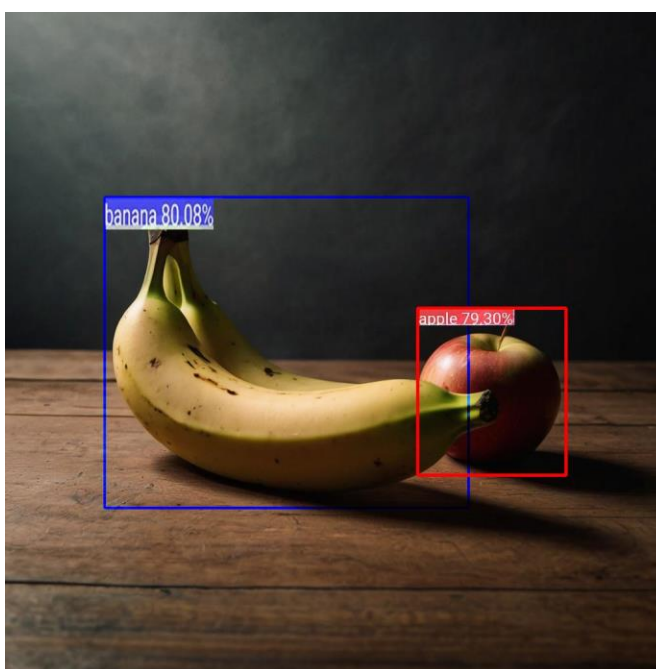


Рисунок 2.3 – Розпізнавання об'єктів

Модель виявлення об'єктів натренована ідентифікувати та визначати місцеположення різних класів об'єктів. Наприклад, таку модель можна навчити на зображеннях з різними видами фруктів, які мають позначки, що вказують на клас

кожного фрукта (наприклад, яблуко, банан або полуниця), та інформацією про їхні точні координати на зображенні.. Після навчання, коли модель отримує нове зображення, вона видає список виявлених об'єктів, координати рамок, що містять кожний об'єкт, і оцінку впевненості у правильності виявлення.

Вхідний підпис

Модель отримує зображення як вхідні дані. Наприклад, зображення повинно мати розмір 300x300 пікселів з трьома кольоровими каналами (червоний, зелений та синій) на кожен піксель. Це зображення необхідно подати моделі у вигляді згладженого буфера, що містить 270 000 байтів (300x300x3). Якщо квантована модель, кожне значення повинно бути представлено одним байтом у діапазоні від 0 до 255.

Вихідний підпис. Модель виводить чотири масиви, що відображаються на індексах 0–4. Масиви 0, 1 і 2 містять інформацію про NNN виявлених об'єктів, де кожен об'єкт представлений відповідним елементом у кожному з цих масивів.

Ця структура виводу дозволяє моделі повідомляти про кілька виявлень одночасно, надаючи координати для обмежувальних рамок, класифікаційні мітки та оцінки впевненості для кожного об'єкта. (таблиця 2.1).

Таблиця 2.1 – Розпізнавання об'єктів

| Індекс | Ім'я | Опис |
|--------|----------------|---|
| 0 | Розташування | Багатовимірний масив [N][4] значень із плаваючою комою від 0 до 1, внутрішні масиви представляють обмежувальні рамки у формі [верхній, лівий, нижній, правий] |
| 1 | Класи | Масив з N цілих чисел (виведених як значення з плаваючою комою), кожне з яких вказує на індекс мітки класу з файлу міток |
| 2 | Результати | Масив із N значень із плаваючою комою від 0 до 1, що представляє ймовірність того, що клас було виявлено |
| 3 | Число виявлень | Ціле значення N |

Наприклад, модель налаштована для розпізнавання полуниці, бананів і яблука. При поданні зображення, вона виводить визначену кількість результатів розпізнавання, як показано нижче. (таблиця 2.2)

Таблиця 2.2 – Приклад класифікації

| Клас | Оцінка | Розташування |
|----------|--------|---------------------|
| Яблуко | 0,92 | [18, 21, 57, 63] |
| Банан | 0,88 | [100, 30, 180, 150] |
| Полуниця | 0,87 | [7, 82, 89, 163] |
| Банан | 0,23 | [42, 66, 57, 83] |
| Яблуко | 0,11 | [6, 42, 31, 58] |

2.3.1 Оцінка достовірності

Для інтерпретації результатів можна звернути увагу на оцінку та розташування кожного знайденого об'єкта. Оцінка, що виражається числом від 0 до 1, також відома як індекс перетину через об'єднання (IoU), показує впевненість у правильності виявлення об'єкта. Чим ближче це число до 1, тим більша впевненість моделі.

Поріг відсікання

Цей поріг можна встановити власноруч, залежно від того, що необхідно уникати: хибні спрацьовування або хибнонегативні результати.

– Хибнопозитивні результати: області на зображення, які невірною визначаються як об'єкти, хоча ними не є.

– Хибнонегативні результати: Реальні об'єкти, які модель не розпізнала через низьку впевненість.

Як приклад, на рисунку 3.4 груша (для виявлення якої модель не була навчена) була помилково визначена як "людина". Це приклад хибнопозитивного результату, який можна уникнути, встановивши відповідний поріг відсікання. У

даному випадку значення 0,6 - 60% ефективно відкидає хибнопозитивний показник.

- Оцінка впевненості: Число від 0 до 1, яке вказує на впевненість моделі у правильності виявлення. Значення ближче до 1 означає вищу впевненість.

- Поріг відсікання: Регулюється для балансування між хибними спрацьовуваннями та хибнонегативними результатами. Підвищення порогу зменшує кількість хибнопозитивних, але може збільшити кількість хибнонегативних і навпаки.

Розташування об'єктів: Координати рамок, що містять кожний виявлений об'єкт, дозволяють точно визначити його місце розташування на зображенні.

Таким чином, налаштування порогу відсікання та аналіз оцінок дозволяють ефективно використовувати модель виявлення об'єктів для досягнення необхідної точності виявлення.

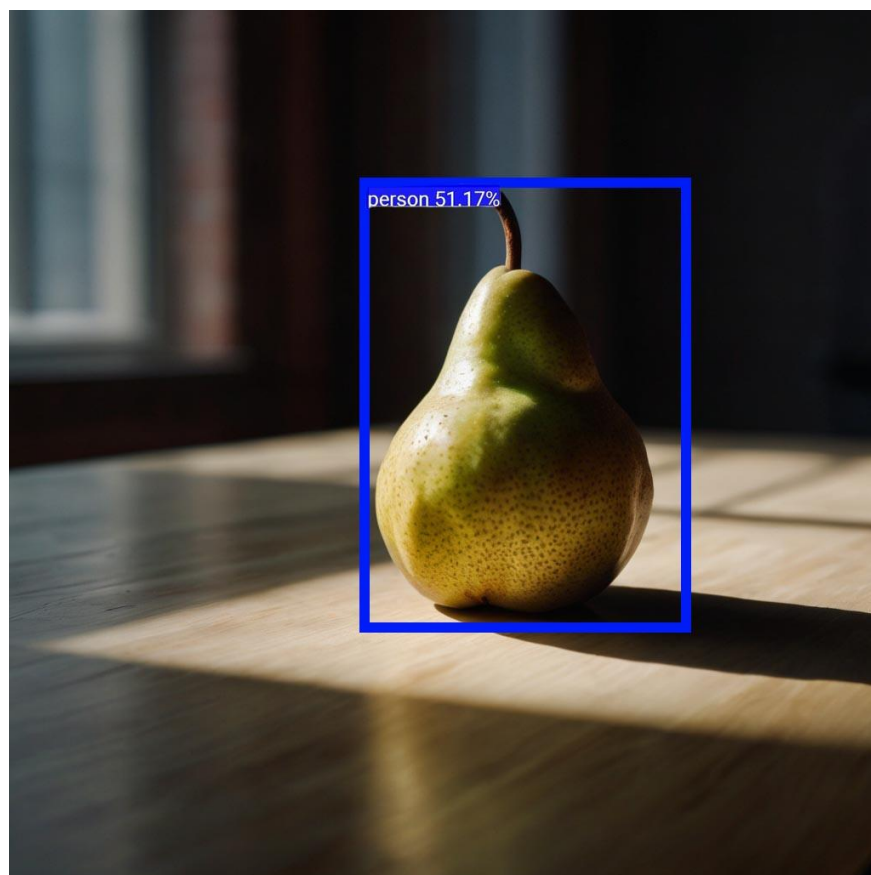


Рисунок 2.4 – Приклад неправильно ідентифікованого об'єкту

2.3.2 Розташування об'єктів на зображенні

Опис моделі розташування об'єктів Модель повертає масив з чотирьох чисел, що відображають прямокутник, який обмежує об'єкт, та його положення відносно зображення. Номери розташування в масиві представлені послідовно: зверху, зліва, знизу, справа.

Пояснення значень в масиві

- Верхнє числа: відстань від верху зображення в пікселях до верхнього краю прямокутника.
- Ліве числа: відстань від лівого краю вхідного зображення до лівого краю прямокутника
- Інші числа: нижній та правий краї прямокутника відображені аналогічним чином.

2.4 Навчання мережі

Збір та підготовка даних Для успішного розпізнавання об'єктів мережу потрібно навчити. Це передбачає збір даних та їх організацію так, щоб нейромережа могла навчатися на них. Підготовка даних - це складне мистецтво, яке включає вирішення проблем, таких як відсутні значення, пошкоджені дані, неправильний формат та мітки. Зазвичай для розпізнавання об'єктів використовують великі набори даних, такі як COCO, ImageNet, Open Images та інші, оскільки більший набір даних забезпечує більшу точність мережі.

Вибір моделі нейронної мережі Після підготовки або вибору набору даних необхідно вибрати модель нейронної мережі, яка буде використовуватися в роботі. Було обрано нейронні мережі MobileNet та EfficientDet Lite (таблиця 3.3) як найкраще рішення для поставлених задач.

Процес навчання моделі виявлення об'єктів залежить від обладнання, на якому він проходить, і може займати значний проміжок часу. Для прискорення цього процесу часто використовують графічні процесори (GPU), що істотно зменшує час навчання в порівнянні з центральними процесорами (CPU) (таблиця

2.3).

Таблиця 2.3 – Параметри нейро мереж

| Архітектура моделі | Розмір, Мб | Час кадру, ms | mAP, % |
|--------------------|------------|---------------|--------|
| EfficientDet-Lite0 | 4.4 | 37 | 25.69 |
| EfficientDet-Lite1 | 5.8 | 49 | 30.55 |
| EfficientDet-Lite2 | 7.2 | 69 | 33.97 |
| MobileNetV1 | 6.7 | 24 | 20.2 |

2.5 Донавчання мережі

Обрані моделі були навчені на наборі даних COCO 2017, тому вони можуть не зовсім відповідати специфічним вимогам. Однак, щоб уникнути перенавчання моделі з нуля і зберегти високу точність, можна скористатися бібліотекою TensorFlow Model Maker.

Ця бібліотека спрощує адаптацію моделей до нових даних за допомогою трансферного навчання, що значно зменшує необхідну кількість навчальних даних і скорочує час навчання.

Для цього потрібно виконати кілька простих кроків у Google Colab:

Встановити необхідні пакети які надані у Додатку А.

2.6 Встановлення моделі на Android смартфон

Для тесту було обрано смартфон з ОС Android. Основною причиною вибору були доступність і не висока складність встановлення.

Заради налаштування на ПК було встановлено Android Studio версії Dolphin. В Android Studio було імпортовано та запущено проєкт, вибрано файл build.gradle, внесено зміни в код, додано файл моделі .tflite до каталогу src/main/assets та виконано наступні кроки.

1. Встановлено змінну з назвою

```

    моделі: val modelName =
        "mobilenetv1.tflite"

```

2. Встановлено параметри моделі, такі як поріг передбачення та розмір набору результатів:

```

    val optionsBuilder =
        ObjectDetector.ObjectDetectorOptions.builder()
            .setScoreThreshold(threshold)
            .setMaxResults(maxResults)

```

3. Вмикаємо прискорення графічного процесора за але код коректно завершується збоєм, якщо прискорення не підтримується на пристрої:

```

    try {
        optionsBuilder.useGpu()
    } catch (e: Exception) {
        objectDetectorListener.onError("GPU is not supported on this device")
    }

```

4. Встановлено данні налаштування даного об'єкта для створення файлу TensorFlow LiteObjectDetector, що містить модель:

```

    objectDetector =
        ObjectDetector.createFromFileAndOptions
        (
            context, modelName, optionsBuilder.build()

```

Рисунок 2.5 – Кроки налаштування ПЗ

Після виконання цих кроків було розпочато синхронізацію з Gradle. У налаштуваннях телефону було активовано режим розробника, після чого телефон підключили до ПК. Після цих дій пристрій відобразився в Android Studio. Після підключення Gradle до смартфона можна тиснути на кнопку Почати. Після збірки Android Studio повідомить BUILD SUCCESSFUL на панелі Build Output.

2.7 Тестування моделі розпізнавання об'єктів

В додатку на телефон було додано 4 моделі ідифікації об'єктів:

- MobileNetV1;
- EfficientDet-Lite0;
- EfficientDet-Lite1;
- EfficientDet-Lite2.

Показники тестового запуску ефективності моделей нейронних мереж зазначені в таблиці 2.4.

Таблиця 2.4 – Результати тестувань

| Назва моделі | mAP, % | Час кадру на GPU, ms | Час кадру CPU, ms |
|----------------|--------|----------------------|-------------------|
| EfficientDet 0 | 33.6 | 146–157 | 98–108 |
| EfficientDet 1 | 54 | 230–280 | 167–180 |
| EfficientDet 2 | 67 | 420–436 | 242–260 |
| MobileNet V1 | 48 | 90–100 | 45–50 |

За результатами дослідження виявлено, що моделі нейронних мереж, які працювали на GPU, у середньому мають більшу швидкість обробки на CPU у 1,66 раза. Під час тестування були виявлені наступні проблеми:

1. Проблема дрібного зображення. На рисунку 2.5 показані вхідні розміри зображень. При зйомці з великої дистанції виникають ускладнення з розпізнаванням об'єктів, таких як автомобілі.

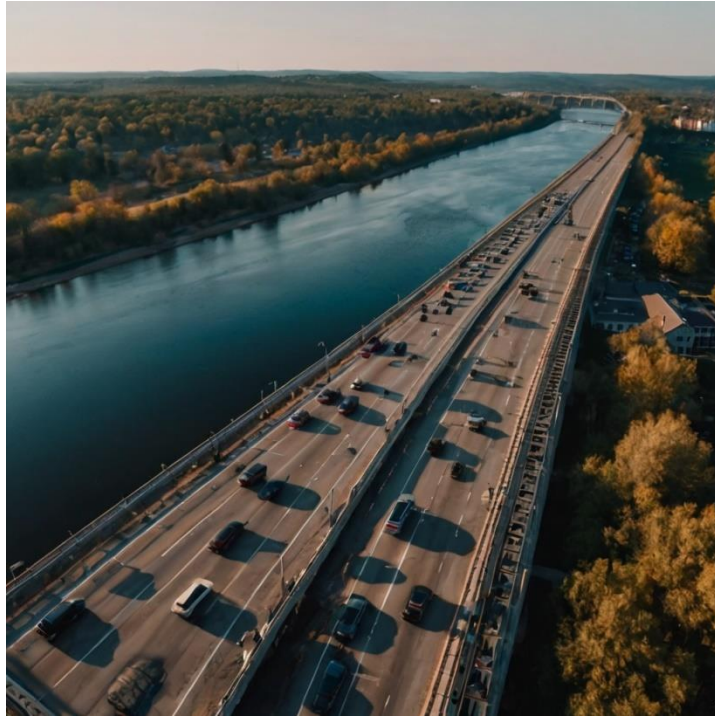


Рисунок 2.6 – Вхідне зображення

Необхідно занадто збільшити роздільну здатність вхідного зображення (рис. 2.6).

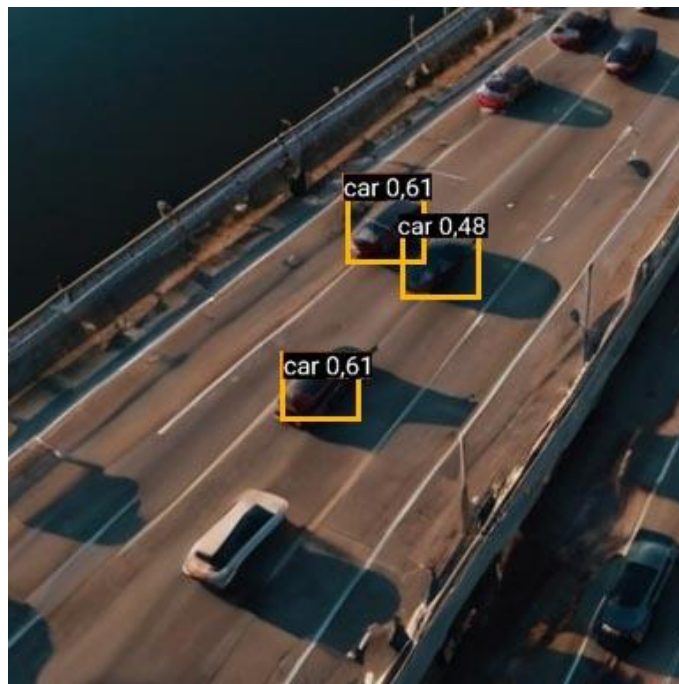


Рисунок 2.7 – Вихідне зображення

Для вирішення вищезазначених проблем можна вжити ряд заходів:

- Збільшення зображення:
- Оптичний зум: необхідна камера з відповідною оптикою, що здатна забезпечити достатнє збільшення зображення.

- Цифровий зум: необхідно використовувати камеру з високою роздільною здатністю та додатковими обчислювальними потужностями.

- Проблема нагріву: Під час тестування програми телефон перегрівається, що призводить до зниження частоти процесора та зменшення продуктивності. Для вирішення цієї проблеми в майбутніх випадках необхідно оснастити центральний процесор більш ефективною системою охолодження, ніж та, яка використовується у телефонах.

2.8 Нейромережі для генерації зображень

Система на базі нейромереж для генерації зображень — це програмне забезпечення, яке використовує нейронні мережі глибокого навчання, натреновані на великому обсязі даних. Такі мережі прогнозують значення пікселів на основі виявлених шаблонів та створюють нові зображення [14].

Існує кілька типів нейромереж для генерації зображень, зокрема:

- Передача стилю: нейромережа застосовує алгоритми глибокого навчання для перенесення стилів між зображеннями.

- Generative Adversarial Networks (GANs): ці мережі створюють нові зображення, навчаючись на існуючих даних.

- Варіаційні автокодери (VAE): також генерують нові зображення на основі навчання на наявних даних.

Deep Dream: використовується для створення сюрреалістичних зображень.

- Передача нейронного стилю: застосовується для створення реалістичних фотографій та картин.

- Генератори стабільної дифузії: створюють масив дифузійних візерунків, які змішуються для отримання кінцевого результату.

Найкраще те, що використання нейромереж для генерації зображень досить просте. Важливо, щоб обрана нейромережа відповідала вашим потребам, тому варто протестувати кілька варіантів для порівняння їхніх можливостей [17].

Переваги та недоліки нейромереж для генерації зображень.

Перш ніж обирати найкращі інструменти для генерації зображень за допомогою нейромереж, варто ознайомитися з їхніми перевагами та недоліками. Це дозволить прийняти виважене рішення щодо їх використання у вашому бізнесі та ефективніше працювати з нейромережею [8].

Основні переваги нейромереж для генерації зображень:

- Швидкість та ефективність: Нейромережі здатні швидко створювати зображення без людського втручання, що економить час і ресурси, особливо при масовому виробництві графіки.

- Економічна ефективність: Використання нейромереж дозволяє зменшити витрати на наймання професійних художників та дизайнерів, що особливо важливо для малих підприємств з обмеженим бюджетом.

- Стабільна якість: Генератори на основі штучного інтелекту забезпечують узгодженість стилю, кольору та якості, що є важливим для брендингу та маркетингових цілей.

- Джерело натхнення: Нейромережі можуть генерувати унікальні стилі та ідеї, корисні для художників і дизайнерів, які прагнуть експериментувати з новими техніками.

- Доступність: Процес створення зображень стає доступним навіть для тих, хто не має технічних або художніх навичок.

- Висока якість облич: Нейромережі відмінно справляються із завданням створення реалістичних облич, іноді важко відрізнити їх від справжніх фотографій.

Недоліки нейромереж для генерації зображень:

- Відсутність оригінальності: Зображення, згенеровані нейромережею, можуть бути позбавлені креативності, оскільки вони базуються на вже існуючих шаблонах та даних.

- Обмежений контроль: Користувачі не завжди мають повний контроль над процесом генерації, що може призвести до непередбачуваних результатів.

- Низька якість: Деякі нейромережі можуть генерувати зображення

низької якості, що може бути проблемою для професійного або комерційного використання.

– Необхідність спеціальних знань: Для ефективного використання деяких генераторів можуть знадобитися спеціальні знання та навички.

Потрібно розглянути декілька прикладів Нейромережі для генерації зображень та визначити їх особливості, позитивні та негативні сторони [15].

2.8.1 Leonardo AI

Цей безкоштовний Text-to-Image генератор на основі Stable Diffusion є чудовим варіантом для новачків або тих, кому Midjourney здався занадто складним. Він дозволяє отримати якісний результат з мінімальними зусиллями. Сервіс надає 150 "монет" на день, що дозволяє створювати до сотні зображень залежно від обраної моделі та налаштувань. Після реєстрації користувачі потрапляють на головну сторінку, де можна оцінити найкращі роботи та дізнатися, як вони створені, просто клікнувши на них.

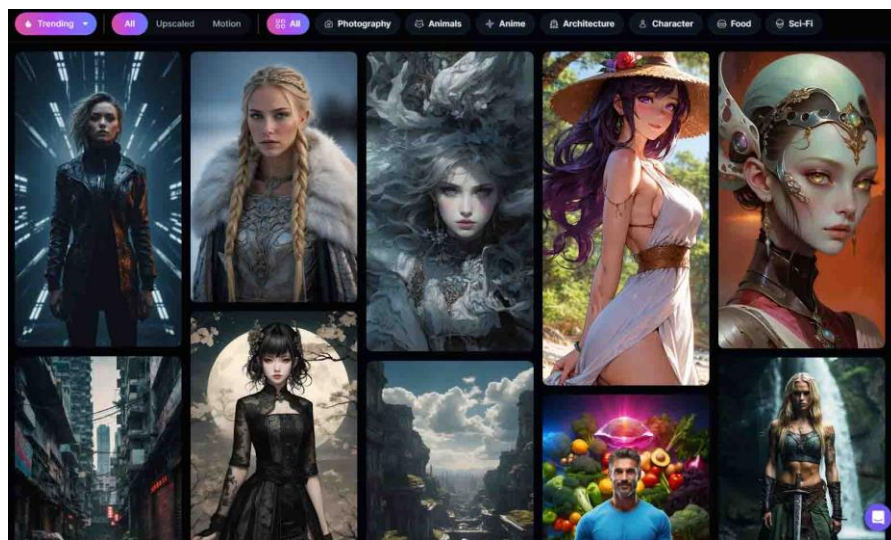


Рисунок 2.8 – Приклади робіт від інших користувачів [12]

Кожна картинку можливо відкрити (Рисунок 2.8), скачати і побачити які налаштування та промт було використано для генерування даного зображення. За допомогою таких прикладів можливо більш точно згенерувати зображення яке уявляється [16].

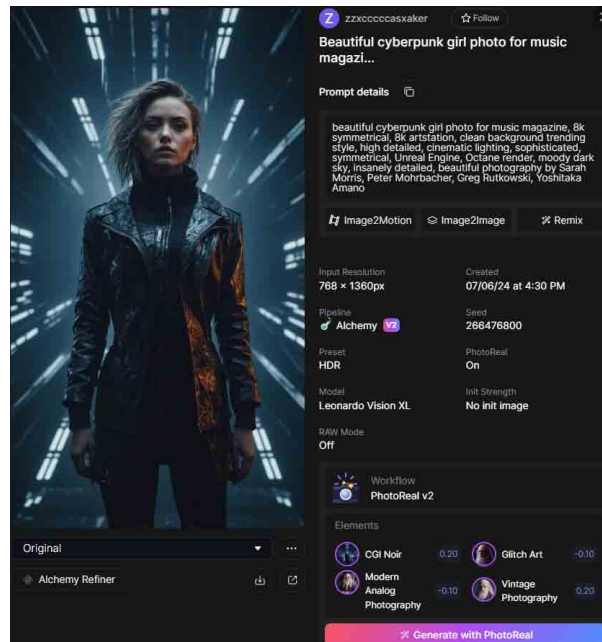


Рисунок 2.9 – Відриті налаштування одного з зображень в галереї

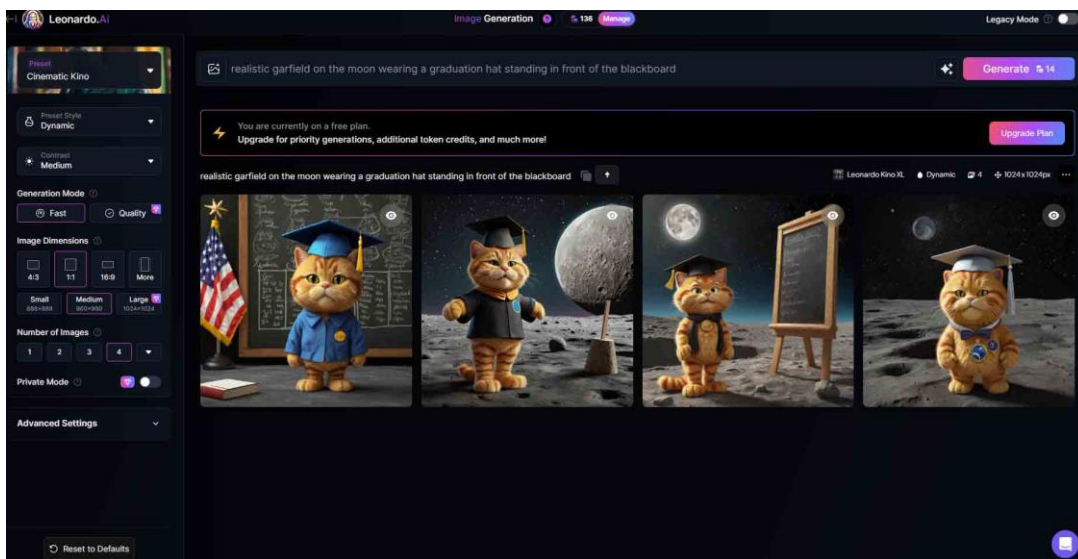


Рисунок 2.10 – Інтерфейс для генерування зображень.

Дуже багато яких функцій налаштувань, але деякі з них заблоковано при використанні безкоштовної версії. Наприклад не можна збільшувати якість зображення, неможливо використовувати приватний режим у якому зображення недоступні іншим користувачам як на (Рисунок 2.7) та інше

2.8.2 BING IMAGE CREATOR

Bing's Image Creator базується на вдосконаленій версії DALL-E, забезпечуючи аналогічну, а можливо, й вищу якість результатів з такою ж швидкістю. Як і DALL-E, цей інструмент безкоштовний для використання. Для доступу до генератора зображень достатньо відвідати веб-сайт і увійти з обліковим записом Microsoft.

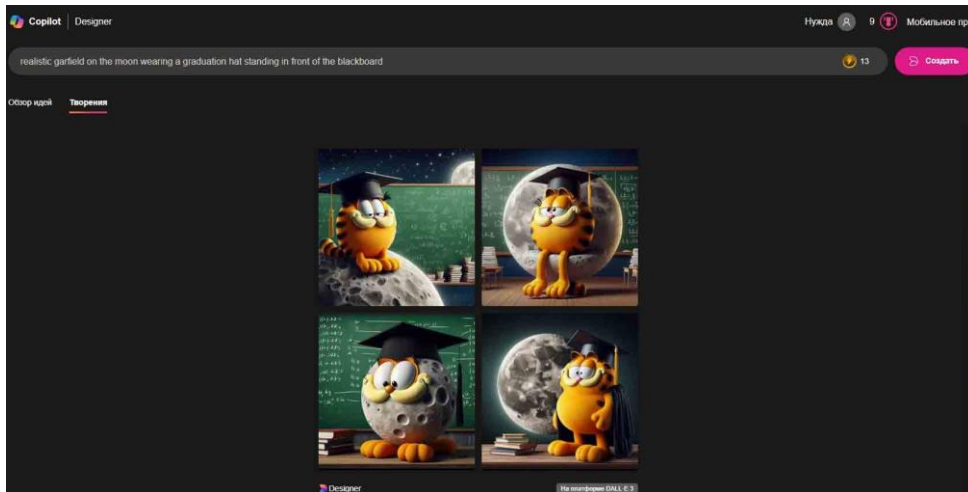


Рисунок 2.11 – Інтерфейс для генерування зображень .

Для створення зображень в інтерфейсі нема жодних налаштувань тому для більш точного генерації потрібно використовувати найбільший плюс цього ШІ-генератора це те о доступ до нього можна отримати там же, де і до чатботу Copilot.

Для доступу до генератора зображень Bing можна відвідати окремий сайт. Однак, якщо у вас є доступ до Copilot, ви можете попросити його створити зображення безпосередньо там. Для цього потрібно перевести Copilot у творчий режим і ввести будь-яку текстову підказку, яку ви хочете перетворити на зображення.

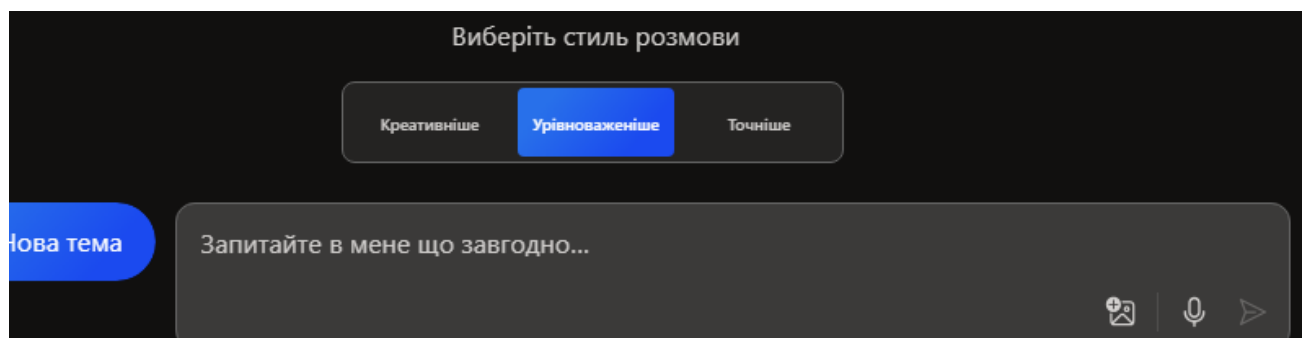


Рисунок 2.12 – Вибір режиму використання чат боту [9]

Ця функція дуже зручна, оскільки можливо задовольнити всі свої потреби у створенні картинок і ШІ-чаті в одному місці.

2.8.3 DREAM ВІД WOMBO

Мобільний додаток для генерування зображень. Завдяки своєму великому успіху, цей додаток здобув перше місце в номінації "Найкращий додаток загалом" на конкурсі Google Play 2022. Ця програма дозволяє створювати художні твори шляхом введення коротких текстових підказок.

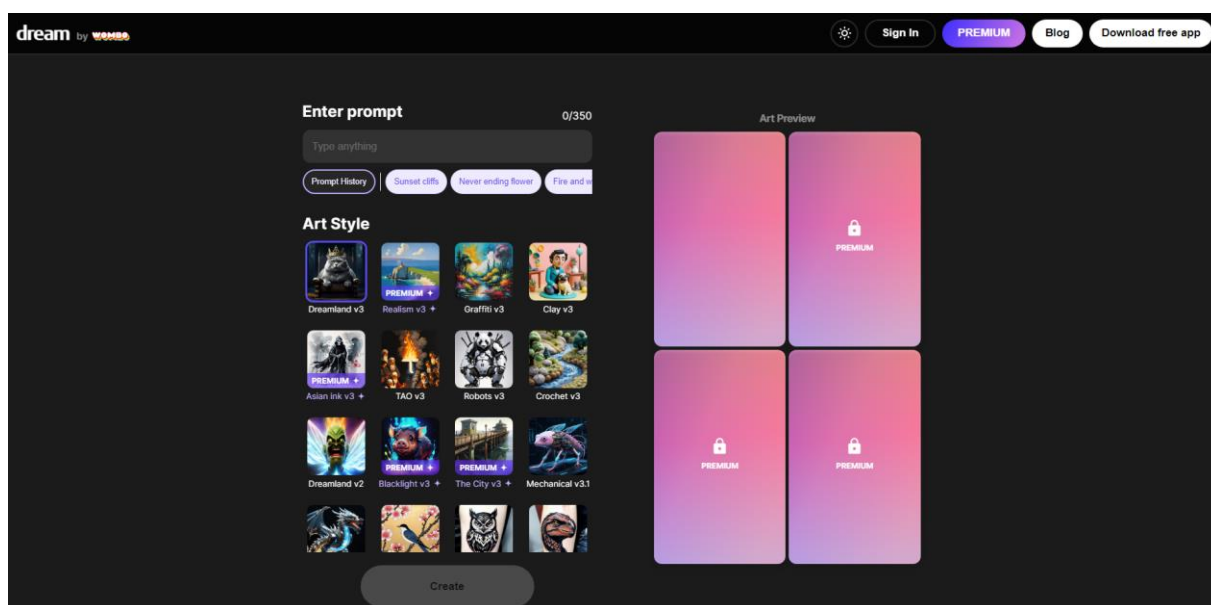


Рисунок 2.13 – Інтерфейс генерації та вибору стилів зображення [10]

Одна з переваг цього генератора зображень – можливість вибору серед різних стилів дизайну, таких як реалізм, експресіонізм, комікси, абстракція, фантазійний стиль, чорнильний малюнок тощо. Крім мобільної версії, існує безкоштовна версія для настільних комп'ютерів, яка дуже проста у використанні. Для розширеного використання програми можна оформити підписку: \$90 на рік, \$10 на місяць або придбати довічний доступ за \$170

2.8.4 CRAIYON

Хоча цей ШІ-генератор спочатку називався DALL-E mini, він не має жодного відношення до OpenAI або DALL-E 2. Насправді, це альтернативна версія з відкритим вихідним кодом. Назва DALL-E 2 mini частково виправдана, оскільки цей генератор виконує ті самі функції, що й DALL-E 2, але з дещо менш точною якістю відтворення.

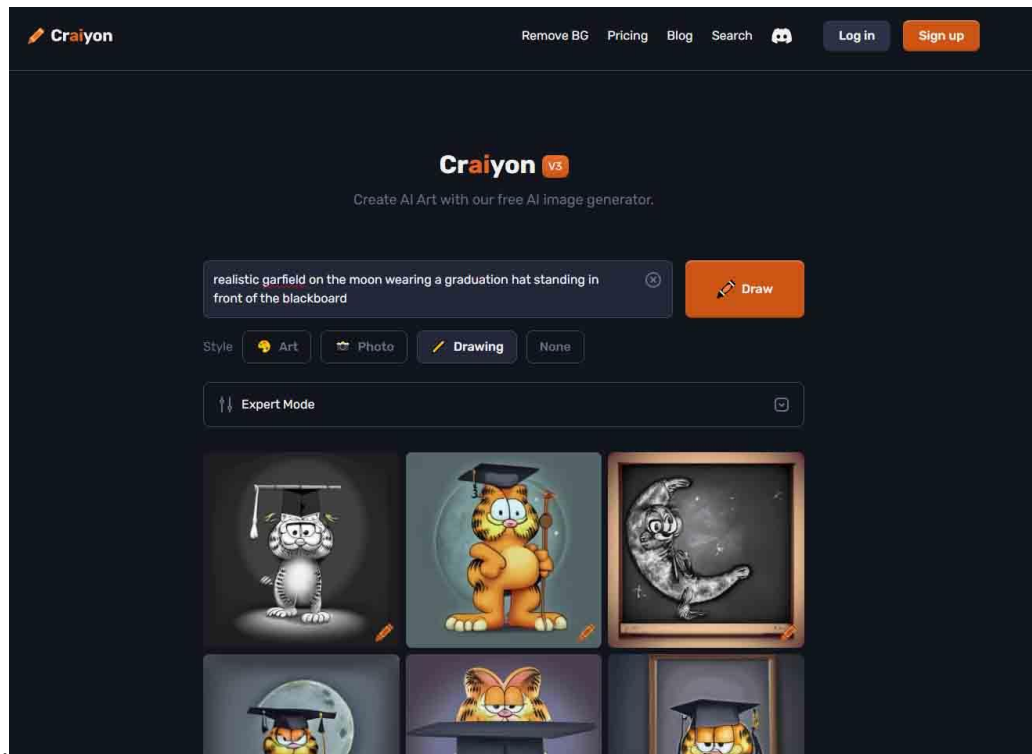


Рисунок 2.14 – Інтерфейс генерації та вибору стилів зображення Craiyon

[11]

На відміну від DALL-E 2, результати, що створюються Craiyon, не мають такої високої якості. Проте, завдяки необмеженій кількості підказок, можна продовжувати експериментувати з налаштуваннями, доки не буде досягнуто бажаного результату. Крім того, сайт надзвичайно простий у використанні, що робить цей ШІ-генератор привабливим і конкурентоспроможним сервісом.

2.8.5 Ideogram

Нейромережа для генерації картинок від колишніх співробітників Google. У

плані функціоналу вона майже нічим не поступається MidJourney чи DALL-E, представляючи широкий спектр можливих стилів і жанрів.

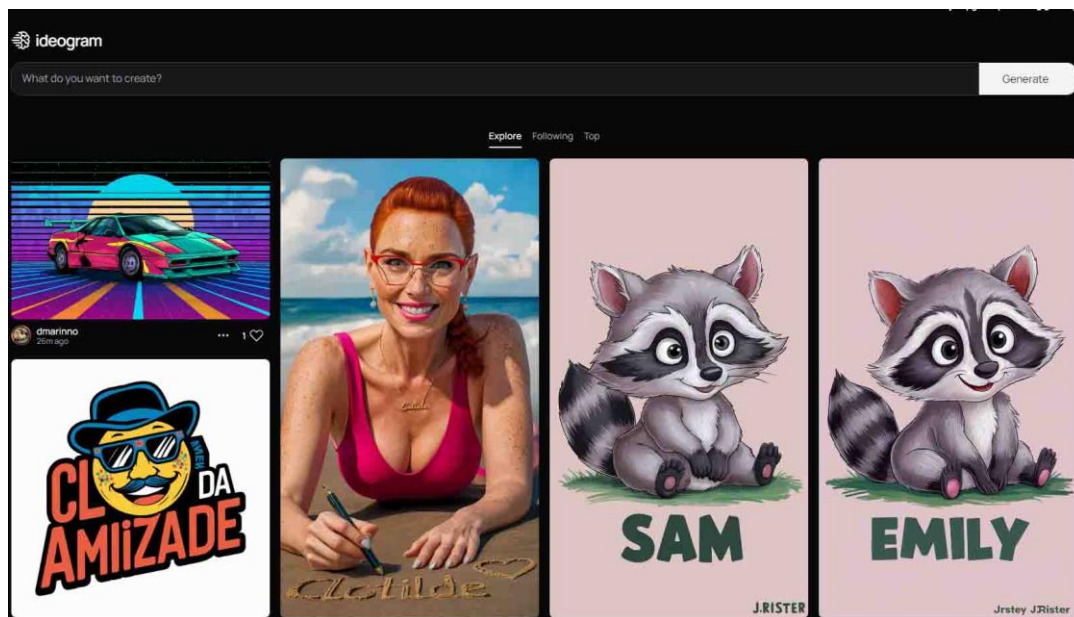


Рисунок 2.15 – Інтерфейс Ideogram генерації та зображення інших авторів [13]

Користувачі також можуть ставити "лайки" чужим роботам, брати їх за основу і створювати на їхній базі власні ремікси або варіації, тобто нейромережа перетворюється на свого роду соціальну мережу. Користувачі зазначають, що це один із найперспективніших інструментів нейрогенерації контенту на сьогоднішній день.

2.9 Розробка додатку генерації зображень за текстовим запитом за допомогою ШІ

OpenAI надав доступ до DALL E через свій API, що означає, що стало можливим використовувати його функціональні можливості у своїх додатках на Python.

У цій роботі знадобиться:

- Використовувати бібліотеку OpenAI Python
- Вивчить дзвінки API, пов'язані зі створенням зображень
- Створить зображення з текстових підказок

- Створить кілька варіантів згенерованого зображення
- Перетворювати відповіді Base64 JSON на файли зображень PNG

Для цього знадобиться невеликий досвід роботи з Python, JSON та операції з файлами.

Потрібно встановити бібліотеку OpenAI Python. Для коректної роботи потрібен Python версії 3.7.1 або вище, створіть та активуйте віртуальне середовище та встановіть бібліотеку OpenAI

```
Python: PS> python --
versionPython 3.12.0
PS> python -m venv venv
PS> .\venv\Scripts\activate
(venv) PS> python -m pip install openai
```

Пакет openai дає доступ до повної версії API OpenAI. Зосередитимось на класі Image, який можна використовувати для взаємодії з DALL E під час створення та редагування зображень з текстових підказок.

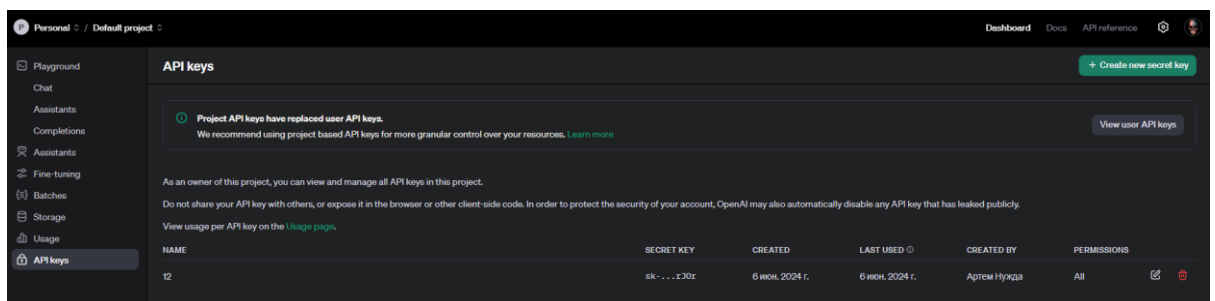


Рисунок 2.16 – API від OpenAI.

Потрібний ключ API для здійснення успішних дзвінків.

Зареєструвавшись в OpenAI API потрібно створити новий ключ API, клацнувши на меню, що розкривається, у своєму профілі і вибравши View API keys:

На цій сторінці можливо керувати своїми ключами API, які дозволяють отримати доступ до послуги, яку OpenAI пропонує через їх API. Можно створювати та видаляти секретні ключі.

Натисніть **Create new secret key**, щоб створити новий ключ API, і скопіюйте значення, показане у спливаючому вікні:

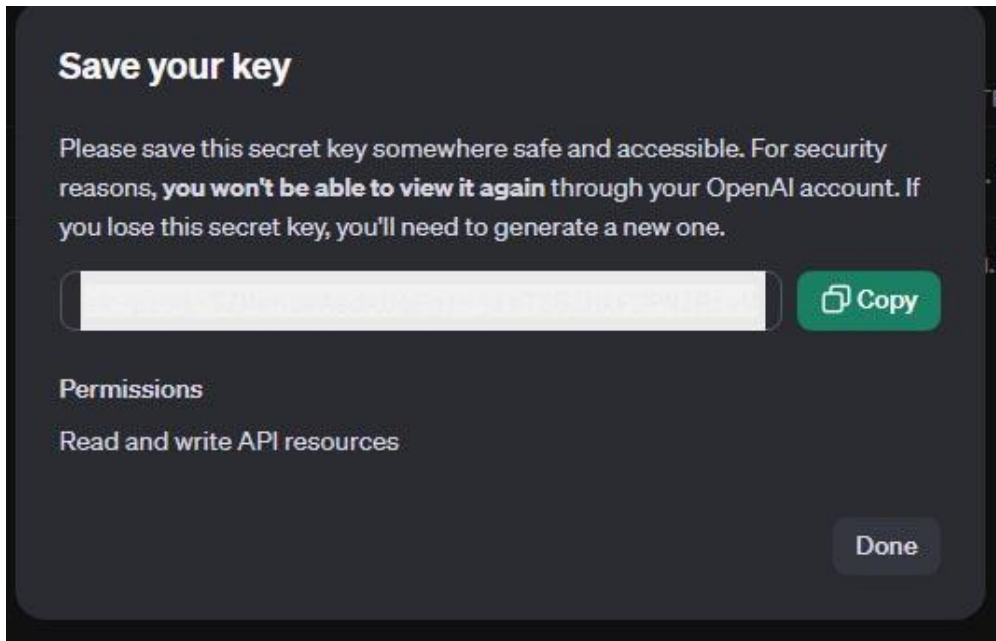


Рисунок 2.17 – Отримання ключа для API

Скопіюємо значення цього ключа, щоб пізніше використати його у проєкті. Побачити значення ключа можна лише один раз.

Збережемо ключ API як змінне середовище

Швидкий спосіб зберегти ключ API і зробити його доступним для скриптів Python - зберегти його як змінне середовище. Виконаємо таку команду:

```
(venv) PS> $ENV:OPENAI_API_KEY = "<your-key-value-here>"
```

За допомогою цієї команди робимо ключ API доступним у змінному середовищі `OPENAI_API_KEY` у поточному сеансі терміналу. Маємо на увазі, що втиратимемо його, якщо закриємо термінал.

Можемо назвати змінну як завгодно, але якщо використовувати ім'я `OPENAI_API_KEY`, запропоноване в документації OpenAI, зможемо

використовувати надані приклади коду без необхідності виконувати будь-яке додаткове налаштування.

Ознайоммося з цінами на DALL·E та інші продукти OpenAI API

Компанія розраховує вартість запитів до API зображень для кожного зображення залежно від якості вихідного зображення:

Дозвіл Ціна за зображення
 256×256 0,016 долар США
 512×512 0,018 долара США
 1024×1024 0,020 долара США

Створемо зображення з текстової підказки за допомогою DALL E від OpenAI

Перевіримо, що все налаштовано та коректно працює, використовуючи бібліотеку openai через інтерфейс командного рядка:

```
(venv) $ openai api image.create -p "a vaporwave computer"
```

Ця команда надішле запит через API і створить одне зображення з текстової підказки комп'ютер вейпорвейв. В результаті отримаємо відповідь JSON, що містить URL-адресу, що вказує на щойно створене зображення:

```
{"created": 1668073562,  
  "data": [{"url": "https://oaidalleapiprodscus.blob.core.windows.net/private/org-QANMxYn3BsMeuAbRT8X3iiu3/user-xSuQTJ0IIVj3dHM4DPymXTg4/img-5GqtVxL86Retwi282RbE8HzA.png?st=2022-11-10T08%3A46%3A02Z&se=2022-11-10T10%3A46%3A02Z&sp=r&sv=2021-08-06&sr=b&rscd=inline&rsct=image/png&skoid=6aaadede-4fb3-4698-a8f6-684d7786b067&sktid=a48cca56-e6da-484e-a814-9c849652bcb3&skt=2022-11-09T14%3A20%3A19Z&ske=2022-11-10T14%3A20%3A19Z&sks=b&skv=2021-08-06&sig=yorbHuIy/qHhWvGPmJrZ8apJptorzpI 0/62VH2lmhcg%3D"} ] }
```

Натискаємо на URL-адресу або скопіюємо та вставимо її у браузер, щоб переглянути зображення. Ось зображення, яке DALL·E придумав на цей запит (Рисунок 2.18).

Примітка. URL-адреса із створеним вами зображенням дійсна лише протягом однієї години, тому обов'язково збережемо зображення на свій комп'ютер, якщо воно знадобиться.



Рисунок 2.18 – Згенероване зображення

Тепер, коли переконалися, що все налаштовано правильно, і отримали уявлення про те, що можна робити з OpenAI API, ми перейдемо до того, як інтегрувати його у скрипт Python.

```
# create.py
import os
import openai
PROMPT = "An eco-friendly computer from the 90s in the style of vaporwave"
openai.api_key = os.getenv("OPENAI_API_KEY")
response = openai.Image.create(
    prompt=PROMPT,
    n=1, size="256x256",
```

```

)print(response["data"][0]["url"])
import os
import openai
PROMPT = "An eco-friendly computer from the 90s in the style of vaporwave"
openai.api_key = os.getenv("OPENAI_API_KEY")
response = openai.Image.create(
    prompt=PROMPT,
    n=1,
    size="256x256",
)print(response["data"][0]["url"])

```

Як і раніше, цей код надсилає автентифікований запит через API, який генерує одне зображення на основі тексту PROMPT. Зауважимо, що цей код додає деякі налаштування, які допоможуть розширити функціональні можливості скрипту:

Рядок 7 визначає текстову підказку як константу.

Для отримання більш точних результатів, додаємо більше тексту для опису зображення, яке хочемо отримати. Крім того, додавання тексту до константи у верхній частині скрипта дозволяє швидко реорганізувати код замість того, щоб збирати текст з введення користувача, оскільки так значення швидше знайти і відредагувати.

Рядок 9 отримує ключ API зі змінного середовища, в якому його зберегли раніше. Оскільки назвали змінну оточення OPENAI_API_KEY, навіть не потрібен цей рядок коду. Бібліотека openai автоматично отримує доступ до значення ключа API з середовища, доки дотримуємось запропонованого імені.

За допомогою цього рядка коду також можемо завантажити його зі змінного середовища з іншим ім'ям.

Рядок 11 створює екземпляр openai.Image і викликає .create(). Наступні кілька рядків містять деякі параметри, які можна передати методу.

Рядок 12 передає значення параметру PROMPT з відповідним ім'ям prompt.

При цьому передається текст DALL E, який буде використаний для створення зображення. Звернемо увагу, що також був надісланий текстовий запит, коли викликали API з інтерфейсу командного рядка.

Рядок 13 — це параметр, який не використовувався раніше. Він передає ціле число 1 параметр n. Цей параметр дозволяє визначити, скільки нових зображень створить Dall-E. Значення n повинно бути від одного до десяти і за умовчанням 1.

У рядку 14 показано ще один новий параметр, який не використовувався під час виклику API із CLI. За допомогою size можна визначити розміри зображення, яке має генерувати DALL·E. Аргумент повинен бути рядком - або "256x256", або "512x512", або "1024x1024". Кожен рядок представляє розміри пікселів зображення, яке отримується. За замовчуванням встановлено максимальну роздільну здатність – 1024x1024.

Також отримати URL-адресу для перегляду згенерованого зображення в Інтернеті.

Для цього переглядаємо відповідь JSON на ключ url у рядку 17 та друкуєте його значення у своєму терміналі.

Коли запускається цей скрипт, отримуємо висновок, аналогічний до попереднього, але тепер не відобразиться всю відповідь JSON, а тільки URL-адресу:

```
(venv) $ python create.py
https://oaidalleapiprodscus.blob.core.windows.net/private/org-
QANMxYn3BsMeuAbRT8X3iiu3/user-xSuQTJ0IIVj3dHM4DPymXTg4/img-
4AMS4wJLFSu6ClQmGDppAeV□.png?st=2022-11-
10T12%3A22%3A46Z&se=2022-11-10T14%3A22%3A46Z&sp=r&sv=2021-08-
06&sr=b&rscd=inline&rsct=image/png&skoid=6aaadede-4fb3-4698-a8f6-
684d7786b067&sktid=a48cca56-e6da-484e-a814-9c849652bcb3&skt=2022-11-
10T10%3A55%3A29Z&ske=2022-11-11T10%3A55%3A29Z&sks=b&skv=2021-08-
06&sig=xJWimMiA1/nGmFMYKUTsJq7G1u4xSL652r/MrzTH0Nk%3D
```

Клацнувши на посилання або вставши його в адресний рядок браузера, щоб переглянути згенероване зображення. Ваше зображення знову виглядатиме інакше, але повинно з'явитися зображення, схоже на текстовий опис, який використовувався в PROMPT



Рисунок 2.19 – Згенероване зображення іншого промту

Надалі можна покращувати код але це повною мірою показує функціонал віл OpenAI.

Висновки до розділу

У ході дослідження та реалізації проекту з розпізнавання та генерації графіки на основі штучного інтелекту було досягнуто результатів, які демонструють практичну цінність і ефективність сучасних AI-технологій у цій сфері.

1. Автоматизація та ефективність: Застосування неймереж дозволило значно прискорити процес створення та обробки зображень. Автоматизація цих

процесів не лише зменшила витрати часу, але й підвищила загальну продуктивність.

2. Економічна доцільність: Використання штучного інтелекту зменшило потребу в залученні великої кількості висококваліфікованих фахівців. Це особливо важливо для малих та середніх підприємств, які прагнуть знизити витрати без втрати якості.

3. Узгодженість та якість: Результати експериментів показали, що генератори графіки на основі нейромереж здатні створювати зображення високої якості, які відповідають заданим стилям та стандартам. Це важливо для забезпечення візуальної узгодженості у брендovаних матеріалах та маркетингових кампаніях.

4. Творчі можливості: Нейромережі продемонстрували здатність генерувати унікальні ідеї та стилі, які можуть стати джерелом натхнення для дизайнерів та художників. Це відкриває нові горизонти для творчих експериментів і пошуку інноваційних рішень.

Однак, під час реалізації проекту було виявлено й кілька обмежень, які слід враховувати при подальшому використанні цих технологій:

1. Контроль над процесом: Деякі інструменти надають обмежені можливості для точного налаштування параметрів генерації, що може призводити до непередбачуваних результатів.

2. Оригінальність та якість: Незважаючи на високий рівень автоматизації, іноді згенерованим зображенням може бракувати оригінальності, оскільки вони базуються на вже існуючих даних і шаблонах.

Загалом, практичні результати проекту підтверджують, що технології розпізнавання та генерації графіки на основі штучного інтелекту мають значний потенціал для покращення продуктивності та якості в різних галузях. Вони дозволяють ефективно вирішувати завдання, пов'язані з обробкою та створенням графічного контенту, забезпечуючи при цьому нові можливості для творчості та інновацій.

ВИСНОВКИ

Технології розпізнавання та генерації графіки на основі штучного інтелекту значно змінюють підходи до створення та обробки зображень у різних галузях. Вони забезпечують швидке та ефективне створення графічного контенту, знижуючи витрати та підвищуючи якість продукції. Завдяки здатності генерувати унікальні стилі та забезпечувати узгодженість візуальних елементів, ці технології стають цінним інструментом для дизайнерів, художників, маркетологів та багатьох інших фахівців.

Попри значні переваги, такі як економічна ефективність, джерело творчого натхнення та доступність для користувачів без спеціальних навичок, існують і певні обмеження. Серед них – можливість генерації зображень низької якості, відсутність повного контролю над процесом та потенційна відсутність оригінальності.

Таким чином, інтеграція штучного інтелекту у процеси розпізнавання та генерації графіки пропонує широкі можливості, але потребує уважного підходу до вибору інструментів та їх налаштування для досягнення бажаних результатів. Зважене використання цих технологій дозволить максимально ефективно реалізувати їх потенціал і уникнути можливих недоліків.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Всі стривожені: як використання ШІ впливає на мистецтво, кіно, музику та літературу URL: <https://suspilne.media/culture/645624-vsi-strivozeni-ak-vikoristanna-si-vplivae-na-mistectvo-kino-muziku-ta-literaturu/> (дата звернення: 05.06.2024)
2. Моркун Н. В., Маринич І. А. Методичні вказівки до виконання кваліфікаційної роботи бакалавру для студентів спеціальності 151 “Автоматизація та комп’ютерно-інтегровані технології”. Кривий Ріг : Видавничий центр КНУ, 2019. 50 с.
3. ДСТУ 3008:2015. Звіти у сфері науки і техніки. Структура і правила оформлення. Київ, ДП «УкрННЦ», 2015. 26с. (Інформація та документація).
4. ДСТУ 8302:2015. Бібліографічне посилання. Загальні вимоги та правила складання Київ, ДП «УкрННЦ», 2016. 16 с. (Інформація та документація).
5. ДСТУ 3582:2013. Бібліографічний опис. Скорочення слів і словосполучень в українській мові. Загальні вимоги та правила. Київ, ДП «УкрННЦ», 2013. 23 с. (Інформація та документація)
6. ДСТУ 3651.0-97 Метрологія. Одиниці фізичних величин. Основні одиниці фізичних величин Міжнародної системи одиниць. Основні положення, назви та позначення Київ, Держстандарт України, 1998. 27 с. (Інформація та документація).
7. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. URL: <https://doi.org/10.48550/arXiv.1704.04861> (дата звернення: 05.06.2024)
8. Нейромережі для генерації фото: 5 безкоштовних сервісів, які підійдуть новачкам. URL: <https://www.unian.ua/techno/yaka-neyromerezha-bezkoshtovno-generuye-kartinki-5-naykrashchih-servisiv-12629253.html> (дата звернення: 05.06.2024)
9. Microsoft Copilot. URL: <https://copilot.microsoft.com> (дата звернення:

05.06.2024)

10. Dream.ai. URL: <https://dream.ai> (дата звернення: 05.06.2024)

11. Craiyon. URL: <https://www.craiyon.com> (дата звернення: 05.06.2024)

12. Leonardo.AI. URL: <https://app.leonardo.ai> (дата звернення: 05.06.2024)

13. Ideogram. AI URL: <https://ideogram.ai/t/explore> (дата звернення: 05.06.2024)

14. УНІАН ІНФОРМАЦІЙНЕ АГЕНТСТВО. URL: <https://www.unian.ua/techno/yaka-neyromerezha-bezkoshtovno-generuye-kartinki-5-naykrashchih-servisiv-12629253.html> (дата звернення: 05.06.2024)

15. Найкращі нейромережі для генерації зображень — майбутнє вже поруч. :URL: <https://is.gd/SauNB8> (дата звернення: 05.06.2024)

16. 5найкращих нейромереж для генерації зображень 2023 року. URL: <https://www.unian.ua/techno/yaka-neyromerezha-bezkoshtovno-generuye-kartinki-5-naykrashchih-servisiv-12629253.html> (дата звернення: 05.06.2024)

17. IT Рейтинг України. URL: <https://it-rating.ua/5-naykraschih-neyromerej-dlya-generatsii-zobrajen-2023-roku> (дата звернення: 05.06.2024)

18. GarageBand for PC Windows 10, 11 Free Download. URL: <https://garagebandonpc.com>(дата звернення: 05.06.2024)

19. Видавництво Ранок Хочу на Марс! URL: <https://www.ranok.com.ua/ru/info-khochu-na-mars-38191.html> (дата звернення: 05.06.2024)

20. Pro 100 Media Краматорська На Amazon з'явилися книжки про “містичну смерть” Вікторії Амеліної, написані, ймовірно, штучним інтелектом, яка загинула під час обстрілу Краматорська. URL: <https://pro100media.com.ua/novyny-kramatorska/na-amazon-z-yavylysyu-knyzhky-pro-mistychnu-smert-viktoriyi-amelinoyi-napysani-jmovirno-shtuchnym-intelektom-yaka-zagynula-pid-chas-obstrilu-kramatorska/> (дата звернення: 05.06.2024)

Лістинг кода

```

!sudo apt -y install libportaudio2
!pip install -q --use-deprecated=legacy-resolver tflite-model-maker
!pip install -q pycocotools
!pip install -q opencv-python-headless==4.1.2.30
!pip uninstall -y tensorflow && pip install -q tensorflow==2.8.0
    – імпортувати необхідні
пакети;import numpy as np
import os
from tflite_model_maker.config import
QuantizationConfig from tflite_model_maker.config
import ExportFormat
from tflite_model_maker import
model_spec from tflite_model_maker
import object_detector import tensorflow as
tf
assert tf.__version
_____.startswith('2')
tf.get_logger().setLevel('ERROR')
from absl import logging
logging.set_verbosity(logging.ERR
OR)
!sudo apt -y install libportaudio2
!pip install -q --use-deprecated=legacy-resolver tflite-model-maker
!pip install -q pycocotools
!pip install -q opencv-python-headless==4.1.2.30

```

```
!pip uninstall -y tensorflow && pip install -q tensorflow==2.8.0
    – імпортувати необхідні
```

```
пакети;import numpy as np
```

```
import os
```

```
from tf_lite_model_maker.config import
```

```
QuantizationConfig from tf_lite_model_maker.config
```

```
import ExportFormat
```

```
from tf_lite_model_maker import
```

```
model_spec from tf_lite_model_maker
```

```
import object_detector import
```

```
tensorflow as tf
```

```
assert tf.__version__.startswith('2')
```

```
tf.get_logger().setLevel('ERROR') from
```

```
absl import logging
```

```
_ logging.set_verbosity(logging.ERROR)
```

```
– підготувати набір даних, який надається у форматі CSV;
```

```
– вибрати архітектуру моделі виявлення
```

```
об'єктів;spec = model_spec.get('efficientdet_lite1')
```

```
– завантажити датасет;
```

```
train_data, validation_data, test_data =
```

```
object_detector.DataLoader.from_csv('gs://cloud-ml-
data/img/openimage/csv/drone_ml_use.csv')
```

```
– проНавчання TensorFlow моделі з навчальними даними;
```

```
model = object_detector.create(train_data,
```

```
model_spec=spec, batch_size=8,
```

```
train_whole_model=True, validation_data=validation_data)
```

```
– оцінка моделі з тестовими
```

```
даними;model.evaluate(test_data)
```

```
– експорт моделі як TensorFlow Lite.
```