

НАЗВАНИЕ ТЕЗИСА

Нематов А. Т.,
Научный руководитель – канд. физ. мат. наук, доцент Графеева Н. Г.
Университет ИТМО
aziznematov.job@gmail.com

Работа выполнена в рамках темы НИР №425041 «Разработка приложения по оцифровке и атрибуции кириллических рукописных текстов с применением методов компьютерного зрения и мультимодальных моделей.».

Введение

В современной архивной науке наблюдается парадокс: темпы оцифровки рукописных фондов многократно опережают возможности экспертов по их атрибуции. Идентификация автора или времени создания кириллической скорописи XV–XVIII веков до сих пор остается трудоемким процессом, зависящим от субъективного опыта палеографа. Существующий зарубежный опыт [1] показывает переход от жестких каскадных систем к сквозному глубокому обучению (end-to-end), что позволяет минимизировать ошибки предварительной обработки. Однако большинство западных решений оптимизировано под латиницу и не учитывает специфику старославянского письма. Проблема создания эффективной информационно-поисковой системы, способной находить стилистические аналогии в огромных массивах неразмеченных изображений, остается открытой.

Основная часть

В качестве решения предлагается метод атрибуции, основанный на извлечении векторных представлений (эмбедингов) признаков почерка без промежуточного этапа полного распознавания текста. Основой подхода является декомпозиция сканов высокого разрешения на патчи фиксированного размера с заданным перекрытием. Это позволяет нейросетевой модели фокусироваться на микро-характеристиках начертания отдельных графем и их соединений.

Для формирования поискового вектора используется архитектура, вдохновленная глобальными регулярными сетями [2], которые эффективно агрегируют локальные признаки в единый дескриптор автора. Новизна метода заключается в применении метрического обучения (metric learning): модель обучается «стягивать» в векторном пространстве фрагменты рукописей одного писца и «отталкивать» разные. Для обеспечения доверия экспертов-историков в систему внедряется механизм визуализации карт активации признаков [3]. Это позволяет наглядно продемонстрировать, какие именно элементы начертания (например, характерный вынос хвоста буквы или наклон) стали решающими для атрибуции. Такой подход превращает нейросеть из закрытого инструмента в ассистента исследователя.

Выводы

Предложенный метод позволяет автоматизировать первичную группировку анонимных рукописей, сокращая время поиска аналогий в архивах в десятки раз. Практическая реализация системы планируется в НИР. Внедрение подобных инструментов в работу государственных архивов позволит существенно облегчить работу исследователей и возможно расширить возможности для масштабных исследований.

Литература

1. Rasyidi H., Farazi M. Beyond the Pipeline: Analyzing Key Factors in End-to-End Deep Learning for Historical Writer Identification. 2025. [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/2510.18671> (дата обращения: 26.02.2026).
2. Wang S. Global Regular Network for Writer Identification. Xian Jiaotong University, 2022. [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/2201.05951> (дата обращения: 26.02.2026).
3. Liao Y., Gao Y., Zhang W. Feature Activation Map: Visual Explanation of Deep Learning Models for Image Classification // IEEE Transactions on Image Processing. 2023. [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/2307.05017> (дата обращения: 26.02.2026).