

## ИССЛЕДОВАНИЕ ГЕНЕРАТИВНЫХ МОДЕЛЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧИ РАСПОЗНОВАНИЯ РУКОПИСНОГО ТЕКСТА НА ПРИМЕРЕ ФРАНЦУЗСКОГО ЯЗЫКА

Скворцова А.Ю. (ИТМО)

Научный руководитель – Трифанов А.И. (ИТМО)

Научный консультант – Михайленко А.Ю. (ИТМО)

**Введение.** Распознавание рукописных текстов представляет собой ключевую задачу в сфере компьютерного зрения и обработки естественного языка. С развитием технологий искусственного интеллекта все большее внимание уделяется применению генеративных сетей для решения этой проблемы. Для таких языков, как английский, уже разработано множество высококачественных решений. Однако при использовании современных генеративных сетей для распознавания рукописных текстов на французском языке возникают определенные трудности. Во-первых, сложная грамматическая структура французского языка усложняет модель и увеличивает время обработки текста. Во-вторых, особенности французского письма создают проблемы с распознаванием диакритических знаков и их отличием от других символов. В-третьих, скорость работы генеративных сетей может оказаться недостаточной для обработки данных в реальном времени, особенно при больших объемах информации. Наконец, требования к вычислительным ресурсам при обучении и использовании этих сетей могут быть значительными, что ограничивает их практическое применение.

**Основная часть.** Для эффективного применения генеративных сетей в задаче распознавания рукописных франкоязычных текстов необходимо провести исследования и оптимизировать методы обучения и работы сетей, чтобы достичь более точных результатов, повысить скорость работы и снизить требования к вычислительным ресурсам. В данном исследовании предлагается рассмотреть существующие архитектуры, адаптировать существующие подходы для английского языка к французскому языку. Провести разбор предобученных моделей для их последующей настройки под конкретную задачу распознавания рукописного текста. Финальным этапом разработки можно считать реконструкцию модели для улучшения показателей скорости и качества распознавания при необходимости.

**Выводы.** Построенная модель может быть применена для распознавания рукописного текста в отсканированных документах, письмах или рукописных записях. Такая модель может также быть использована для распознавания и анализа рукописных подписей, что может быть полезно в банковском секторе, юридических документах и других областях. Эта модель применима и в искусстве: распознавание рукописных текстов, написанных в различных стилях каллиграфии, может оказаться полезным для художников, дизайнеров и любителей каллиграфии. Отдельно стоит выделить медицинскую сферу: такая модель может распознавать рукописные записи в медицинских картах или отчетах, что поможет в улучшении точности и скорости анализа медицинских данных. Также модель сможет стать основой для разработки генеративных моделей для других языков с письменной структурой, сложнее англоязычной.

**Список использованных источников:**

1. *Sharon Fogel, Hadar Averbuch-Elor, Sarel Cohen, Shai Mazor and Roei Litman* ScrabbleGAN: Semi-Supervised Varying Length Handwritten Text Generation // arXiv. 2020. V. 1.
2. *Jorge Juan Sueiras Revuelta* Continuous Offline Handwriting Recognition using Deep Learning Models // arXiv. 2021. V. 1.
3. *Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser* Attention Is All You Need // arXiv. 2023. V. 7.