

РАЗРАБОТКА АЛГОРИТМА КЛАССИФИКАЦИИ И СЕГМЕНТАЦИИ СЭМ-ИЗОБРАЖЕНИЙ СФЕРИЧЕСКИХ НАНОЧАСТИЦ НА ПОВЕРХНОСТИ БИОСОВМЕСТИМЫХ МАТЕРИАЛОВ

Д.И. Рябкин, к.ф.-м.н., доцент, НИУ МИЭТ, RyabkinDI@yandex.ru;

Д.Ф. Вольхин, НИУ ВШЭ, deneal123@mail.ru.

УДК 004.855.5

Аннотация. Представленный авторами алгоритм объединяет в себе различные задачи компьютерного зрения, позволяющие автоматизировать измерение диаметра Ферета сферических наночастиц на СЭМ-изображениях, снижая временные затраты и количество субъективных ошибок. В работе исследованы современные методы генерации синтетических изображений и автоматизации разметки данных, которые решают проблемы переобучения на малых наборах данных и снижают экономические затраты.

Ключевые слова: компьютерное зрение; синтетические данные; СЭМ; наночастицы; диаметр Ферета; интерфейс.

DEVELOPMENT OF AN ALGORITHM FOR CLASSIFICATION AND SEGMENTATION OF SEM IMAGES OF SPHERICAL NANOPARTICLES ON THE SURFACE OF BIOCOMPATIBLE MATERIALS

D.I. Ryabkin, PhD, Associate Professor, NRU MIET;

D.F. Volkhin, NRU HSE.

Annotation. The algorithm combines various computer vision tasks to automate the measurement of the Feret diameter of spherical nanoparticles in SEM images, reducing time costs and the number of subjective errors. The study explored modern methods for generating synthetic images and automating data annotation, which addressing the problems of overfitting on small datasets and reducing economic costs.

Keywords: computer vision; synthetic data; SEM; nanoparticles; Feret diameter; interface.

Введение

Автоматизация анализа наноматериалов, в частности, измерение диаметра сферических наночастиц, представляет собой актуальную задачу в области нанотехнологий. Ручные методы измерения наночастиц по изображениям сканирующей электронной микроскопии (СЭМ) требуют значительных временных затрат, подвержены субъективным ошибкам и ограничивают возможности проведения масштабных исследований. Точные данные о размере наночастиц, такие как диаметр Ферета, критически важны для оценки их физических свойств, что, в свою очередь, оказывает влияние на их применение в медицине, электронике и других направлениях. Это обуславливает необходимость разработки методов, которые позволили бы минимизировать человеческий фактор и ускорить анализ изображений.

Для решения данной проблемы в рамках исследования предлагается алгоритм [1], который автоматизирует процесс измерения сферических наночастиц. В основе решения – компьютерное зрение, которое позволяет отказаться от трудоемкой ручной обработки. Одной из ключевых задач являлось формирование синтетических наборов данных для тренировки нейронных сетей, так как ограниченность доступных экспериментальных данных затрудняет

разработку надежных алгоритмов классификации, детекции и сегментации. Генерация синтетических данных с использованием современных методов машинного обучения позволяет увеличить объем тренировочных данных, что способствует улучшению обобщающей способности моделей и позволяет решать проблему переобучения.

Алгоритм включает этапы фильтрации изображений без частиц, детекции объектов интереса, сегментации для получения контура, а также распознавание текста, помогающее определить масштаб, который используется для получения реального размера наночастиц. Кроме того, разработанный инструмент предполагает наличие интуитивно понятного интерфейса, что делает его доступным для специалистов различных направлений.

Экспериментальные данные для исследования были получены в лабораторных условиях с использованием установки *HELIOS NANOLAB 650*. СЭМ-изображения готовились двумя методами: в первом случае наночастицы наносились методом напыления на кремниевые подложки, во втором – порошок наночастиц распределялся на углеродный скотч, который затем приклеивался к кремниевым подложкам. Итоговый набор данных включает 422 изображения, представляющие наночастицы на различных поверхностях.

Разметка экспериментального набора данных проводилась вручную с целью выделения трех классов: 1) *carbonetape (cnt)* – изображения, где наночастицы нанесены на углеродный скотч; 2) *siliconsubstrate (sss)* – изображения, полученные методом напыления наночастиц на кремниевые подложки; 3) *nonparticles (nop)* – изображения, не содержащие сферических наночастиц, также вручную определялись ограничивающие боксы.

Маски сегментации получены при помощи инструмента [2] через указание *key points*. Точность разметки данных имеет ключевое значение для последующего обучения и тестирования алгоритмов – в процессе разметки возникали конфликтные ситуации, при которых было трудно однозначно присвоить класс или выделить частицу. Исходя из этого, результаты метрик нужно интерпретировать через призму возможных неоднозначностей разметки.

Таким образом, данное исследование направлено на разработку комплексного инструмента для анализа изображений наночастиц, который, благодаря использованию синтетических данных и современных методов машинного обучения, способен повысить скорость и надежность оценки размеров наночастиц, а также минимизировать влияние субъективных факторов, свойственных ручному анализу.

Материалы и методы

Рассмотрены шесть методов формирования синтетических наборов данных: с использованием программного обеспечения *Blender* [3] и *Automatic (Stable Diffusion)* [4], методами дообучения *DreamBooth* [5] и *LoRA* [6] модели *Stable Diffusion*, генеративно-состязательной (*GAN*) [7] и диффузионно-трансформерной (*DiT*) [8] нейронными сетями. Примеры генерации изображений: А – *Blender*; Б – *Automatic (SD)*; В – *DreamBooth (SD)*; Г – *LoRA (SD)*; Д – *GAN*; Е – *Transformer (DiT)* – для каждого метода представлены на рис. 1.

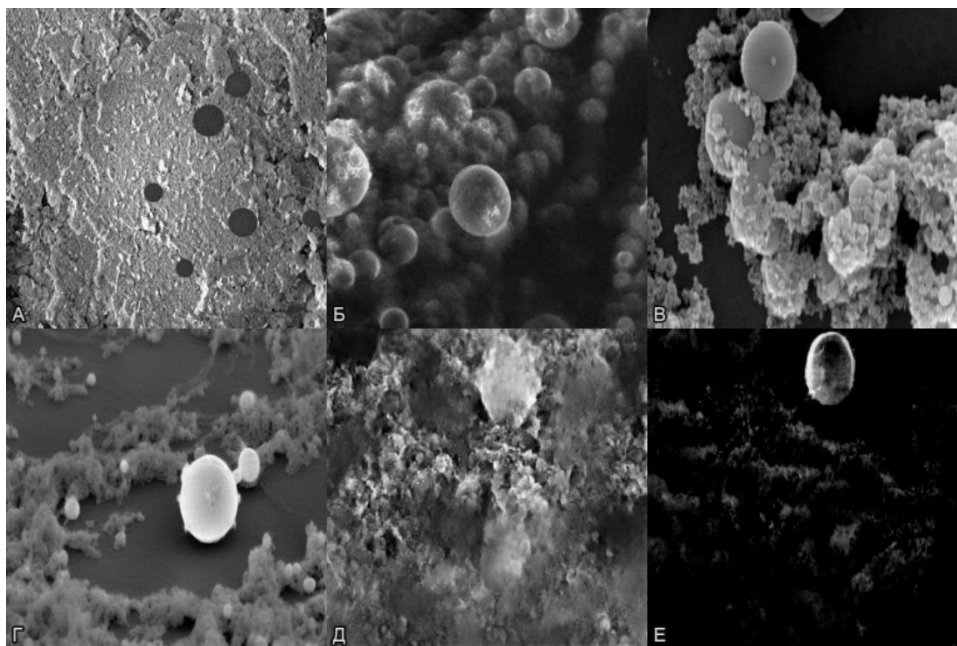


Рисунок 1

Для задачи детекции и сегментации использовалась автоматизированная разметка с помощью детектора *Grounding-Dino* [9] – для выделения ограничивающих рамок и разметка с помощью семейства моделей *Segment Anything* [10] – для сегментирования сфер соответственно. Модели обладают повышенной обобщающей способностью, но не годятся для специализированных задач. Ко всему прочему, эти модели обладают низкой скоростью обработки из-за своих размеров.

Наночастицы являются сложными объектами для задачи детекции и сегментации изображений по нескольким причинам:

Во-первых, они могут находиться на поверхностях с разнообразными текстурами, что создает сложности при выделении их на фоне. Это означает, что модели должны адаптироваться к различным текстурным характеристикам фона для точного выделения наночастиц.

Во-вторых, на поверхности наночастиц возможен шум, который может быть ложно интерпретирован моделями детекции как часть объекта или дополнительные объекты, что ведет к неточным результатам.

И, наконец, форма наночастиц далеко не всегда идеально сферическая, что отличается от объектов, на которых могут быть обучены некоторые обобщенные модели.

Для решения задачи классификации использовались современные архитектуры, «голова» классификации для всех моделей была одинаковая. Она состояла из слоя глобального среднего пулинга, который усредняет признаки в выходном тензоре по всем пространственным измерениям, двух полносвязных слоев с 4048 нейронами и выходного слоя с тремя нейронами. Для оценки качества обучения моделей использовалась функция потерь категориальная кросс-энтропия, а в качестве основной метрики выбрана *accuracy*, так как классы сбалансированы. В табл. 1 представлены метрики после 31 эпохи обучения.

Таблица 1.

Название модели	Метрики качества обучения				
	Коэффициент корреляции Мэтьюса (<i>MCC</i>)	Истинность (<i>accuracy</i>)	Точность (<i>precision</i>)	Полнота (<i>recall</i>)	F1 мера (<i>f1-score</i>)
<i>VGG16</i>	0,606±0,004	0,765±0,005	0,830±0,002	0,671±0,004	0,701±0,005
<i>DenseNet201</i>	0,327±0,003	0,627±0,008	0,665±0,004	0,477±0,001	0,485±0,003
<i>InceptionV3</i>	0,543±0,007	0,726±0,008	0,766±0,007	0,647±0,003	0,677±0,005
<i>ResNet50</i>	0,582±0,004	0,752±0,007	0,869±0,003	0,643±0,002	0,691±0,004
<i>Xception</i>	0,571±0,006	0,744±0,008	0,835±0,003	0,652±0,004	0,695±0,005
<i>MobileNet</i>	0,566±0,004	0,737±0,006	0,863±0,003	0,618±0,003	0,663±0,004
<i>EfficientNetB3</i>	0,689±0,005	0,817±0,005	0,831±0,003	0,765±0,006	0,789±0,005
<i>EfficientNetB2</i>	0,600±0,006	0,751±0,009	0,850±0,005	0,664±0,003	0,701±0,005
<i>EfficientNetB4</i>	0,615±0,004	0,779±0,004	0,836±0,002	0,682±0,003	0,727±0,004
<i>inceptionResNetV2</i>	0,414±0,007	0,622±0,006	0,679±0,003	0,628±0,006	0,620±0,005
<i>NasNetMobile</i>	0,555±0,007	0,721±0,009	0,861±0,003	0,614±0,005	0,655±0,008
<i>ResNet101</i>	0,313±0,005	0,541±0,005	0,609±0,013	0,533±0,004	0,496±0,004

Процесс сегментации частиц был реализован в два последовательных этапа. На первом этапе использовалась модель *YOLOv8* [11] для выделения областей интереса – прямоугольных фрагментов, в которых предположительно находятся частицы. На основе координат этих фрагментов вырезались соответствующие части исходного изображения и масштабировались до размера, соответствующего входному слою модели сегментации *Unet* [12]. Метрики качества обучения модели детекции наночастиц № 1 представлены в табл. 2, где приведены значения метрик для последних пяти из 49 эпох. Модели № 2-4 применялись для детекции значения масштаба из колонтитула СЭМ.

Таблица 2.

Номер модели №	Название метрик			
	Точность (<i>precision</i>)	Полнота (<i>recall</i>)	Средняя точность при <i>IoU</i> =0,5 (<i>mAP50</i>)	Средняя точность при <i>IoU</i> =(0,5-0,95) (<i>mAP50-95</i>)
1	0,607±0,005	0,128±0,003	0,377±0,009	0,142±0,002
2	0,999±0,001	0,999±0,001	0,995±0,002	0,885±0,008
3	0,999±0,001	0,999±0,001	0,995±0,005	0,995±0,003
4	0,999±0,001	0,999±0,001	0,995±0,002	0,995±0,001

Далее для распознавания масштаба, указанного в колонтитуле изображения, была применена модель *ABINet* [13], оптимизированная для задач оптического распознавания текста (*OCR*). Эта модель продемонстрировала высокую точность распознавания, позволяя точно идентифицировать размерный масштаб, что является критически важным для корректного расчета параметров частиц. Метрики качества работы *ABINet* на тестовой выборке приведены в табл. 3.

Таблица 3.

Название модели	Название метрик			
	F1-мера (<i>f1-score</i>)	Истинность (<i>accuracy</i>)	Точность (<i>precision</i>)	Полнота (<i>recall</i>)
<i>ABINet</i>	0,973±0,002	0,952±0,003	1,000±0,000	0,952±0,003

На следующем этапе использовались различные модификации модели *Unet* для выделения контуров частиц непосредственно из областей интереса, определенных на первом этапе. Такой подход позволил добиться точной локализации и качественной сегментации контуров частиц. Обучение модели сегментации проводилось в течение 31 эпохи, и оценка его эффективности проводилась с использованием функции потерь на основе бинарной кросс-энтропии, а основным показателем качества работы модели являлось отношение пересечения к объединению (*IoU*). Метрики качества модели для данного этапа приведены в табл. 4.

Таблица 4.

Название модели	Метрики качества обучения				
	Коэффициент корреляции Мэтьюса (<i>MCC</i>)	Отношение пересечения к объединению (<i>IoU</i>)	Истинность маски (<i>accuracy</i>)	Площадь под <i>ROC</i> -кривой	Площадь под <i>PR</i> -кривой
<i>SegHRnet</i>	0,17±0,06	0,08±0,03	0,56±0,03	0,56±0,03	0,16±0,06
<i>UnetPlus</i>	0,16±0,05	0,10±0,02	0,24±0,06	0,57±0,03	0,19±0,06
<i>PrismaNet</i>	0,24±0,07	0,16±0,05	0,40±0,07	0,60±0,03	0,24±0,07
<i>PrismaMattingNet</i>	0,21±0,07	0,14±0,04	0,29±0,07	0,59±0,03	0,22±0,07
<i>FastDeepMatting</i>	0,16±0,05	0,12±0,03	0,23±0,05	0,58±0,03	0,21±0,05
<i>UnetZF</i>	0,27±0,07	0,16±0,04	0,34±0,07	0,63±0,03	0,27±0,07
<i>Vnet</i>	0,21±0,06	0,14±0,04	0,38±0,06	0,58±0,03	0,22±0,06
<i>BCDUnet</i>	0,23±0,07	0,15±0,04	0,40±0,06	0,59±0,03	0,24±0,07
<i>Unet</i>	0,26±0,06	0,13±0,03	0,46±0,05	0,60±0,03	0,26±0,06
<i>ResAttUnet</i>	0,12±0,05	0,07±0,03	0,38±0,06	0,54±0,02	0,11±0,05
<i>UefficientNetB3</i>	0,37±0,07	0,22±0,05	0,48±0,07	0,67±0,04	0,36±0,07

Применение многоступенчатого подхода, включающего детекцию и сегментацию, продемонстрировало высокую надежность, позволяя достигать результатов точнее, чем при использовании просто детекции.

Для вычисления диаметра Ферета [14] использован алгоритм, основанный на анализе полигона, полученного из бинарной маски частицы на изображении. Полигон представлен в виде списка точек, описывающих контур частицы. Алгоритм проходит по всем парам точек в диапазоне 180 градусов полигона и вычисляет расстояние между ними. Для этого используется формула расстояния между двумя точками на плоскости. Диаметр Ферета определяется как среднее арифметическое значений между медианой и наибольшим расстоянием. Такой подход обоснован тем, что модель может выделить только часть сферы, и это приводит к выбросам. Итоговый диаметр Ферета получен путем умножения рассчитанного значения на масштаб изображения, выраженный в нанометрах на пиксель.

На рис. 2 представлен пример работы объединенного алгоритма: А – 7 частиц; Б – 7 частиц; В – 12 частиц; Г – 21 частица; Д – 12 частиц; Е – 13 частиц. В левом верхнем углу каждой рамки указан номер частицы, снизу каждой рамки расположено значение рассчитанного диаметра в микрометрах для соответствующей частицы.

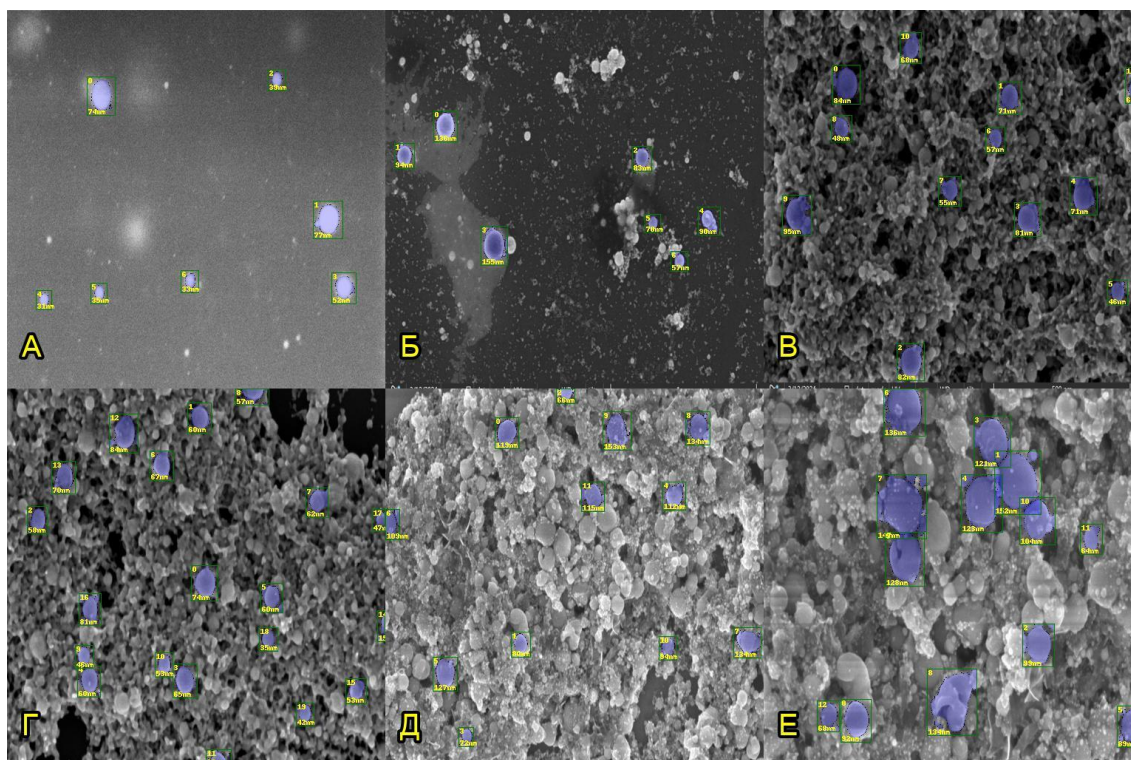


Рисунок 2

Результаты

Методы формирования синтетических наборов данных оценены с помощью таких метрик как *FID* (дистанция Фреше) [15], *SSIM* (структурное сходство), *PSNR* (пиковое соотношение сигнал/шум), *COSIM* (косинусное сходство) и *EMD* (расстояние Вассерштейна), результаты оценки которых приведены в табл. 5.

Таблица 5.

Метод	Метрики оценок результатов генерации изображений				
	<i>FID</i>	<i>SSIM</i> × 10 ⁻⁴	<i>PSNR</i> × 10 ⁻¹	<i>COSIM</i> × 10 ⁻³	<i>EMD</i> × 10 ⁻³
<i>Blender</i>	317±3	9941±2	613±1	892±3	73±5
<i>Automatic</i>	271±2	9931±2	607±1	862±4	56±4
<i>LoRA</i>	157±1	9932±4	609±2	875±6	77±4
<i>DreamBooth</i>	128±2	9925±3	605±2	838±6	65±5
<i>GAN</i>	221±2	9924±1	603±1	840±3	44±3
<i>DiT</i>	237±2	9781±7	564±1	290±4	343±6

Метод *DreamBooth* дообучения диффузионной-нейронной сети *Stable Diffusion* показал себя наилучшим образом в сравнении с другими методами, он использовался для формирования основной обучающей выборки – набора данных 4353 изображений с метками классов, боксами и бинарными масками. Также стоит отметить, что методы *GAN* и *DiT* имеют высокий потенциал, но в связи с дорогостоящими вычислениями и нехваткой разнообразия в данных, модели на основе этих архитектур не были обучены достаточно долго, что оказало влияние на результаты генераций и впоследствии метрик.

В конечном варианте алгоритма использовались модель классификации *EfficientNetV3* [16], модель сегментации *UEfficientNet* [17], модели *Yolo* (№ 1-4)

для детекции областей интереса и модель распознавания текста *ABINet*. На рис. 3 представлена блок-схема объединенного алгоритма.

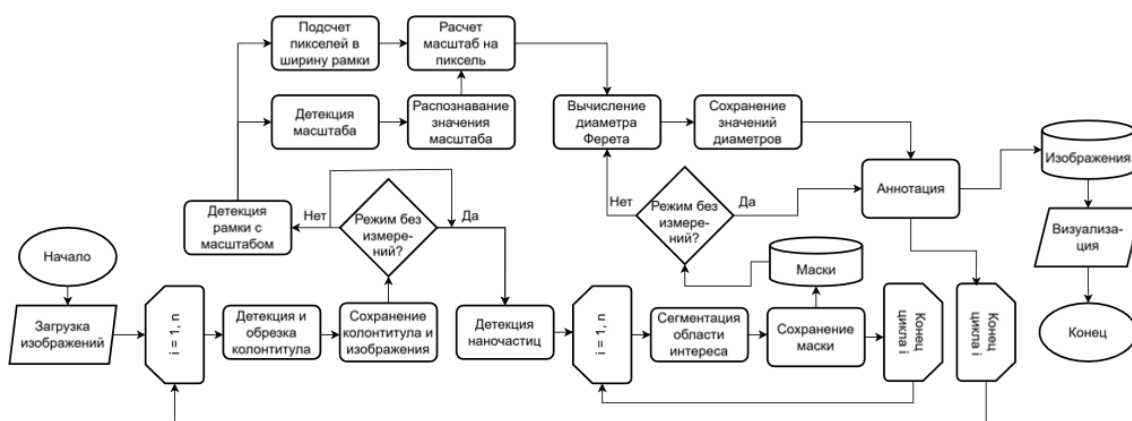


Рисунок 3

Для удобства использования и визуализации результатов работы алгоритма разработан пользовательский интерфейс на основе библиотеки *Streamlit* [18]. Веб-интерфейс (рис. 4) позволяет пользователю взаимодействовать с алгоритмом без необходимости написания кода, достаточно просто загрузить одно или сразу группу изображений.

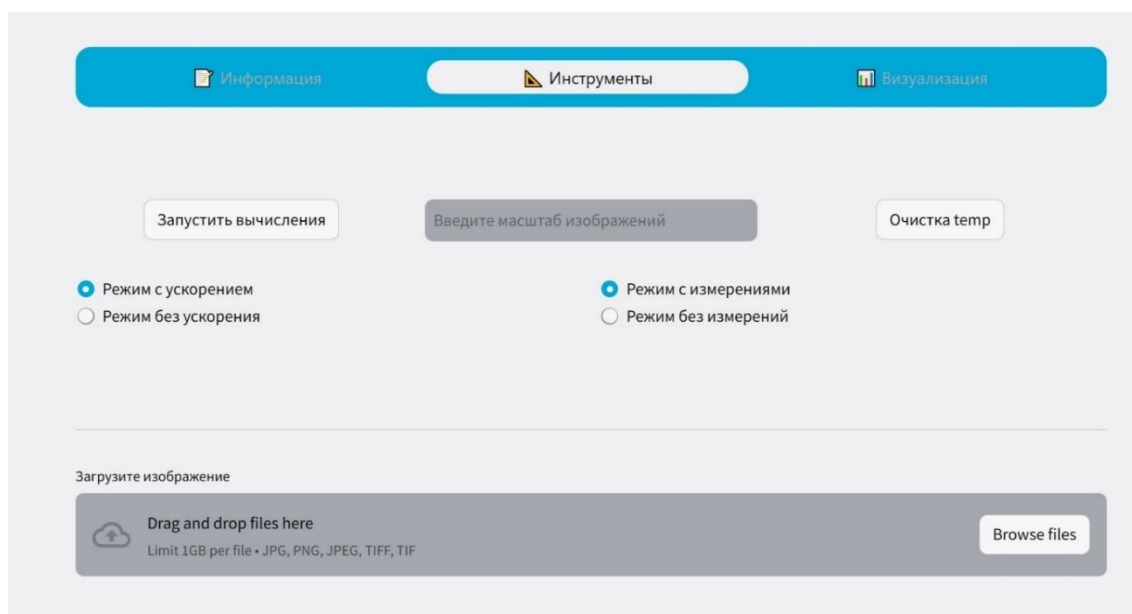


Рисунок 4

Для оценки эффективности разработанного алгоритма вычисления размерных характеристик сферических наночастиц было проведено сравнение с традиционным методом измерения, основанным на ручном измерении с использованием электронной линейки.

Результаты сравнительного анализа показали, что точность алгоритма сопоставима с точностью ручного метода измерения, что отражено на гистограмме распределения среднего значения диаметра для каждой наночастицы по отдельности (рис. 5), а также на гистограмме распределения среднего значения диаметра нескольких наночастиц (от 2 до 10) для каждого изображения по отдельности (рис. 6). Это говорит о том, что разработанный алгоритм способен обеспечить сопоставимую точность при вычислении размеров наночастиц.

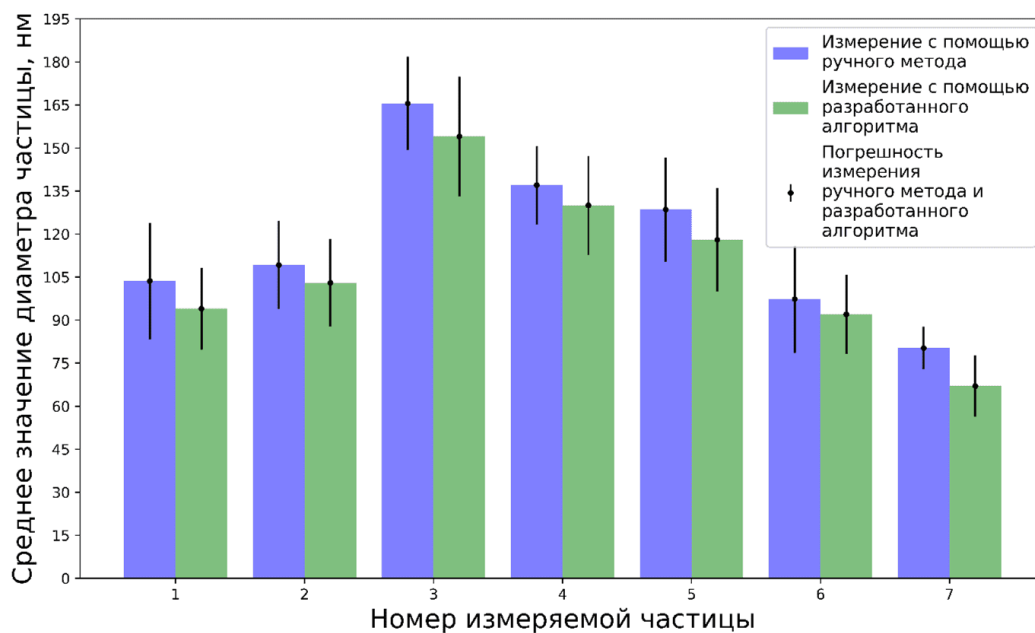


Рисунок 5

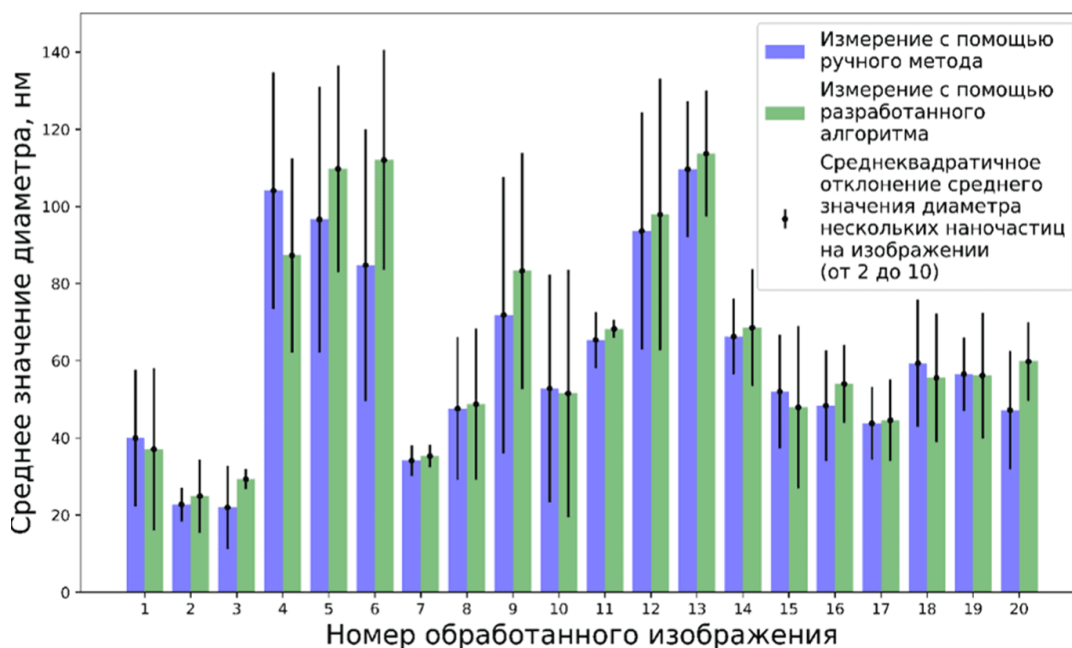


Рисунок 6

Однако одним из ключевых преимуществ алгоритма является его высокая скорость работы. Как видно из табл. 6, алгоритм работает значительно быстрее по сравнению с ручным методом измерения. Например, для анализа 10 изображений с несколькими частицами, время выполнения алгоритма составило всего $14,5 \pm 5,7$ секунд, в то время как ручной метод занял в среднем $2116,7 \pm 1325,1$ секунд.

Таблица 6.

Метод	Затраченное время (сек)	
	1 частица	10 изображений
Ручной	48 ± 10	$2116,7 \pm 1325,1$
Алгоритм	-	$14,5 \pm 5,7$

Заключение

В ходе исследования была проведена оценка эффективности передовых методов дообучения моделей глубокого обучения для генерации синтетических изображений, таких как *DreamBooth* и *LoRA*, в приложении к модели *Stable Diffusion*. Данные методы продемонстрировали свою эффективность, хотя они и обладают некоторыми ограничениями (качество генерации и требования к вычислительным ресурсам), они взаимоисключающие. Автоматизация процесса разметки данных с использованием моделей *Grounding Dino* и *Segment Anything* ускорила подготовку обучающих наборов данных, что особенно важно для создания высокоточных систем компьютерного зрения.

Результаты экспериментов с моделями *EfficientNetB3* и *UefficientNetB3* подчеркнули необходимость внимательного подхода к выбору архитектуры сети для специфичных задач классификации и сегментации изображений. В процессе работы также были рассмотрены модели *YOLO* и *ABINet*, которые доказали свою эффективность в задачах детекции и распознавания текста.

Сравнение разработанного алгоритма [19] с традиционным методом измерения размеров наночастиц показало его значительные преимущества в скорости и сопоставимого значения точности, что открывает перспективы для ускорения аналитических процессов в научных исследованиях. Применение машинного обучения и синтетических данных позволило сократить временные и финансовые затраты, что крайне важно при проведении исследований.

Литература

1. Вольхин Д.Ф. Разработка алгоритма классификации и сегментации сферических наночастиц на поверхности биосовместимых материалов // В книге: Микроэлектроника и информатика. Сборник материалов (тезисов) XXXI международной конференции, 2024. – С. 14.
2. URL <https://github.com/uncbiag/SimpleClick> (дата обращения – май 2024 г.).
3. Kharin A.Yu. Deep learning for scanning electron microscopy: Synthetic data for the nanoparticle's detection // *Ultramicroscopy*, 2020. – V. 219. – № 113125.
4. Nguyen L.X., Sone Aung P., Le H.Q., Park S. B., Hong C. S. A new chapter for Medical Image Generation: The Stable Diffusion Method // *ICOIN*, 2023. – P. 483-486.
5. Nataniel Ruiz, Yannzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, Kfir Aberman. DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subjects-Driven Generation // *CVPR*, 2023. – P. 22500-22510.
6. Gu Yuchao and Wang, Xintao and Wu, Jay Zhangjie and Shi, Yujun and Chen, Yunpeng and Fan, Zihan and XIAO, WUYOU and Zhao, Rui and Chang, Shuning and Wu, Weijia and Ge, Yixiao and Shan, Ying and Shou, Mike Zheng. Mix-of-Show: Decentralized Low-Rank Adaptation for Multi-Concept Customization of Diffusion Models // *Advances in Neural Information Processing Systems*, 2023. – V. 36. – P. 15890-15902.
7. Nandhini Abirami R. and Durai Raj Vincent, P. M. and Srinivasan, Kathiravan and Tariq, Usman and Chang, Chuan-Yu. Deep CNN and Deep GAN in Computational Visual Perception-Driven Image Analysis // *Complexity*, 2021. – V. 2021. – №. 1. – P. 5541134.
8. William Peebles, Saining Xie. Scalable Diffusion Models with Transformers // *ICCV*, 2023. – P. 4195-4205.
9. Son Jinhwan and Jung, Heechul. Teacher-Student Model Using Grounding DINO and You Only Look Once for Multi-Sensor-Based Object Detection // *Applied Sciences*, 2024. – V. 14. – № 6.

10. Kirillov A., Mintun E., Ravi Nikhila and Mao, Hanzi and Rolland, Chloe and Gustafson, Laura and Xiao, Tete and Whitehead, Spencer and Berg, Alexander C. and Lo, Wan-Yen. Segment Anything // ICCV, 2023. – P. 4015-4026.
11. Palanivel N. The Art of YOLOv8 Algorithm in Cancer Diagnosis using Medical Imaging // ICSCAN, 2023. – P. 1-6. DOI: 10.1109/ICSCAN58655.2023.10395046.
12. Ronneberger O.F., Philipp B.T. U-Net: Convolutional Networks for Biomedical Image Segmentation // MICCAI, 2015. – P. 234-241.
13. Fang Shancheng and Xie, Hongtao and Wang, Yuxin and Mao, Zhendong and Zhang, Yongdong. Read Like Humans: Autonomous, Bidirectional and Iterative Language Modeling for Scene Text Recognition // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021. – P. 7098-7107.
14. Dražić S., Sladoje N., Lindblad J. Estimation of Feret's diameter from pixel coverage representation of a shape // Pattern Recognition Letters, 2016. – V. 80. – P. 37-45.
15. Heusel M., Ramsauer H., Unterhiner T., Nessler B., Hochreiter S. GANs trained by a two time-scale update rule converge to a local nash equilibrium // NIPS17, 2017. – V. 17. – № 12. – P. 6629-6640.
16. Koonce B. EfficientNet // Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization, 2021. – P. 109-123.
17. He S., Zou Y., Li B. *et al.* An image inpainting-based data augmentation method for improved sclerosed glomerular identification performance with the segmentation model EfficientNetB3-Unet // *Sci Rep*, 2024. – V. 14. – P. 2045-2322.
18. URL Streamlit documentation (дата обращения – май 2024 г.).
19. URL deneal123/MicroscopeAI (дата обращения – май 2024 г.).