

ПЕРЕНЕСЕННЯ СТИЛЮ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

В.І. МАЦ¹, Л.М. ЛЮБЧИК²

¹*магістрант кафедри "Комп'ютерної математики та аналізу даних", НТУ «ХПІ», Харків, УКРАЇНА.*

²*доцент кафедри, канд. техн. наук, НТУ «ХПІ», Харків, УКРАЇНА*

**email: vladislavmats8@gmail.com.*

За останні 10 років відбувся великий прорив в областях машинного навчання, штучних нейронних мереж, і, як наслідок, у комп'ютерному зорі. Можливо, найкращим прикладом цього буде проект "Asirra" від Microsoft Research. Проект був заснований у 2006 році як засіб для відрізнення ботів від людей в інтернеті. За його задумкою користувачу ставилась задача відповісти: кіт чи собака знаходиться на фотографії. Зображення брались з постійно зростаючої бази, що збиралася в притулках для тварин. Ця система безпеки вважалася майже ідеальною, бо задача відрізнення kota від собаки вважалася AI-повною, тобто для її вирішення необхідно створення повноцінного штучного інтелекту. У той час кращі алгоритми мали якість близько 60%, тоді як випадковий вибір має якість 50%. Тому ймовірність, що алгоритм зможе вірно підписати 15 зображень була мізерною (близько 0.047%), в той час як для людини ця задача не складає жодних проблем. Результатом проекту стало його закриття у 2014 році за причиною вирішення поставленої задачі. Найкращі алгоритми на основі глибоких штучних нейронних мереж мали якість більше 99%, і, як наслідок, 15 зображень відповідали вірно з ймовірністю 86%, а 50 - з ймовірністю 60.5%[2]. Кількість помилок у цій надзвичайно складній для комп'ютера задачі зменшилась у 40 разів та стала абсолютно мізерною. І все це лише за 8 років. Звичайно, задача відрізнення котів від собак є далеко не найважливішим здобутком глибинного навчання та комп'ютерного зору. Є безліч більш важливих та корисних для людства задач: детекція об'єктів на зображеннях, розпізнавання облич, пішоходів та дорожніх знаків, а також багато іншого. Але це завдання є дуже важливим з теоретичної точки зору - воно вказало на фундаментальні проблеми та обмеження комп'ютерного зору, вирішення яких дуже допомогло вирішенню й інших задач.

У наш час більшість задач комп'ютерного зору, пов'язаних з розумінням та класифікацією того, що відбувається на зображеннях, мають гарний розв'язок, але з певними недоліками. У деяких схожих задачах комп'ютер навіть перевершив людину. Може створитися враження, що комп'ютерний зір буде повністю вирішеною областю в найближчі роки. Але це не так. Людина вмє набагато більше, аніж просто розуміти, що відбувається на зображенні - вона вмє створювати нові зображення. Тому одна з найбільш швидко розвиваючихся областей у комп'ютерному зорі останніх років - генерація зображень. І вона виявляється більш складною та відкриває відповіді на такі

питання: як людина створює нові образи та як навчити комп'ютер робити те ж саме. І також, що не менш важливо, ця область досі є найменш вивченою, і вона має чимало відкритих питань та задач. Перші значущі успіхи в ній з'явилися лише в останні 5 років, але навіть вони відкривають величезну кількість перспектив та технологій, що сильно допоможуть людству у майбутньому.

Як тему для даної роботи я обрав завдання з області генерації зображень - задачу перенесення стилів. Під нею зазвичай мають на увазі наступне завдання: ми маємо 2 зображення: семантичне (зазвичай фотографія) та стилістичне (зазвичай картина художника), і наша задача полягає в тому, щоб згенерувати третє зображення, таке, щоб його семантичний зміст був якомога ближче до семантичного зображення, а стилістичне наповнення до стилістичного.

Основними цілями роботи були:

1. Реалізувати алгоритм перенесення стилю.
2. Провести експерименти з різними нейронними мережами та обрати найкращу архітектуру для використання її в нашому алгоритмі.
3. Провести експерименти зі взяттям семантичного змісту зображення з різних шарів нейронної мережі, порівняти їх якість роботи та обрати найкращий для використання у нашому алгоритмі.
4. Провести експерименти зі взяттям стилю зображення з різних шарів нейронної мережі, порівняти їх якість роботи та обрати найкращий для використання у нашому алгоритмі.
5. Провести експерименти зі знищенням локальності стилістичного тензору різними способами (взяття середнього фільтру та матриця Грама), порівняти їх якість роботи та обрати найкращий для використання у нашому алгоритмі.
6. Провести експерименти з різними гіперпараметрами (число ітерацій, параметр важливості втрат), а також порівняти якість та швидкість роботи алгоритму з їх різними комбінаціями.
7. Запустити фінальний алгоритм на різних парах зображень та візуалізувати результат його роботи.

У результаті роботи був реалізований та досліджений алгоритм перенесення стилю за допомогою штучних нейронних мереж. Для цього були опрацьовані та порівняні історичні методи перенесення стилю зі способами, що ґрунтуються на глибоких нейронних мережах. Також були проведені експерименти для визначення впливу різних параметрів алгоритму, таких як архітектура нейронної мережі, способи виділення та перенесення семантичного змісту та стилю зображення, кількості ітерацій та параметрів важливості втрат на якість та швидкість роботи алгоритму. Як результат були визначені найкращі параметри алгоритму та представлені візуалізації його роботи на зображеннях з різними семантичними змістами та стилями.

Список літератури:

1. "A Neural algorithm of artistic style". - 2015 [Електронний ресурс]. Режим доступу: <https://arxiv.org/pdf/1508.06576>
2. Ian Goodfellow, Yoshua Bengio, Aaron Courville. "Deep learning". – 2016