

УДК 004.89

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ ДИСТИЛЛЯЦИИ ЗНАНИЙ К ЗАДАЧЕ КЛАССИФИКАЦИИ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА ПО СНИМКАМ МРТ

Ян Цзяфэн (ИТМО)

Научный руководитель – кандидат технических наук Русак А.В. (ИТМО)

Введение. Автоматизированные модели мультиклассовой классификации могут значительно повысить эффективность ранней диагностики опухолей головного мозга (ОГМ), что имеет решающее значение для выбора наилучших вариантов лечения. В последнее время модель Vision Transformer (ViT) [1] показала впечатляющую эффективность в задачах классификации изображений. Однако из-за того, что ручная аннотация опухолей рентгенологами требует значительных временных и финансовых затрат, типичные наборы данных для классификации ОГМ содержат всего лишь сотни или тысячи изображений. Известно, что модели-трансформеры показывают лучшие результаты при обучении на больших объемах данных. Когда данных недостаточно, их производительность часто отстает от производительности сверточных нейронных сетей, поскольку механизм внимания требует моделирования долгосрочных зависимостей. Данная работа направлена на исследование возможности применения архитектуры Vision Transformer и технологии дистилляции знаний к задаче мультиклассовой классификации ОГМ по снимкам магнитно-резонансной томографии (МРТ).

Основная часть. В исследовании используется комбинированный набор данных [2], содержащих 7022 МРТ-снимка с четырьмя классами опухолей: глиома, менингиома, гипофиз и отсутствие опухоли. Для обеспечения единообразия и совместимости с архитектурой модели изображения были изменены до размеров 224 x 224 пикселей. Кроме того, были использованы методы аугментации данных, включая случайное вращение с максимальным углом 10 градусов и случайное горизонтальное переворачивание. Для решения поставленной задачи мы использовали модель Data-efficient image Transformer (DeiT) [3] и VGG-16 в качестве базовой модели. Чтобы обеспечить обучение на небольших наборах данных, модель DeiT использует стандартную архитектуру кодировщика, используемую в ViT, и добавляет токен дистилляции знаний для соответствия смягченным прогнозам учителя. Реализация моделей основана на библиотеках PyTorch, Sci-kit Learn, vit-pytorch и timm на языке программирования Python. Все модели обучались на графических процессорах NVIDIA GeForce RTX 3070 в течение 100 эпох с использованием оптимизатора Adam, размером пакета 32 и скоростью обучения 0,00007. Модель DeiT продемонстрировала впечатляющие результаты, превосходя модель VGG-16 по достоверности (99,31% против 96,89%) и макро F1 оценке (99,28% против 96,80%). Это подтверждает ее способность эффективно работать с небольшими наборами данных при использовании метода дистилляции знаний.

Выводы. В данной работе мы оценили модель DeiT с применением дистилляции знаний для мультиклассовой классификации опухолей головного мозга по снимкам МРТ, используя VGG-16 в качестве базовой модели. Окончательное значение F1-меры модели DeiT составило 99,28%. Результаты показывают, что предобученная модель DeiT с использованием метода дистилляции знаний во время обучения достигла значительно лучшей производительности по сравнению с VGG-16. Это подтверждает возможность применения трансформеров в области анализа медицинских изображений. Дальнейшие исследования, направленные на улучшение обобщения модели, могут привести к тому, что трансформеры достигнут или превзойдут эффективность CNN в задачах медицинского анализа изображений.

Список использованных источников:

1. Dosovitskiy A., Beyer L., Kolesnikov A., et al. An image is worth 16x16 words:

Transformers for image recognition at scale // arXiv preprint arXiv:2010. – 11929. – 2020.

2. Brain Tumor MRI Dataset. – URL:

<https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>.

3. Touvron H., Cord M., Douze M., et al. Training data-efficient image transformers & distillation through attention // International conference on machine learning. PMLR. – 2021. – 10347–10357.