

УДК 004.9

Миронова Н.О.¹, Сергієнко О. С.²

¹канд. техн. наук, доц. ЗНТУ

²студ. гр. КНТ-415 ЗНТУ

ПРОГРАМНА РЕАЛІЗАЦІЯ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ ДЛЯ ОБРОБКИ ЗОБРАЖЕНЬ

На сьогодні, обробка медіа даних, що представляють собою статичні зображення, відео чи навіть онлайн відео трансляції у реальному часі, є дуже актуальною темою. Обсяги медіа ростуть, необхідність у їхній обробці також. Багато хто з нас використовує засоби обробки зображень та відео у своєму повсякденному житті, ведучи блог у Instagram, Twitter, Facebook чи навіть YouTube, де простота, швидкість та автоматичність процесу обробки медіа робить його зручним та легким але обмеженим.

На допомогу прийшло спеціалізоване професійне програмне забезпечення, як Adobe Photoshop, Sony Vegas Pro та інше ПЗ для обробки відео та зображень. Але використання цих програм потребує певних навичок та має під собою ручну обробку та редагування, що підвищує складність використання та витрачає багато часу.

Засобом автоматизації певних рутинних ручних дій користувача для обробки зображень і не тільки можуть виступати алгоритми машинного навчання, а саме – нейронні мережі та лінійні алгоритми. За їх допомогою можна автоматизувати такі процеси обробки зображень як комплексне покращення якості, корекція кольору, накладання стилів та їх трансформування, масок та багато іншого лише в декілька простих кроків, без необхідності ручного редагування медіа даних.

Розроблене програмне забезпечення має автоматизувати певний процес обробки медіа, а саме – глибоке накладання стилю на окреме зображення чи відео, в декілька кроків, без участі користувача. Таке ПЗ дає змогу перетворити звичайне відео – у мультфільм, ваше фото – у ваш портрет, ніби

намальований видатним художником. Треба лише обрати стиль та натиснути одну кнопку. Таке трансформування стилю може використовуватися у сфері художньої обробки фото чи відео і особливо – у сфері розваг. Наприклад, накладання стилю в історіях користувачів Instagram у режимі реального часу лише одним рухом.

Також, розроблене ПЗ має бути крос-платформним та незалежним від інтерфейсної частини, представляти собою ядро обробки медіа даних, абстрагуючись від конкретної реалізації графічного інтерфейсу користувача.

Виходячи із вимог, розроблений трансформатор стилю представляє собою програмне забезпечення, розроблене на Python, з певним API для користування, що реалізує алгоритм згортувальної нейронної мережі (CNN) – безпосередньо використовує натреновану мережу для кожного конкретного стилю а також може тренувати нові конфігурації мережі для нових стилів. В якості інструменту для розробки нейронної мережі використовується Google TensorFlow Framework, функція втрат – VGG19, бібліотеки NumPy, SciPy, MoviePy. Для оптимізації процесів обробки медіа та тренування мереж, існує підтримка GPU з пакетом Nvidia CUDA та cuDNN (за замовчуванням встановлено CPU з BLAS/MKL).

Так як трансформатор стилю представляє собою крос-платформна програма для нейронної мережі, вона має лише API у вигляді параметрів запуску, отримуючи необроблений медіа-файл на вході та генеруючи оброблений файл на виході. Для нього може бути реалізований будь-який GUI, як у вигляді веб-сервісу, так і у вигляді додатку для якоїсь платформи чи у вигляді модуля складного комплексного програмного забезпечення. Для простоти демонстрації прикладу реалізації програми на практиці, в якості фінального зразка рішення з GUI, було використано додаток для Windows, розроблений на C++/Qt.