

СРАВНЕНИЕ РАЗМЕРА БАТЧА ДЛЯ MLP-MIXER ДЛЯ ДАТАСЕТА CIFAR-10

Н. С. МОНТИК, А. П. ГОЛОВАЧ

УО «Брестский государственный технический университет», Брест, Беларусь,
nikolay.montik@gmail.com

Введение. Объектом исследования данной работы является архитектура нейронной сети MLP-Mixer, применяемая для решения задач классификации изображений на примере датасета CIFAR-10. MLP-Mixer представляет собой относительно новую архитектуру, которая, в отличие от традиционных сверточных нейронных сетей или трансформеров, основана исключительно на многослойных перцептронах. Такой подход позволяет исследовать возможности чисто линейных преобразований данных без использования сложных операций. Актуальность обусловлена тем, что современные методы глубокого обучения активно развиваются, и появляются новые архитектуры, которые стремятся упростить моделирование при сохранении высокой точности.

Эффективность обучения таких моделей существенно зависит от выбора гиперпараметров, одним из ключевых из которых является размер батча. Правильный подбор этого параметра может существенно влиять на скорость сходимости, стабильность обучения и итоговое качество модели. Целью данной работы является сравнительный анализ влияния различных размеров батча на процесс обучения модели.

Материалы и методы. CIFAR-10 — это набор данных, состоящий из 60000 цветных изображений размером 32×32 , разделённых на 10 классов (по 6000 изображений на каждый класс). Датасет разбит на обучающую выборку (50000 изображений) и тестовую (10000 изображений).

Для реализации поставленной задачи использовался язык программирования Python, библиотека pytorch. Значение параметра seed установлено равным 42. В качестве оптимизатора выбран Adam. Для повторяемости эксперимента недетерминированные алгоритмы CUDA, shuffle и т. д. были отключены, что несколько снизило точность. Обучение проводилось в течение 10 эпох.

Результаты и обсуждение. В качестве метрик оценки производительности использовались длительность обучения и точность. Результаты проведенных экспериментов представлены в таблице 1.

Таблица 1 – Сравнение точности и длительности обучения

Метрика	Размер батча					
	8	16	32	64	128	256
Точность, %	61,66	62,26	63,25	63,48	59,85	59,95
Длительность обучения, с.	844,26	361,23	286,17	182,62	142,92	137,01

Заключение. На основе проведенных экспериментов установлено, что размер батча, равный 32 или 64, обеспечивает оптимальный баланс между точностью и временем обучения для данной задачи.