

УДК 004.85:004.931

КОНТРОЛЬ ЗА АПАРАТНИМ ЗАБЕЗПЕЧЕННЯМ МОБІЛЬНОГО ПРИБОРУ НА ОПЕРАЦІЙНІЙ СИСТЕМІ IOS ДЛЯ ВИКОНАННЯ ЗАДАЧ РОЗПІЗНАВАННЯ ОБРАЗІВ

Хрипунов І.М.

Науковий керівник – канд. техн. наук, доц. Каук В.І.

Харківський національний університет радіоелектроніки, каф. ПІ
м. Харків, Україна

тел. +38(050) 027-00-75

This work is devoted to the control of hardware support of a mobile device on the iOS platform during the operation of machine learning models for pattern recognition. The main libraries for implementing the integration of the model into the application are considered. The main capabilities of hardware control, which are available within the considered libraries for integration, are analyzed.

Можливість сприймати зовнішній світ у формі образів сприяє передумовам дослідження властивостей величезної кількості об'єктів завдяки ознайомленню з кінцевою їх кількістю, а об'єктивна ознака засадничої властивості образів допомагає створювати модель їх розпізнавання.

Для того, щоб отримати можливість інтеграції моделі для розпізнавання образів в мобільний додаток для iOS необхідно використати додаткові бібліотеки. На сьогоднішній день найбільш популярними інструментами для реалізації даної задачі є бібліотека CoreML від компанії Apple та TensorFlowLite, що є продуктом компанії Google.

Отже, використовуючи один з вищезазначених інструментів, розробник отримує можливість реалізувати інтеграцію моделі машинного навчання у мобільний додаток і будувати бізнес логіку застосунку, базуючись на цьому. Наступним етапом розробки є підвищення ефективності роботи інтегрованої моделі. Адже розпізнавання образів є досить трудомісткою задачею, яку можна пришвидшити за рахунок використання особливостей апаратного забезпечення iOS пристрою.

У разі інтеграції моделі машинного навчання за допомогою CoreML, розробник отримує можливість використати GPU, CPU або ANE (Apple Neural Engine) девайсу для виконання задачі розпізнавання образів.

Apple Neural Engine (або ANE) — це тип NPU, що розшифровується як Neural Processing Unit. Це як GPU, але замість прискорення графіки NPU прискорює операції нейронної мережі, такі як згортки та множення матриць. Більшість нових iPhone та iPad мають Neural Engine, спеціальний процесор, який робить моделі машинного навчання дуже швидкими.

Під час створення екземпляру моделі для подальшого використання та передбачень за допомогою інструменту від Apple, необхідно створити відповідні конфігурації, в межах яких розробник отримує можливість визначити, які самі частини апаратного забезпечення пристрою необхідно вико-

ристовувати для обробки зображень. Ось список доступних налаштувань:

- all (параметр, який дозволяє моделі використовувати всі доступні обчислювальні блоки, включаючи нейронні механізми);
- cpuOnly (параметр, який обмежує дозволяє моделі використовувати лише ЦП);
- cpuAndGPU (параметр, який дозволяє моделі використовувати обидва CPU та GPU, але не нейронний двигун);
- cpuAndNeuralEngine (параметр, який дозволяє використовувати моделі як центральний процесор, так і нейронний механізм, але не графічний процесор).

Вибір налаштування на етапі конфігурації моделі дає нам можливість контролювати її продуктивність та безпосередньо впливати на те, яке апаратне забезпечення буде використане системою для роботи моделі.

Контроль апаратного забезпечення за допомогою моделі TensorFlow Lite реалізовано за допомогою делегатів. Делегати дозволяють апаратне прискорення моделей TensorFlow Lite за допомогою вбудованих прискорювачів, таких як графічний процесор і процесор цифрових сигналів (DSP). За замовчуванням TensorFlow Lite використовує ядра ЦП, оптимізовані для набору інструкцій ARM Neon. Однак ЦП — це багатоцільовий процесор, який не обов'язково оптимізований для важкої арифметики, яка зазвичай зустрічається в моделях машинного навчання.

Для контролю апаратного забезпечення, що використовує модель доступний делегат GPU. Він оптимізований для запуску 32-розрядних і 16-розрядних моделей на основі плаваючої пам'яті, де є графічний процесор. Він також підтримує 8-розрядні квантовані моделі та забезпечує GPU продуктивність на рівні з їх версіями з плаваючою машиною. Іншим доступним інструментом є делегат Core ML, який дає можливість прискорити роботу моделі для 32 або 16-розрядних моделей із плаваючою комою.

Таким чином, вибір правильного делегату при роботі з TFLite дає розробнику можливість впливати на апаратне забезпечення для підвищення ефективності роботи моделі.

Список використаних джерел:

1. Заяць, В.М., & Камінський, Р.М. (2004). *Методи розпізнавання образів*. Нац. ун-т «Львів. політехніка».
2. *Core ML Tools*: сайт. URL: <https://coremltools.readme.io/docs> (дата звернення 10 березня 2022 р.).
3. *TensorFlow Lite*: сайт. URL: <https://www.tensorflow.org/lite> (дата звернення 12 березня 2022 р.).