

УДК 004.023

**Разработка алгоритма поиска архитектуры свёрточных сетей Колмогорова-Арнольда на основе эволюционной оптимизации**

**Латыпов В.В. (ИТМО)**

**Научный руководитель – кандидат физико-математических наук Хватов Александр Александрович (ИТМО)**

**Введение.** Последние несколько лет стала актуальна и популярна в исследованиях по машинному обучению тема поиска новых архитектур нейронных сетей (например, Mamba, KAN). Сети Колмогорова-Арнольда (KAN) — особенно популярная в последние полгода альтернатива многослойному перцептрону (MLP), введённая в [1], основанная на обучении коэффициентов разложения сплайновых функций активации по базису В-сплайнов.

Статья ConvKAN [2] — вводит свёрточные KAN: свёрточный слой Колмогорова-Арнольда формулируется перенесением логики сплайновых активаций в операцию линейной свёртки.

**Основная часть.** Для обоснования новых архитектур необходимо показать, что они либо вычислительно эффективнее, либо превосходят существующие по некоторой мере качества для некоторого класса задач. Чаще всего архитектуры новых сетей строятся под влиянием гайдлайнов для старых, которые могут для них не подходить, следовательно, новые архитектуры могут быть отброшены, не реализовав свой потенциал.

Цель — получить модели свёрточных KAN без априорных предположений об архитектуре и более честно сравнить свёрточные KAN и CNN на задачах классификации и предсказания временных рядов, локализовав область поиска применений KAN. Достичь этого позволяет автоматизированный поиск с помощью алгоритма поиска нейросетевой архитектуры (NAS), находящего оптимальную архитектуру даже в условиях отсутствия гайдлайнов для неё в сообществе.

С этой целью реализован эволюционный алгоритм NAS для задачи классификации изображений и предсказания многомерных временных рядов с физическими данными и произведено сравнение архитектур на основе этого алгоритма.

Разработаны кодировка с учётом специфики различных архитектур, эволюционные операторы и правила валидации. Внедрён многокритериальный эволюционный алгоритм, где в качестве метрик используются качество и метрика сложности моделей: количество обучаемых параметров или количество FLOPs.

Для полноты сравнения предложен свёрточный KAN слой с альтернативной геометрией и архитектура на его основе. Также реализованы транспонированные свёрточные слои KAN для преобразования физических данных.

Произведён отбор типов слоёв для каждого вида архитектуры. Для оценки результатов происходит статистическая оценка полученных алгоритмом NAS моделей.

В качестве эволюционной кодировки для задачи предсказания многомерных временных рядов физических данных используется шаблон архитектуры кодировщик-декодировщик, так как для физических данных использование только обычных свёрточных слоёв искажает локальную информацию, внося ошибку в производные поля высших порядков.

Проведены эксперименты на часто используемых бенчмарках для NAS (MNIST/Fashion MNIST, CIFAR10, EuroSAT) для задачи классификации и на датасете OSI SAF Global Sea Ice Concentration (SSMIS) для задачи предсказания временных рядов.

Для большинства видов архитектур и датасетов алгоритм поиска архитектуры находит модели, доминирующие по Парето модели, предложенные в [2].

Выполненное в данной работе сравнение на основе результатов NAS для свёрточных KAN и CNN позволяет, в отличие от [3], корректно определить применимость KAN к работе с изображениями, т.к.

- В данной работе используются свёрточные, а не полносвязные KAN.
- В данной работе используется NAS на множестве произвольных ациклических графов-паттернов соединения сети — более эффективный алгоритм, чем grid search, а также более широкое и свободное от произвольности выбора исследователей пространство поиска.

**Выводы.** Предложен алгоритм эволюционного поиска архитектуры свёрточных KAN для задач классификации изображений и предсказания многомерных временных рядов. Произведено сравнение оптимальных конфигураций архитектур, найденных алгоритмом. В ходе работы реализованы различные новые типы свёрточных KAN слоёв. Произведено сравнение с существующими работами на известных датасетах, превзойдены метрики качества и сложности моделей.

#### **Список использованных источников:**

1. Liu Z., Wang Y., Vaidya S. и др. KAN: Kolmogorov-Arnold Networks [Электронный ресурс] // arXiv.org. 2024. URL: <https://arxiv.org/abs/2404.19756>
2. Bodner A.D., Tepsich A.S., Spolski J.N. и др. Convolutional Kolmogorov-Arnold Networks [Электронный ресурс] // arXiv.org. 2024. URL: <https://arxiv.org/abs/2406.13155>
3. Yu R., Yu W., Wang X. KAN or MLP: A Fairer Comparison [Электронный ресурс] // arXiv.org. 2024. URL: <https://arxiv.org/abs/2407.16674>