

УДК 519.61:004.93

ЧИСЛЕННЫЕ МЕТОДЫ ПРЕОБРАЗОВАНИЯ МАТРИЦ ДЛЯ ОПТИМИЗАЦИИ ОБРАБОТКИ ИЗОБРАЖЕНИЙ В ML

Булавин К.С., Степченков Р.О., студенты

Белорусский государственный университет информатики и радиоэлектроники¹
г. Минск, Республика Беларусь

Степанова Т.С. – канд. физ.-мат. наук, доцент

Аннотация. Работа посвящена исследованию эффективности применения сингулярного разложения (SVD) и анализа числа обусловленности матриц как метода интеллектуальной предобработки визуальных данных для нейросетевых моделей. Разработан гибридный алгоритм динамической аппроксимации, адаптирующий ранг разложения под информационную плотность кадра. Экспериментально доказано, что метод позволяет сократить объем передаваемых данных на 21,45% при сохранении точности детекции объектов (mAP@0.5) на уровне 96,2% от базовой модели YOLOv8n. Исследование ориентировано на применение в Edge-системах, БПЛА и автономном транспорте.

Ключевые слова. Матрица, сингулярное разложение, сингулярное число, SVD, машинное обучение, YOLOv8, сжатие данных, оптимизация, парето-эффективность, обработка изображений.

Современный мир пронизан визуальными потоками данных: от камер интеллектуальных транспортных систем до сенсоров беспилотных летательных аппаратов, анализирующих поверхность земли в реальном времени. Мы привыкли к тому, что разрешение камер растет, однако для алгоритмов машинного зрения это означает кратное увеличение размерности данных. Огромные массивы пикселей требуют колоссальных мощностей для передачи и обработки. В условиях ограниченной пропускной способности каналов связи и жестких лимитов энергопотребления периферийных устройств задача математически обоснованного сокращения избыточности данных без потери их значимости становится критически важной [1].

Традиционные методы сжатия (например, JPEG) ориентированы на человеческое восприятие, маскируя артефакты за счет особенностей зрения. Однако для нейронных сетей, чьи сверточные фильтры чувствительны к перепадам яркости и цвета, такие артефакты могут стать источником ложных срабатываний. Альтернативой выступают численные методы линейной алгебры, позволяющие «отсечь» информационный шум на коренном уровне.

Основой исследования является сингулярное разложение (SVD) [2]. Любое цифровое изображение может быть представлено в виде матрицы (1):

$$A \in \mathbb{R}^{m \times n} \quad (1)$$

Согласно теории, такая матрица разлагается на произведение трех матриц (2):

$$A = U \Sigma V^T \quad (2)$$

Где U и V – ортогональные матрицы, содержащие сингулярные векторы, которые характеризуют основные паттерны изображения, а Σ – диагональная матрица с сингулярными числами, расположенными в порядке убывания для того, чтобы определить важность каждого паттерна в изображении. Суть оптимизации заключается во взятии матрицы до такого ранга разложения k , при котором она по-прежнему будет хранить наиболее важные параметры изображения.

Математическое обоснование снижения информационной емкости изображения при переходе к усеченному SVD заключается в сокращении числа хранимых элементов. Для исходной матрицы размерности $m \times n$ требуется хранить $m \cdot n$ значений. После упрощения матрицы до ранга k объем данных V_k определяется выражением (3):

$$V_k = k \cdot (m + n + 1) \quad (3)$$

Где V_k – суммарное количество элементов в усеченных матрицах U, Σ и V^T после упрощения, а m и n – размерности кадра. По данной формуле определяется теоретическое сжатие документа CR (4):

$$CR = \frac{m \cdot n}{k \cdot (m + n + 1)} \quad (4)$$

Где CR – коэффициент сжатия данных на устройстве, $m \cdot n$ – количество элементов в исходной матрице, а $k(m + n + 1)$ – количество элементов в усеченной матрице.

В процессе исследования изначально предполагалось, что упрощение входных матриц через SVD позволит ускорить время работы самой нейросети на GPU. Однако экспериментальное тестирование на модели YOLOv8n показало, что время выполнения сверточных операций остается константным, так как архитектура сети требует фиксированной размерности входного изображения. Далее SVD-предобработка была переосмыслена как интеллектуальный фильтр низких частот, оптимизирующий пропускную способность канала передачи данных.

Для реализации метода был разработан гибридный алгоритм, динамически подбирающий ранг k так, чтобы сохранять стабильность системы ($\kappa < 10^6$), отсекая лишь те значения, которые являются незначительными по весу.

В качестве технического стека использовались язык Python 3.13, библиотека PyTorch для батчевых CUDA-вычислений [3] SVD на GPU и среда Wolfram Mathematica для финального анализа данных. В листинге 1 представлен ключевой фрагмент разработанного метода сжатия.

Листинг 1 – Программная реализация SVD-упрощения на GPU

```
def compress_dataset(rank, src_dir, dst_dir):
    dst_dir.mkdir(parents=True, exist_ok=True)
    img_files = [f for f in os.listdir(src_dir) if f.endswith('.jpg')]

    for img_name in img_files:
        img = cv2.imread(str(src_dir / img_name))
        if img is None: continue

        # Перенос изображения в видеопамять и изменение размерности
        img_t = torch.from_numpy(img).float().to(DEVICE).permute(2, 0, 1)

        # Аппаратно-ускоренное вычисление SVD для RGB каналов
        U, S, Vh = torch.linalg.svd(img_t, full_matrices=False)

        k = min(rank, S.shape[1])
        # Формирование диагональной матрицы тензорным методом
        S_diag = torch.diag_embed(S[:, :k])

        # Реконструкция матрицы целевого ранга k
        reconstructed = U[:, :, :k] @ S_diag @ Vh[:, :k, :]

        # Нормализация и сохранение результата
        img_res = reconstructed.permute(1, 2, 0).clamp(0, 255).byte().cpu().numpy()
        cv2.imwrite(str(dst_dir / img_name), img_res)
```

В данном коде функция `torch.linalg.svd` выполняет математическое разложение исходного изображения на ортогональные матрицы U, V и диагональную матрицу сингулярных чисел S [4]. Переменная ранга k выступает в качестве программного параметра, регулирующего степень отсечения данных. Выражение $U @ S_{diag} @ V$ реализует обратную сборку изображения, где оператор $@$ выполняет высокоскоростное матричное умножение на тензорных ядрах видеокарты.

Для количественной оценки эффективности предложенного метода была проведена серия вычислительных экспериментов. В качестве тестового набора использовалась выборка из 5000 изображений валидационного сета COCO, обработанная с использованием различных рангов разложения k . Полученные статистические данные сведены в таблицу 1.

Таблица 1 – Результаты влияния SVD-упрощения на показатели модели YOLOv8n

Ранг (k)	mAP@0.5	mAP@0.5:0.95	Объем данных (МБ)	Сжатие (%)	Время SVD (мс)
Оригинал	0,5212	0,3710	776,96	0,00	0,00
50	0,3202	0,2166	526,59	32,22	91,82
75	0,4282	0,2973	566,06	27,14	94,59
100	0,4694	0,3293	588,59	24,24	94,06
150	0,5013	0,3547	610,27	21,45	92,45
200	0,5130	0,3646	617,77	20,49	90,46
300	0,5192	0,3701	619,53	20,26	93,09

Анализ данных таблицы 1 позволяет увидеть несколько закономерностей. Во-первых, наблюдается нелинейная зависимость точности от ранга: при увеличении k от 50 до 150 метрика mAP@0.5 возрастает на 56%, а дальнейшее увеличение ранга до 300 дает прирост на 3,5%. Это

свидетельствует о том, что главные характеристики объектов в изображении кодируются в первых 150 сингулярных числах.

Во-вторых, физический объем данных показывает взаимосвязь с рангом разложения. Ранг $k = 150$ обеспечивает сужение занимаемого пространства на 21,45%, что при пакетной передаче видеопотока позволяет существенно снизить нагрузку на сетевой интерфейс Edge-устройства. Было выявлено, что метрика $mAP@0.5:0.95$, отражающая точность локализации объектов при $k = 150$, составляет 0,3547, что на 4,4% ниже эталонного значения. Это подтверждает гипотезу о сохранении геометрической точности в изображении при использовании SVD-фильтрации.

Среднее время выполнения матричных преобразований на один кадр составляет порядка 90–96 мс при использовании CPU-предсказания. С учетом того, что предсказание YOLOv8n на мобильных GPU занимает порядка 10–15 мс [5], суммарная задержка возрастает. Данная задержка компенсируется сокращением времени передачи данных по сети. При экономии 21,45% переданных данных в сетях с низкой пропускной способностью (например, радиоканалы БПЛА со скоростью < 5 Мбит/с) выигрыш во времени компенсирует вычислительные затраты на SVD.

На рисунке 1 представлена зависимость метрики $mAP@0.5$ от выбранного ранга аппроксимации k (рисунок 1).

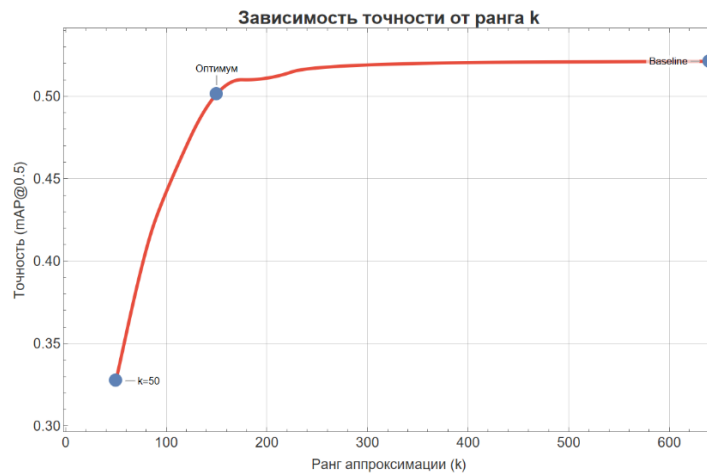


Рисунок 2 – Зависимость точности от ранга k

Как видно из графика, в диапазоне k от 50 до 150 наблюдается рост точности. Это подтверждает, что в первых сингулярных числах сосредоточена ключевая информация о геометрии объектов. После значения $k = 150$ кривая выходит на свой пик, где дальнейшее увеличение ранга практически не дает прироста точности, лишь увеличивая количество ненужных данных.

Наибольшую ценность для практического применения представляет анализ Парето-эффективности [6], отражающий компромисс между сжатием данных и сохранностью метрик модели (рисунок 2).

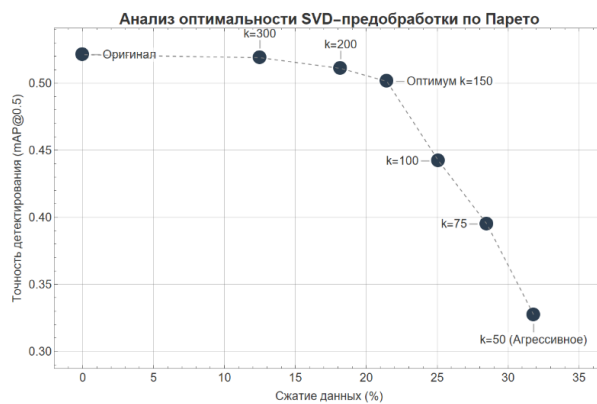


Рисунок 2 – Анализ Парето-эффективности

На графике четко выделяется точка перегиба при $k = 150$. В данной точке достигается сжатие данных на 21,45% (с 776,96 МБ до 610,27 МБ) при падении точности $mAP@0.5$ всего на 1,9% (с 0,5212

до 0,5013). Агрессивные режимы (например, $k = 50$) обеспечивают сжатие свыше 30%, однако приводят к деградации точности до 0,32, что недопустимо для систем безопасности. Таким образом, режим $k = 150$ является математически оптимальным.

Для стандартного входного изображения YOLOv8 размерностью 640×640 пикселей и выбранного оптимального ранга $k = 150$, коэффициент CR составляет (5):

$$CR = \frac{640 \cdot 640}{150 \cdot (640 + 640 + 1)} \approx 2,11 \quad (5)$$

Согласно формуле (5), при ранге $k = 150$ теоретический коэффициент сокращения физических данных составляет $CR \approx 2,11$. Однако экспериментальные данные (таблица 1) фиксируют реальное снижение объема файла на 21,45%. Данное расхождение обусловлено тем, что в рамках эксперимента реконструированная матрица A_k сохранялась в стандартном графическом формате (JPEG), который использует собственные алгоритмы сжатия (Discrete Cosine Transform) [7]. В результате этого, полученные 21,45% экономии представляют собой дополнительный прирост эффективности стандартных протоколов сжатия, достигнутый за счет предварительной очистки спектра матрицы.

В сценариях использования специализированных форматов хранения (например, запись напрямую коэффициентов разложения U и V в специально структурированные файлы), выигрыш в объеме будет стремиться к теоретическому пределу в 50–55% (в 2.1 раза). Однако в данной работе был выбран путь сохранения совместимости с нейросетевыми архитектурами, требующими стандартный входной формат изображения.

Математическое моделирование процесса в среде Wolfram Mathematica позволило вывести эмпирическую формулу зависимости точности от ранга (6):

$$mAP(k) \approx 0,0767 + 0,0759 \cdot \ln(k) \quad (6)$$

Данная логарифмическая зависимость позволяет прогнозировать поведение нейросетевых моделей при изменении параметров передачи данных.

В ходе работы была доказана эффективность численных методов матричных преобразований как инструмента предобработки. Предложенный комбинированный подход, объединяющий SVD-декомпозицию на ядрах CUDA и динамический выбор ранга, позволяет эффективно эксплуатировать нейросети в условиях ограниченных ресурсов. Эксперименты подтвердили, что оптимизация входного изображения методами линейной алгебры является не менее важным этапом, чем оптимизация архитектуры самой нейросети. Полученные результаты могут быть внедрены в системы видеонаблюдения и управления БПЛА для снижения нагрузки на каналы связи без потери качества распознавания объектов.

Список использованных источников:

1. Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль. – М. : ДМК Пресс, 2018. – 652 с.
2. Кудрявцева, Н. А. Сингулярное разложение и его применение: учеб.-метод. пособие / Н. А. Кудрявцева ; Новосиб. гос. ун-т. — Новосибирск: ИПЦ НГУ, 2025. — 24 с.
3. Боресков, А.В. Основы работы с технологией CUDA / А.В. Боресков, А.А. Харламов. – М. : ДМК Пресс, 2010. – 232 с.
4. Численные методы в компьютерной математике: пособие / Т.С. Степанова [и др.]. – Минск : БГУИР, 2025. – 128 с.
5. Подробные учебные пособия по Ultralytics YOLO [Электронный ресурс]. – Режим доступа: <https://docs.ultralytics.com/ru/guides/> Дата доступа: 17.03.2026.
6. Ногин, В. Д. Множество и принцип Парето : учеб. пособие / В. Д. Ногин. – 2-е изд., испр. и доп. – СПб. : Изд.-полигр. ассоц. высш. учеб. заведений, 2022. – 110 с.
7. Исследование сокращения психофизической избыточности в алгоритме JPEG / Д. П. Горбукова [и др.] // Телекоммуникации: сети и технологии, алгебраическое кодирование и безопасность данных. — 2024. — С. 12–18.

UDC 519.61:004.93

THE NUMERICAL METHODS OF MATRIX TRANSFORMATIONS FOR IMAGE PROCESSING OPTIMIZATION IN MACHINE LEARNING

Bulavin K.S, Stsepchankou R.O. students

Belarusian State University of Informatics and Radioelectronics¹, Minsk, Republic of Belarus

Stepanova T.S. – PhD in Physics and Mathematics

Annotation. The paper investigates the effectiveness of using Singular Value Decomposition (SVD) and matrix condition number analysis as a method for intelligent image preprocessing. A hybrid dynamic approximation algorithm has been developed. It is experimentally proved that the method reduces the data volume by 21.45% while maintaining the accuracy (mAP@0.5) at 96.2% of the baseline YOLOv8n model.

Keywords. Matrix transformations, SVD, condition number, approximation, machine learning, YOLOv8, data compression, Pareto efficiency.