

РАЗРАБОТКА АЛГОРИТМА КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ С ВСТРОЕННЫМИ АФФИННЫМИ ПРЕОБРАЗОВАНИЯМИ ДЛЯ УМЕНЬШЕНИЯ РАЗМЕРА НЕЙРОННОЙ СЕТИ

Петрова В. В.¹

Научный руководитель – канд. техн. наук Ходненко И. В.¹

¹Университет ИТМО

vika04vvp@mail.ru

Введение

В современных сверточных нейронных сетях устойчивость к поворотам в задачах классификации изображений часто достигается за счёт увеличения архитектурной сложности или вычислительных затрат. Так, существующие подходы решают задачу вращательной инвариантности за счёт перебора трансформаций на этапе инференса (Pre-rotation Only at Inference-Stage [1]) или усложнения архитектуры нейронных сетей (Spatial Transformer Networks [2], Invariant Integration [3], Oriented Response Networks [4] и др.), что приводит к росту вычислительных затрат и зависимости от большого объёма размеченных данных. В условиях ограниченных вычислительных ресурсов возникает необходимость создания компактных моделей классификации изображений, которые интегрируют учёт ориентации объектов непосредственно в архитектуру нейронной сети и позволяют снизить требования к размеру модели.

Основная часть

Предлагается компактный метод классификации изображений рукописных цифр, обеспечивающий устойчивость к вращательным преобразованиям за счёт интеграции оценки ориентации непосредственно в архитектуру нейронной сети. Модель реализована в виде сверточной нейронной сети с двумя выходными головами, предназначенными для классификации и регрессии угла поворота изображения. Предсказанный угол используется на этапе инференса для выборочной коррекции ориентации входных данных при низкой уверенности классификации, что позволяет снизить вычислительные затраты по сравнению с перебором трансформаций. Обучение проводится на синтетических данных с контролируемыми геометрическими параметрами, что уменьшает зависимость от объёма размеченных реальных данных и повышает воспроизводимость результатов.

Выводы

Предложен компактный алгоритм классификации изображений с учётом ориентации, объединяющий задачи классификации и оценки угла поворота в рамках одной нейронной модели. Использование выборочной коррекции ориентации позволяет повысить устойчивость к вращениям при снижении вычислительных затрат. Достигнут прирост точности классификации на 3.47 % при обучении на синтетических данных по сравнению с базовой моделью без коррекции ориентации при тестировании на MNIST.

Литература

1. Fan Y., Zhang P., Han J., Liu D., Tang J., Zhang G. Pre-rotation Only at Inference-Stage: A Way to Rotation Invariance of Convolutional Neural Networks [Электронный

- ресурс]. – Режим доступа: <https://link.springer.com/article/10.1007/s44196-024-00490-z> (дата обращения: 19.11.2025).
2. Jaderberg M., Simonyan K., Zisserman A., Kavukcuoglu K. Spatial Transformer Networks [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1506.02025> (дата обращения: 19.10.2025).
 3. Rath M., Condurache A. P. Improving the Sample-Complexity of Deep Classification Networks with Invariant Integration [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/2202.03967> (дата обращения: 19.10.2025).
 4. Zhou Y., Ye Q., Qiu Q., Jiao J. Oriented Response Networks [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1701.01833> (дата обращения: 19.10.2025).