

Введение. Современные методы классификации изображений опираются на использование глубоких нейронных сетей, выбор архитектуры которых существенно влияет на качество модели. Разработка метода автоматизированного выбора архитектуры представляет собой актуальную задачу, поскольку позволяет минимизировать человеческие усилия [1][2]. В данной работе предлагается новый метод выбора архитектуры нейронной сети, основанный на концепции CLIP-подобных моделей, где на один вход подается изображение, а на другой — информация о характеристиках нейросетей [3]. Важной особенностью предложенного метода является разделение информации о сетях на несколько категорий, каждая из которых имеет свой набор параметров и весов, что позволяет учитывать уровень значимости различных характеристик.

Основная часть. Разработанная архитектура использует два входа: первый предназначен для обработки изображений, второй — для обработки метаданных о нейросетях. Входные данные второго типа подразделяются на четыре группы:

- Архитектурные характеристики (Ширина и глубина нейронной сети)
- Сложность (Количество операций с плавающей точкой и параметров)
- Базовые характеристики (Тип архитектуры и год создания)
- Архитектурные особенности (Нормализация, сквозные связи и т.д.)

Каждая из групп параметров обрабатывается отдельными слоями с индивидуальными весами, что позволяет модели определить наиболее значимые характеристики. Для обучения системы используется 1000 датасетов и портфолио из 100 моделей. Эксперименты проводились на открытых датасетах из Kaggle.

Выводы. Разработан метод выбора архитектуры нейросети для задач классификации изображений, который позволяет находить подходящую модель для любой задачи классификации изображений. Данный метод может быть применен в различных областях, таких как медицинская диагностика, автоматизация процессов в промышленности, распознавание лиц, анализ спутниковых снимков.

Список использованных источников:

1. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks //Advances in neural information processing systems. - 2012. - Т. 25.
2. He K. et al. Deep residual learning for image recognition //Proceedings of the IEEE conference on computer vision and pattern recognition. - 2016. - С. 770-778.
3. Radford A., Kim J. W., Hallacy C. Learning Transferable Visual Models From Natural Language Supervision // arXiv preprint. – 2021. – № arXiv:2103.00020.