

ОЦИФРОВКА СТАРОСЛАВЯНСКИХ РУКОПИСНЫХ ТЕКСТОВ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ

Никишин А.П. (ИТМО), Вдовиченко М.С. (ИТМО)

Научный руководитель – кандидат физико-математических наук, доцент Графеева Н.Г. (ИТМО)

Введение. Цифровизация и распознавание рукописных исторических документов позволят сохранить культурное наследие и обеспечить доступ к историческим архивам для научного сообщества и широкой публики. Несмотря на значительные достижения в области глубокого обучения, существующие системы распознавания рукописного текста испытывают сложности при обработке изображений из-за разнообразия стилей письма, плохого состояния сохранности документов и необходимости адаптации к текстам из различных исторических периодов. В данной научной работе рассматривается процесс оцифровки рукописного текста послания антиохийского патриарха Петра III Доминику, архиепископу Градо [1], с применением методов глубокого обучения.

Основная часть. В рамках нашей работы рассматриваются двухстраничные изображения. Такие изображения нередко содержат значительные искажения и дефекты, характеризующиеся низким разрешением и высоким уровнем шума, особенно по краям.

Для оцифровки изображений мы предлагаем следующие этапы работы:

- 1) Этап предобработки изображений:
 - a. Разделение двухстраничных изображений на отдельные страницы.
 - b. Коррекция перекоса с использованием градиента яркости по оператору Собеля.
 - c. Определение текстового корпуса и выделение текстовых линий на основе гистограмм интенсивности.
 - d. Бинаризация изображений.
- 2) Этап реализации моделей глубокого обучения:
 - a. Использование сверточных нейронных сетей (CNN) для выявления пространственных иерархий и локальных особенностей изображений.
 - b. Использование рекуррентных нейронных сетей (RNN), в частности сети с длительной кратковременной памятью (LSTM), для учета последовательный контекст, что позволит обойти ограничения присущие CNN в обработке последовательного контекста.
 - c. Использование гибридных моделей, таких как CNN-BiLSTM [2,3] (или CNN-BiLSTM-CTC) и комбинированных моделей, включающих vision transformer и text transformer [4].

Выводы. В данном исследовании был разработан алгоритм для устранения дефектов изображений на этапе предварительной обработки, а также были протестированы различные модели глубокого обучения, включая гибридные модели CNN-LSTM и архитектуры трансформеров. Эти модели демонстрируют эффективность в работе как с визуальной, так и с текстовой информацией, что открывает новые возможности для улучшения точности распознавания рукописных текстов.

Список использованных источников:

1. Сизиков А.В. Послание патриарха Антиохийского Петра III об опресноках в славянской письменности // Библия и христианская древность. — 2020. — С. 15—55.
2. Lulseged Yimer H., Dereje Degefa H., Cristani M., Cunico F. Learning based Ge'ez character handwritten recognition // arXiv e-prints. — 2024. — arXiv-2411.
3. Hamdan M., Rahiche A., Cheriet M. HTR-JAND: Handwritten Text Recognition with Joint Attention Network and Knowledge Distillation // arXiv preprint arXiv:2412.18524. — 2024.
4. Dolfing H.J. Whole page recognition of historical handwriting // arXiv preprint arXiv:2009.10634. — 2020.