

МАТЕМАТИЧЕСКИЕ МЕТОДЫ АДАПТИВНОГО ПОВЫШЕНИЯ КАЧЕСТВА БИОМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Гриценко Н.Ю., аспирант

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Батура М.П. – д.т.н., профессор

Аннотация. В данной работе исследуются математические методы адаптивного повышения качества биомедицинских изображений. Были рассмотрены современные методы, такие как: адаптивная фильтрация (вейвлет-фильтрация), глубокое обучение (сверточные нейронные сети (CNN), генеративно-сопоставительные сети (GAN)), гибридные методы (нейросети в связке с вейвлет-анализом), трансформеры (ViT, PVT).

Ключевые слова. Обработка изображений, нейронные сети, глубокое обучение, трансформеры.

Введение. Развитие медицинского оборудования способствует увеличению объема и качества данных, однако получаемые изображения часто содержат искажения, такие как шум и размытие, что затрудняет их анализ и диагностику. Для решения этих проблем активно применяются современные математические методы, нейросетевые модели и методы машинного обучения, направленные на повышение качества изображений.

Одним из ключевых математических подходов является адаптивная вейвлет-фильтрация, основанная на вейвлет-преобразовании. Этот метод позволяет разложить изображение на набор базисных функций (вейвлетов), локализованных как в пространственной, так и в частотной областях, что обеспечивает эффективное удаление шума при сохранении важных деталей, таких как контуры и текстуры. Основные этапы вейвлет-фильтрации включают разложение сигнала на коэффициенты, их обработку (например, пороговую фильтрацию) и обратное преобразование для восстановления изображения. Преимущества метода заключаются в его универсальности, локальной адаптации к уровню шума и способности минимизировать артефакты, что делает его применимым для различных типов биомедицинских изображений.

Нейросети, как вычислительные модели, играют важную роль в обработке биомедицинских изображений благодаря своей способности выявлять сложные нелинейные зависимости в данных. Нейросеть (англ. neural network) – это математическая модель, которая имитирует работу человеческого мозга и состоит из множества взаимосвязанных искусственных нейронов. Эти нейроны способны обрабатывать большие массивы данных, находить в них сложные закономерности и обучаться на примерах, что делает нейросети актуальными для решения задач классификации, сегментации и восстановления изображений.

Глубокое обучение – это подраздел машинного обучения, основанный на обучении представлений, где модели автоматически извлекают определенные признаки из данных.

Среди методов глубокого обучения, применимых для повышения качества изображений, выделяются сверточные нейронные сети (CNN), которые автоматически извлекают локальные признаки, такие как края, текстуры и формы объектов, из изображений, что делает их эффективными для задач классификации, сегментации и восстановления.

Генеративно-сопоставительные сети (GAN), состоящие из генератора и дискриминатора, используются для увеличения разрешения изображений, восстановления поврежденных данных и создания синтетических изображений, что особенно актуально в условиях ограниченности реальных данных.

Оба метода обладают высокой адаптивностью и могут быть применены для работы с различными типами медицинских изображений, включая рентген, МРТ и УЗИ.

Гибридные методы, сочетающие вейвлет-анализ и нейросетевые модели, представляют собой эффективный подход для обработки сложных сигналов и изображений. Вейвлет-преобразование используется для предобработки данных, удаления шума и выделения информативных признаков, что упрощает задачу обучения нейросетей. Такие методы находят применение, например, в анализе ЭЭГ-сигналов, где вейвлет-анализ помогает выделить шаблоны активности мозга, а нейросети классифицируют состояние пациента. Ключевой особенностью является то, что гибридные подходы позволяют сохранять важные детали изображений.

Трансформеры, основанные на механизме внимания, представляют собой современный подход к обработке изображений, учитывающий глобальные зависимости в данных. Архитектуры, такие как Vision Transformer (ViT) и Pyramid Vision Transformer (PVT), эффективно решают задачи повышения качества, восстановления и анализа медицинских изображений. В отличие от CNN, трансформеры лишены

предвзятости, вызванной локальными свертками, что позволяет им учитывать долгосрочные контекстные зависимости. PVT, с его пирамидальной структурой, особенно эффективен для задач, требующих высокой детализации, таких как сегментация органов или опухолей. Трансформеры также применяются для увеличения разрешения изображений (Super-Resolution) и генерации синтетических данных, что делает их ценным инструментом в условиях ограниченности реальных биомедицинских изображений.

Современные математические методы, включая адаптивную вейвлет-фильтрацию, нейросетевые модели (CNN, GAN), гибридные методы и трансформеры (ViT, PVT), позволяют эффективно повышать качество биомедицинских изображений. Их применение способствует улучшению диагностической точности, автоматизации обработки данных и решению задач в условиях ограниченности исходных данных, что делает их актуальными в современной биомедицине.

Вейвлет-фильтрация. Этот метод сочетает мощь вейвлет-преобразования для анализа изображений в разных частотных диапазонах с адаптивным подходом к фильтрации, основанном на локальных характеристиках изображения. Ключевая идея заключается в том, чтобы не применять фильтрацию ко всему изображению целиком одинаково, а адаптировать процесс фильтрации к различным областям изображения в зависимости от их свойств. Это особенно важно для биомедицинских изображений, где шум и полезные детали могут быть распределены неравномерно [1].

Метод адаптивной вейвлет-фильтрации начинается с вейвлет-разложения. Исходное изображение подвергается дискретному вейвлет-преобразованию, которое раскладывает изображение на несколько поддиапазонов, соответствующих различным частотам и ориентациям. Обычно выделяют низкочастотный поддиапазон, который содержит общую структуру изображения, его "гладкую" версию, похожую на размытую копию оригинала, и высокочастотные поддиапазоны, содержащие детали изображения, такие как края, текстуры и шум. Различные уровни разложения позволяют выделить детали разных масштабов. Вейвлет-разложение обеспечивает мультимасштабное представление изображения, где информация о разных частотах и пространственных областях разделена. На рисунке 1 показано изображение, на примере которого продемонстрировано преобразование. На рисунке 2 показаны пример всех разложений изображения данным методом.

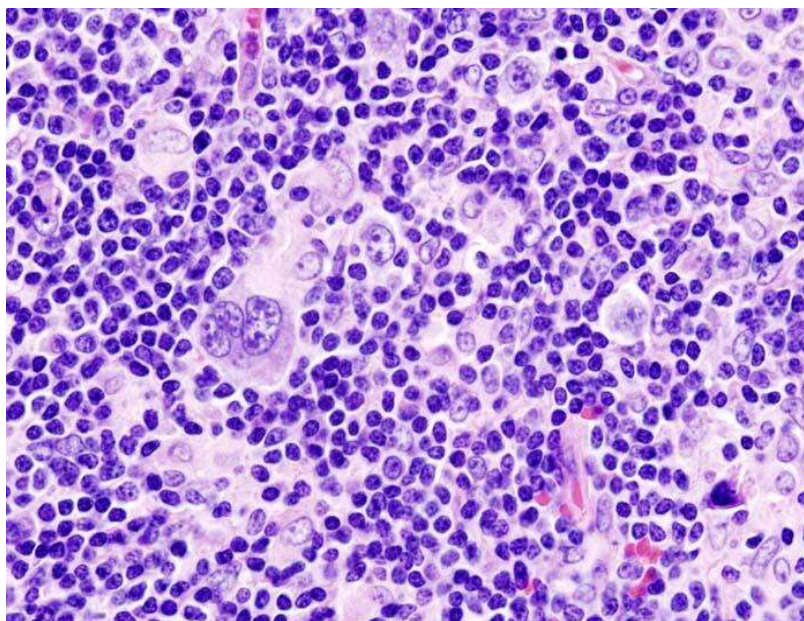


Рисунок 1 – Оригинальное изображение биоптата лимфоузла

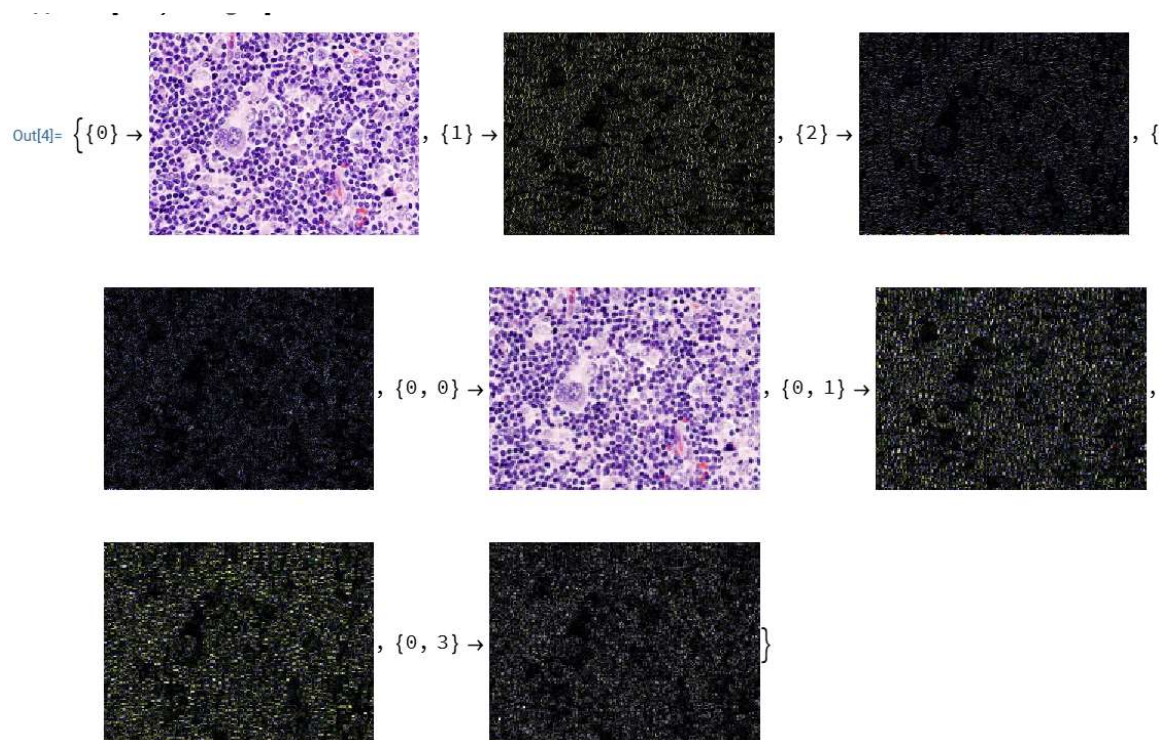


Рисунок 2 – Разложение изображения

Далее происходит локальная оценка характеристик в вейвлет-домене. Именно на этом этапе начинается локальная обработка. Вместо глобальной фильтрации, анализируются вейвлет-коэффициенты внутри небольших локальных областей для каждого поддиапазона. Оцениваются локальные характеристики, которые могут указывать на наличие шума или полезных деталей. Типичные характеристики для оценки включают локальную дисперсию (мера разброса значений коэффициентов в окне, высокая дисперсия может указывать на детали или шум, низкая - на гладкую область), локальную энергию и статистические моменты, такие как асимметрия и эксцесс, которые могут дать дополнительную информацию о распределении коэффициентов и характере сигнала в локальной области.

На основе локальных характеристик, определенных на предыдущем этапе, принимается решение о степени фильтрации для каждого локального окна и каждого поддиапазона, что составляет этап адаптивного фильтрования вейвлет-коэффициентов на основе локальных оценок. Адаптивность проявляется в том, что фильтрация не является фиксированной, а изменяется от области к области. Например, при адаптивном пороговом квантовании, порог устанавливается локально для каждого окна и поддиапазона, определяя, какие вейвлет-коэффициенты будут считаться шумом и будут обнулены или ослаблены.

На последнем этапе выполняется обратное вейвлет-преобразование. После адаптивного фильтрования вейвлет-коэффициентов, выполняется обратное дискретное вейвлет-преобразование которое преобразует модифицированные вейвлет-коэффициенты обратно в пространственную область, формируя фильтрованное изображение. Обратное преобразование изображения из рисунков 1 и 2 показано на рисунке 3.

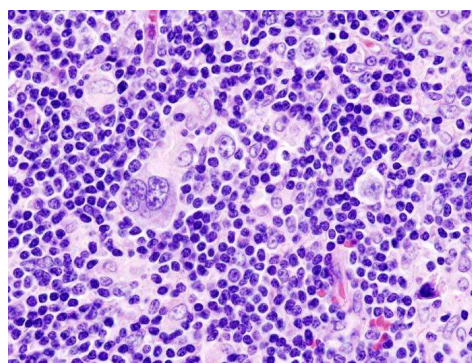


Рисунок 3 – Обратно преобразованное изображение

Адаптивная фильтрация хорошо справляется с неоднородным шумом, часто встречающимся в биомедицинских изображениях, учитывая локальные вариации шума. В результате, по сравнению с глобальными методами фильтрации, адаптивная вейвлет-фильтрация часто обеспечивает улучшенное визуальное качество, аккуратно удаляя шум и сохраняя при этом важную информацию об изображении.

Основной недостаток такой фильтрации – вычислительная сложность, так как локальная обработка и адаптивная оценка характеристик требуют дополнительных вычислений по сравнению с глобальными методами. Агрессивное адаптивное фильтрование, особенно пороговое квантование, иногда может вносить артефакты в изображение, особенно в виде ступенчатых или "блочных" эффектов, если параметры адаптации не подобраны корректно.

Глубокое обучение с использованием CNN сетей и GAN. Метод глубокого обучения для адаптивного улучшения качества изображений с CNN и GAN состоит из нескольких основных компонентов.

Сверточная нейронная сеть (CNN) используется в качестве генератора (G) в GAN, отвечая за преобразование входного изображения с дефектами в улучшенное изображение. Архитектура CNN для генератора может варьироваться, но обычно включает сверточные слои, которые извлекают признаки из входного изображения с помощью обучаемых фильтров, позволяя сети изучать иерархические признаки [2]. Целью CNN в роли генератора является обучение преобразованию дефектного входного изображения в выходное, максимально похожее на идеальное, высококачественное изображение.

Чаще всего используются CNN архитектуры U-net, которая была создана в 2015 году для сегментации биомедицинских изображений. Архитектура сети представляет собой полносвязную сверточную сеть, модифицированную так, чтобы она могла работать с меньшим количеством примеров (обучающих образов) и делала более точную сегментацию.

Еще одна CNN используется в качестве дискриминатора в GAN. Дискриминаторная сеть предназначена для классификации входного изображения на "реальное" (высококачественное) и "поддельное" (улучшенное генератором). Архитектура дискриминатора включает сверточные слои для извлечения признаков, слои активации для нелинейности и пуллинг-слои для уменьшения размерности. Цель дискриминатора – научиться отличать улучшенные генератором изображения от реальных высококачественных изображений, максимизируя свою способность к правильной классификации.

GAN объединяет генератор и дискриминатор в рамках состязательного обучения. Генератор стремится обмануть дискриминатор, создавая улучшенные изображения, выглядящие как "реальные" и высококачественные, в то время как дискриминатор пытается не быть обманутым, точно различая "реальные" и "поддельные" изображения. Обучение GAN основано на двойной функции потерь: потеря дискриминатора, которая стремится классифицировать реальные изображения как "реальные", а поддельные как "поддельные", и потеря генератора, которая стремится генерировать изображения, классифицируемые дискриминатором как "реальные".

Обучение на разнообразных данных позволяет модели обобщать знания и применять их к новым изображениям. CNN автоматически извлекают признаки, описывающие дефекты и особенности изображения. Состязательное обучение GAN заставляет генератор создавать визуально реалистичные улучшения. Условные GAN могут использовать дополнительную информацию, такую как тип дефекта, уровень дефекта или семантическая информация об изображении, в качестве условия для адаптации улучшения. Например, условием может быть тип дефекта, позволяя обучить отдельные модели для каждого типа или использовать информацию о типе дефекта для обработки нескольких типов.

Процесс обучения начинается со сбора данных, включающего набор пар "дефектное - идеальное" изображение, где "дефектные" изображения могут быть получены искусственно или быть реальными изображениями. Затем определяется архитектура CNN для генератора и дискриминатора, используя существующие архитектуры или разрабатывая собственные. Далее, выбирается функция потерь для GAN. После этого генератор и дискриминатор обучаются итеративно в состязательном режиме, используя оптимизаторы для обновления весов сетей. Наконец, модель валидируется и оценивается на тестовом наборе данных с использованием различных метрик и субъективной оценки качества.

Преимущества метода включают адаптивность к специфическим дефектам, высокое качество результатов, автоматическое обучение и возможность улучшения различных типов дефектов.

Основные недостатки метода – требовательность к данным, потребность в вычислительных ресурсах, нестабильность обучения GAN, возможность появления артефактов и сложность интерпретации работы модели.

Гибридные методы (нейросети в связке с вейвлет-анализом). Эти методы стремятся объединить сильные стороны обоих подходов: способность вейвлет-анализа эффективно представлять изображения в различных частотных диапазонах и мощь нейросетей в обучении сложным нелинейным зависимостям и адаптации к данным. Нейронные сети, особенно CNN, превосходно справляются с обучением сложным преобразованиям изображений, извлечением признаков высокого уровня и адаптацией к различным типам дефектов, представленным в обучающих данных. Стандартные CNN могут неявно обрабатывать частотную информацию, что может быть неоптимально для дефектов,

связанных с шумом и артефактами сжатия. Вейвлет-анализ, напротив, предлагает математически обоснованный и эффективный способ разложения изображений на частотные поддиапазоны, позволяя анализировать и обрабатывать изображение на разных уровнях детализации, выделяя низкочастотные и высокочастотные компоненты. Вейвлет-преобразование обладает хорошими свойствами локализации, но традиционные вейвлет-методы улучшения качества часто являются линейными, ограничивая их адаптивность. Гибридные методы стремятся объединить эти подходы, используя вейвлет-анализ для предобработки или декомпозиции изображений, а нейросети для адаптивного и нелинейного преобразования вейвлет-коэффициентов или субполос, улучшая интерпретируемость, повышая эффективность обучения, качество результатов и снижая вычислительные затраты.

Существует несколько основных типов гибридных методов.

Первый тип использует вейвлет-преобразование как предобработку для нейронной сети. В этом подходе изображение сначала подвергается вейвлет-преобразованию, в результате чего получается набор вейвлет-коэффициентов или субполос. Затем эти вейвлет-коэффициенты, или субполосы, используются в качестве входных данных для нейронной сети. Нейросеть обучается преобразовывать эти вейвлет-коэффициенты таким образом, чтобы после обратного вейвлет-преобразования получилось улучшенное изображение.

В рамках этого подхода существует несколько вариантов реализации. Например, нейросеть может обрабатывать непосредственно все вейвлет-коэффициенты как вектор или многоканальное изображение, используя, скажем, CNN для обработки каждой субполосы как отдельного канала. Другой подход заключается в обработке субполос как отдельных изображений, подавая их на вход CNN, где сеть может обрабатывать их независимо или совместно, например, через конкатенацию признаков карт. Альтернативно, можно использовать нейросеть для селективной обработки только определенных субполос, таких как высокочастотные, наиболее подверженные шуму или артефактам, оставляя низкочастотные субполосы без изменений или обрабатывая их традиционными методами.

Второй тип гибридных методов представляет собой каскадные методы, где вейвлет-анализ и нейросеть применяются последовательно. Например, изображение сначала может быть обработано традиционными вейвлет-методами улучшения качества, такими как вейвлет-пороговая фильтрация шума, а затем результат подается на вход нейросети для дальнейшего улучшения. В другом варианте последовательности, изображение сначала обрабатывается нейросетью для первичного улучшения, а затем результат подвергается вейвлет-анализу для точной настройки деталей или устранения артефактов, внесенных нейросетью.

Гибридные методы имеют ряд преимуществ, включая улучшенное представление частотной информации, потенциальную повышенную интерпретируемость, возможность комбинирования сильных сторон вейвлет-анализа и нейросетей, потенциальное улучшение качества результатов и возможность снижения вычислительных затрат. Однако существуют и недостатки, такие как сложность архитектуры и обучения.

Трансформеры. Трансформеры, изначально разработанные для обработки естественного языка, оказали революционное влияние и на компьютерное зрение, продемонстрировав впечатляющие результаты в различных задачах, включая улучшение качества изображений.

Vision Transformer (ViT) и Pyramid Vision Transformer (PVT) представляют собой ключевые архитектуры трансформеров, адаптированные для работы с изображениями, предлагая новые возможности для адаптивного улучшения качества, отличные от традиционных сверточных нейронных сетей и гибридных методов [1].

Трансформеры, благодаря механизму внимания, обладают способностью эффективно моделировать долгосрочные зависимости и захватывать глобальный контекст изображения, что критически важно для адаптивного улучшения, поскольку дефекты часто не локальны и зависят от контекста всего изображения. В отличие от CNN с их локальными сверточными операциями, трансформеры лучше справляются с глобальным контекстом.

Vision Transformer (ViT) для адаптации к улучшению качества изображений требует некоторых изменений в архитектуре, изначально предназначенной для классификации. Входное представление в ViT меняется: вместо обработки целого изображения, ViT разбивает его на патчи, небольшие квадратные фрагменты, которые затем "разворачиваются" в векторы и рассматриваются как "токены".

Для обучения ViT применяются функции потерь, измеряющие разницу между улучшенным и идеальным изображениями. Это могут быть пиксельные потери, такие как среднеквадратичная ошибка или средняя абсолютная ошибка, перцептуальные потери, основанные на признаках предобученных сетей, или состязательные потери в рамках GAN, где ViT выступает генератором. Обучение ViT требует парных наборов данных "дефектное - идеальное" изображение, где "дефектные" изображения могут быть искусственно созданы или быть реальными дефектными изображениями. Аугментация данных также важна для увеличения разнообразия обучающих данных.

Pyramid Vision Transformer (PVT) представляет собой усовершенствование ViT, разработанное для задач плотного предсказания и также эффективно адаптируется для улучшения качества изображений, предлагая ряд преимуществ. PVT использует пирамидальную структуру, аналогичную CNN, для

извлечения признаков на разных масштабах, обрабатывая изображение на ранних стадиях с большими патчами и меньшим разрешением, и на поздних стадиях – с меньшими патчами и большим разрешением. Такая пирамида признаков позволяет эффективно захватывать как глобальный контекст, так и локальные детали, что особенно важно для улучшения качества.

В отличие от ViT, обрабатывающего изображение на одном масштабе, пирамидальная структура PVT более эффективно использует вычислительные ресурсы и повышает качество. В PVT размер патчей постепенно уменьшается по мере продвижения по слоям трансформера, используя сверточные слои с шагом для уменьшения разрешения признаков карт между стадиями пирамиды. Это прогрессивное уменьшение патчей позволяет PVT постепенно увеличивать разрешение и детализацию признаков. Благодаря пирамидальной структуре и уменьшению патчей, PVT может быть более вычислительно эффективным, чем ViT, особенно для изображений высокого разрешения. Для адаптации PVT к улучшению качества, используются подходы, аналогичные ViT, включая энкодер-декодер архитектуру, где энкодер PVT извлекает пирамиду признаков, а декодер восстанавливает улучшенные патчи, механизм внимания для моделирования зависимостей и адаптивной обработки, и аналогичные функции потерь и обучающие данные.

Преимущества трансформеров в улучшении качества изображений включают превосходное глобальное контекстное понимание, адаптивность к пространственно-вариативным дефектам, эффективное моделирование нелокальных взаимодействий и потенциально более высокое качество результатов, а также гибкость архитектуры. Однако, существуют и недостатки. Высокие вычислительные затраты, особенно для ViT, и требовательность к большим объемам данных являются значительными вызовами. Обучение и настройка трансформеров также могут быть сложнее, чем CNN, и интерпретируемость, хотя и присутствует благодаря вниманию, может быть менее прямой, чем у CNN.

Список использованных источников:

1. Elsevier [Electronic resource] : Mallat, S. A Wavelet Tour of Signal Processing: The Sparse Way. – Mode of access: <https://www.elsevier.com/books/a-wavelet-tour-of-signal-processing/mallat/978-0-12-374370-1>. Date of access: 17.03.2025
2. nature [Electronic resource] : LeCun, Y., Bengio, Y., Hinton, G. Deep Learning. – Mode of access: <https://www.nature.com/articles/nature14539>. Date of access: 17.03.2025
3. arXiv [Electronic resource] : Dosovitskiy, A. et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. – Mode of access: <https://arxiv.org/abs/2010.11929>. Date of access: 19.03.2025

MATHEMATICAL METHODS FOR ADAPTIVE QUALITY ENHANCEMENT OF BIOMEDICAL IMAGES

Gritsenko N.Y., Postgraduate Student

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

Batura M.P. – Doctor of Technical Sciences, Professor

Annotation. This work investigates mathematical methods for adaptive quality enhancement of biomedical images. Modern methods have been considered, such as: adaptive filtering (wavelet filtering), deep learning (convolutional neural networks (CNN), generative adversarial networks (GAN)), hybrid methods (neural networks in conjunction with wavelet analysis), and transformers (ViT, PVT).

Keywords. Image processing, neural networks, deep learning, transformers