

Тюгаев Н.П.

студент

Самарский национальный исследовательский университет

**ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ
СОПОСТАВЛЕНИЯ ЛОКАЛЬНЫХ ПРИЗНАКОВ НА
ИЗОБРАЖЕНИЯХ**

*Аннотация: В работе исследованы нейросетевые модели для сопоставления локальных признаков на изображениях. Рассмотрены детекторы и дескрипторы ключевых точек DISK, SuperPoint, SIFT, а также методы сопоставления LightGlue, KeyNet-AdaLAM и LoFTR. Проведена программная реализация методов на Python и выполнено сравнительное тестирование на изображениях в помещении и локаций вне помещений. Установлено, что LoFTR показала наилучшую точность для сцен в помещении (0,921), LightGlue – для сцен вне помещений (0,914).
Ключевые слова: компьютерное зрение, ключевые точки, нейронные сети, сопоставление признаков, LightGlue, LoFTR, DISK.*

Tyugaev N.P.

student

Samara National Research University

**INVESTIGATION OF NEURAL NETWORK MODELS FOR LOCAL
FEATURE MATCHING ON IMAGES**

Abstract: The paper investigates neural network models for local feature matching on images. Keypoint detectors and descriptors DISK, SuperPoint, and SIFT, as well as matching methods LightGlue, KeyNet-AdaLAM, and LoFTR are considered. Software implementation of the methods in Python was carried out and comparative testing on indoor and outdoor images was performed. LoFTR achieved the highest accuracy for indoor scenes (0.921), while LightGlue performed best for outdoor scenes (0.914).

Keywords: computer vision, keypoints, neural networks, feature matching, LightGlue, LoFTR, DISK.

Введение

Нейросетевые модели для поиска соответствующих точек на изображениях становятся всё более важными благодаря их способности решать сложные задачи компьютерного зрения. Эти технологии находят применение в создании карт, 3D-реконструкции, автономном управлении транспортом и дополненной реальности. В условиях, где традиционные алгоритмы сталкиваются с трудностями из-за изменений освещения, угла съёмки или текстурных особенностей сцены, нейросетевые подходы демонстрируют более высокую точность и надёжность [1].

Целью данной работы является исследование нейросетевых моделей для сопоставления локальных признаков на изображениях. Задачами являются: анализ и сравнение существующих методов, их программная реализация, а также тестирование и анализ работоспособности реализованных методов.

1. Анализ методов поиска соответствующих точек на изображениях

В работе рассмотрены детекторы и дескрипторы ключевых точек: DISK – метод на основе глубокого обучения с оптимизацией через reinforcement learning непосредственно для задачи сопоставления [2]; SuperPoint – CNN-модель с самообучением на синтетических и реальных данных [3]; SIFT – классический метод на основе экстремумов в пространстве масштабов и гистограмм ориентированных градиентов [4].

Для сопоставления локальных признаков исследованы: LightGlue – адаптивная нейронная сеть на механизмах self-attention и cross-attention, динамически подстраивающая вычисления под сложность пары изображений, адаптивный механизм [5] которой показан на рисунке 1; KeyNet-AdaLAM – комбинация нейросетевого детектора и алгоритма фильтрации с учётом локальных аффинных преобразований [6]; LoFTR –

трансформерная модель плотного сопоставления, работающая без явного извлечения дескрипторов [7].

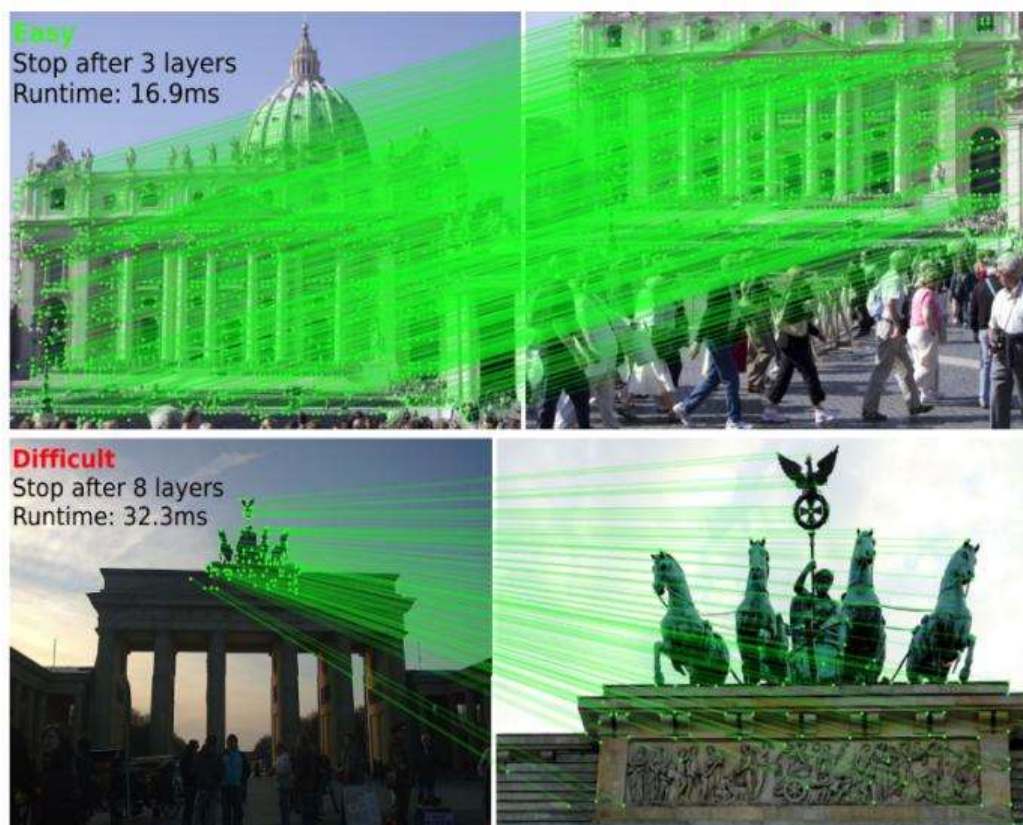


Рисунок 1 – Адаптивный механизм LightGlue

2. Реализация и тестирование методов

Все три метода реализованы на Python с использованием библиотеки Kornia и платформы Google Colab. Тестирование проводилось на парах изображений в помещении и локаций вне помещений. Количественные результаты сравнения методов на изображениях в помещении представлены в таблице 1, а для изображений сцен вне помещений – в таблице 2.

Таблица 1 – Количественные результаты сравнения методов на изображениях в помещении

Метод	Соответствий	Точность
LoFTR	467	0,921
LightGlue+DISK	308	0,873
KeyNet-AdaLAM	46	0,518
SIFT	84	0,724

Согласно таблице 1, на изображениях в помещении SIFT как классический метод уступил нейросетевым методам LightGlue и LoFTR, показавшему лучший результат.

Таблица 2 – Количественные результаты сравнения методов на изображениях сцен вне помещений

Метод	Соответствий	Точность
LoFTR	358	0,887
LightGlue+DISK	411	0,914
KeyNet-AdaLAM	127	0,675
SIFT	218	0,789

Согласно таблице 2, на изображениях сцен вне помещений лидирует LightGlue. SIFT выявил больше соответствий, чем KeyNet-AdaLAM, но значительно уступил LightGlue и LoFTR.

Примеры работы LoFTR на парах изображений в помещении приведены на рисунке 2, а примеры работы LightGlue на парах изображений вне помещений – на рисунке 3.

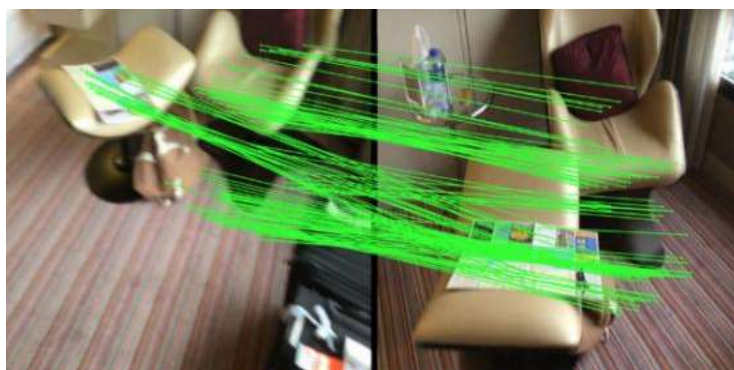


Рисунок 2 – Примеры работы LoFTR на парах в помещении изображений

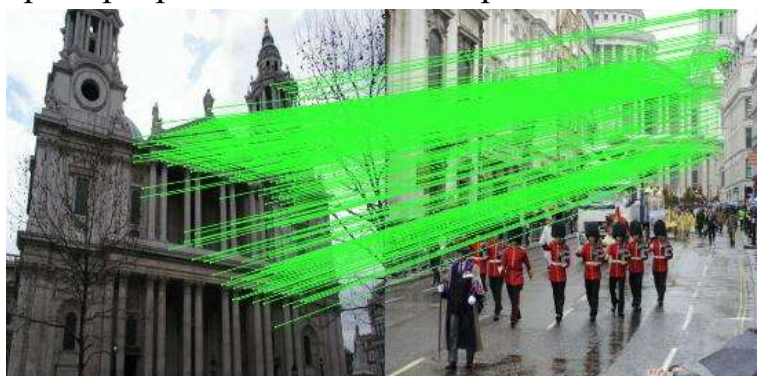


Рисунок 3 – Примеры работы LightGlue на парах изображений сцен вне помещений

Заключение

В работе проведено количественное сравнение нейросетевых методов сопоставления локальных признаков на изображениях. Рассмотрены детекторы DISK, SuperPoint, SIFT и методы сопоставления LightGlue, KeyNet-AdaLAM, LoFTR. Все методы реализованы на Python и протестированы на парах изображений в помещении и сцен вне помещений.

По результатам тестирования LoFTR показала наилучшую точность на сценах в помещении (0,921), а LightGlue лидировала на сценах вне помещений (0,914). Классический метод SIFT уступил нейросетевым подходам по всем показателям: точность не превысила 0,789, а количество соответствий было значительно ниже. KeyNet-AdaLAM продемонстрировала наихудшие результаты из рассмотренных методов.

Использованные источники:

1. Детекция ключевых точек лица с помощью капсульных нейронных сетей / А. А. Бойцев, Д. Г. Волчек, Е. Н. Магазенков [и др.]. // Научно-технический вестник информационных технологий, механики и оптики. – 2023. – №3. – С. 506–517. – URL: <https://cyberleninka.ru/article/n/detektsiya-klyuchevyhtochek-litsa-s-pomoschyu-kapsulnyh-neyronnyh-setey> (дата обращения: 25.02.2026).
2. Tyszkiewicz, M. Disk: learning local features with policy gradient / M. Tyszkiewicz, P. Fua, E. Trulls // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 1–12. – URL: <https://proceedings.neurips.cc/paper/2020/file/a42a596fc71e17828440030074d15e74-Paper.pdf> (дата обращения: 25.02.2026).
3. DeTone, D. Superpoint: Self-supervised interest point detection and description / D. DeTone, T. Malisiewicz, A. Rabinovich // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. – 2018. –

- Р. 224–236. – URL: https://openaccess.thecvf.com/content_cvpr_2018_workshops/papers/w9/DeTone_SuperPoint_Self-Supervised_Interest_CVPR_2018_paper.pdf (дата обращения: 26.02.2026).
4. Сафонов, А. С. Построение sift-дескрипторов и нахождение особых точек на изображениях / А. С. Сафонов // Известия ТулГУ. Технические науки. – 2017. – №2. – С. 183–187. – URL: <https://cyberleninka.ru/article/n/postroenie-sift-deskriptorov-i-nahozhdenie-osobyh-tochek-na-izobrazheniyah> (дата обращения: 26.02.2026).
5. Lindenberger, P. LightGlue: Local Feature Matching at Light Speed / P. Lindenberger, P. Sarlin, M. Pollefeys. – DOI: 10.48550/arXiv.2306.13643 // ArXiv : [сайт] / Cornell University. – 2023. – URL: <https://arxiv.org/abs/2306.13643> (дата обращения: 10.03.2026).
6. AdaLAM: Revisiting Handcrafted Outlier Detection / L. Cavalli, V. Larsson, M. Oswald [et al.]. – DOI: 10.48550/arXiv.2006.04250 // ArXiv : [сайт] / Cornell University. – 2020. – URL: <https://arxiv.org/abs/2006.04250> (дата обращения: 10.03.2026).
7. Efficient LoFTR: Semi-Dense Local Feature Matching with Sparse-Like Speed / Y. Wang, X. He, S. Peng [et al.]. – DOI: 10.48550/arXiv.2403.04765 // ArXiv : [сайт] / Cornell University. – 2024. – URL: <https://arxiv.org/abs/2403.04765> (дата обращения: 15.03.2026).