

## ВЕКТОРНЫЕ ПРЕДСТАВЛЕНИЯ ШРИФТОВ

Щурик М. В., Копылов Д. Е.

**Ключевые слова:** стилевые эмбендинги, классификация шрифтов, компьютерные шрифты, сверточные нейронные сети

Актуальность работы обусловлена необходимостью автоматического анализа стилей шрифтов при обработке документов. Современные исследования применяют сверточные нейронные сети (CNN) в различных прикладных задачах, однако сохраняется ключевая проблема – высокая размерность получаемых векторных представлений, что делает их неэффективными для обработки больших объемов документов [1]. В данной работе предлагается подход к созданию компактных векторных представлений шрифтов с использованием CNN, что позволит улучшить обработку документов [2].

Важную часть исследования составляет подготовка данных. Поскольку единой классификации шрифтов нет, нами были выделены основные признаки, представленные в таблице «табл. 1».

Признак	Вариации
Начертание	Тонкий, Обычный, Жирный
Наклон	Прямой, Курсив
Моноширинность	Да/Нет
Засечки	Гротеск, Антиква, Брусковый
Рукописность	Да/Нет
Форма	Прямоугольный/Круглый

Таблица 1. Классификация признаков шрифтов

Было сгенерировано три набора данных: стандартные шрифты из Microsoft Word, вариативные шрифты с регулируемыми параметрами (жирность, наклон, ширина) и их комбинации.

Архитектура модели включает в себя два компонента: CNN для генерации 8-мерных векторных представлений и модель DiffModel для их сравнения «рис. 1». Модель CNN была вдохновлена архитектурой VGG-NET [3, с. 955]. Данная архитектура представляет собой блоки, состоящие из сверточных слоев и функции MaxPooling «рис. 2». DiffModel – модель, которая на выходе представляет собой бинарный классификатор.

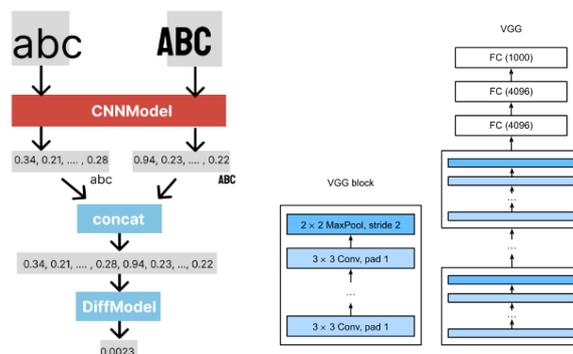


Рис. 1. Архитектура модели Рис. 2. Архитектура VGG-NET

В работе рассматривается две модели CNN: с одним и двумя блоками VGG. Обучение проводится на смешанном датасете, тестирование на разных наборах данных. Результаты представлены в таблице «табл. 2».

	Стандартный датасет	Вариативный датасет	Смешанный датасет
1 блок VGG	0.8854	0.9054	0.8954
2 блока VGG	0.8759	0.8762	0.8760

Таблица 2. Результат тестирования моделей

Для решения этих задач используется библиотека PyTorch на языке программирования Python [4]. Анализ показал, что одноблочная архитектура VGG наиболее эффективна. Её планируется интегрировать в систему обработки документов для оптимизации и ускорения решения задач.

### Список литературы

1. Jiang S., Wang Zh., Hertzmann A. et al. Visual font pairing // IEEE Transactions on Multimedia. 2019. Vol. 22, N 8. P. 2086-2097.
2. Lecun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE. 1998. Vol. 86, N 11. P. 2278-2324.
3. Swapna M., Sharma Y., Prasad B. CNN Architectures: Alex Net, Le Net, VGG, Google Net, Res Net // International Journal of Recent Technology and Engineering (IJRTE). 2020. Vol. 8, N 6. P. 953-959.
4. PyTorch documentation [Электронный ресурс]. – URL: <https://docs.pytorch.org/docs/stable/> (дата обращения: 21.01.2025).