

**Предобработка изображений как способ повышения  
эффективности нейросетевых моделей**

**Назаренко А.А. (ИТМО)**

**Научный руководитель – к.ф.-м.н., доцент, директор ВШЦК, Михайлова Е.Г.**

**Научный консультант – к.ф.-м.н., доцент ВШЦК, Самарин А.В. (ИТМО)**

**Введение.**

Предобработка изображений – критический этап в компьютерном зрении, направленный на улучшение качества входных данных для нейросетевых моделей [4]. Её значимость обусловлена:

- Чувствительностью нейросетей к шумам, артефактам и вариациям в данных (например, размытие, засветы, зашумленность).

- Необходимостью адаптации к специфике задач: от медицинской визуализации (сегментация рентгеновских снимков, детектирование микроорганизмов в микроскопии) до анализа видеопотока (распознавание поведения студентов на экзамене в разных условиях съемки).

- Повышением устойчивости моделей за счет стандартизации данных (цветовая коррекция, устранение искажений).

Цель данной работы – сравнить современные методы предобработки, выделив их сильные и слабые стороны для различных сценариев. Упор сделан на легковесность подходов – необходимо, чтобы модель предобработки была максимально простой и вычислительно эффективной, потому что она будет использоваться вкупе с другими ресурсоемкими подходами. Кроме того, область легковесных нейросетевых подходов мало исследована.

**Основная часть.**

В ходе исследования существующих решений были выделены их достоинства и недостатки. Были рассмотрены классические способы предобработки изображений: аугментация данных, нормализация гистограммы и схожие. Кроме того, в работе рассмотрены нейросетевые подходы на базе генеративного компьютерного зрения (Pix2Pix [6] EnhanceGAN [7]), а также комбинированные легковесные подходы, сочетающие в себе использование нейросетевого кодировщика и классические дифференцируемые фильтры, моделирующие инструменты программы фоторедактора.

Ключевые наблюдения:

- Классические методы (аугментация данных, нормализация гистограммы, и другие) подходят для базовой стандартизации, но не решают проблему низкого качества данных. Достоинства: простота в реализации, повышение устойчивости к вариациям изображений, отсутствие артефактов и вычислительная эффективность. Недостатки: не улучшается исходное качество изображений, нет адаптации к контексту изображения.

- GAN (Pix2Pix [6], EnhanceGAN [7] и другие) эффективны для задач с высокими требованиями к детализации, но неприменимы в медицине из-за артефактов. Достоинства: высокое качество работы на сложных данных, улучшение детализации изображений, возможность обучаемого повышения разрешения исходного изображения. Недостатки: тяжеловесность (миллионы параметров), большая вычислительная нагрузка как на обучении, так и на предсказании, присутствие нейросетевых артефактов на генерациях.

- LFIEM [1] и UniFi [2] сочетают интерпретируемость, скорость и легковесность, что делает их идеальными для мобильных решений. LFIEM – легковесная модель (101К параметров), использующая фильтры и регуляризацию согласованности. Достоинства: интерпретируемость, высокая скорость работы, отсутствие артефактов, портируемость на edge devices. Недостатки: ограниченный набор фильтров, генератор параметров вызывается несколько раз. UniFi же, в отличие от LFIEM, основан на базе универсальных обучаемых ядер свертки и позволяет вызывать генератор только один раз. Из недостатков только

ограниченный набор фильтров, а также требование обучения на целевых данных.

- Unpaired UniFi [3] открывает возможности для работы с неразмеченными данными, но требует сложного обучения модели. Особенность подхода в том, что это легковесная модель (47K параметров), используются дифференцируемые фильтры. Подход к обучению: GAN Unsupervised с мультимодальным CLIP-дискриминатором. Подход позволяет получить более хорошие метрики качества. Достоинства: интерпретируемость, высокая скорость работы, отсутствие артефактов, портируемость на edge devices. Недостатки: ограниченный набор фильтров, сложный процесс обучения.

Рассмотренные методы предобработки были интегрированы в разные задачи, например, подход LFIEM был интегрирован в задачу классификации удостоверений личности [4], что позволило увеличить F1-меру со значений 0.91 до 0.96. В задаче классификации фасадов зданий F1-мера возрастает с 0.63 до 0.68 за счет коррекции освещения [4].

### **Выводы.**

В данной работе проведен обзор современных методов предобработки изображений перед их использованием в нейронных сетях. Предобработка во многих задачах является критическим этапом, так как она устраняет шумы и артефакты, стандартизирует данные и адаптирует изображения под специфику задачи. В качестве вывода можно сказать, что для embedded-систем (мобильные устройства, IoT), а также для обеспечения оптимального использования ресурсов серверов следует использовать легковесные методы (UniFi, LFIEM), которые сочетают скорость, низкую ресурсоемкость и отсутствие артефактов. Некоторые из рассмотренных методов предобработки были внедрены в различные задачи компьютерного зрения и позволили повысить точность предсказаний моделей нейросети.

### **Список использованных источников:**

1. Tatanov O., Samarin A. LFIEM: Lightweight Filter-based Image Enhancement Model // 2020 25th International Conference on Pattern Recognition (ICPR). — 2021. — С. 873–878.
2. Samarin, A., Nazarenko, A., Savelev, A., Toropov, A., Dzestelova, A., Mikhailova, E., Motyko, A. and Malykh, V., 2024. A Model Based on Universal Filters for Image Color Correction. *Pattern Recognition and Image Analysis*, 34(3), pp.844-854.
3. Samarin, A., Nazarenko, A., Toropov, A., Kotenko, E., Dzestelova, A., Mikhailova, E., Malykh, V., Savelev, A. and Motyko, A., 2024, October. Universal Filter-Based Lightweight Image Enhancement Model with Unpaired Learning Mode. In *2024 36th Conference of Open Innovations Association (FRUCT)* (pp. 711-720). IEEE.
4. Lightweight model for automatic image enhancement: [Электронный ресурс]. URL: (Дата обращения: 10.01.2024)
5. Samarin, A., Toropov, A., Dzestelova, A., Nazarenko, A., Kotenko, E., Mikhailova, E., Malykh, V., Savelev, A., Motyko, A. and Dozortseva, A., AutoML Applications for Bacilli Recognition by Taxonomic Characteristics Determination over Microscopic Images.
6. Isola, P., Zhu, J., Zhou, Y. and Efros A., 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).
7. Deng, Y., Loy, C.C. and Tang, X., 2018, October. Aesthetic-driven image enhancement by adversarial learning. In *Proceedings of the 26th ACM international conference on Multimedia* (pp. 870-878).