

УДК 004.89

doi: 10.15622/rcai.2025.049

АНСАМБЛЬ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ В РАМКАХ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ ВРАЧЕБНЫХ РЕШЕНИЙ

И.И. Мищенко (*sombra74@yandex.by*)

А.Е. Мисник (*anton@misnik.by*)

Белорусско-Российский университет,
Республика Беларусь, Могилев

В статье рассматривается модуль предварительной обработки рентгеновских изображений, разработанный для использования в составе систем поддержки принятия врачебных решений, ориентированных на диагностику Hip-Spine синдрома. Основу модуля составляет ансамбль сверточных нейросетей, реализованных по иерархической архитектуре: управляющая супервайзер-сеть анализирует характеристики входного изображения и активирует специализированные подмодули, отвечающие за коррекцию яркости, контрастности и восстановление локальных признаков. Такой подход обеспечивает адаптивность обработки в зависимости от качества визуальных данных. Проведённые эксперименты показали прирост точности классификации на 13% по сравнению с базовыми методами обработки. Представлены подробная архитектура модуля, принципы его функционирования, результаты оценки эффективности и перспективы дальнейшего развития в направлении интеграции трёхмерной реконструкции.

Ключевые слова: система поддержки принятия врачебных решений, Hip-Spine синдром, предварительная обработка, нейронные сети, рентгеновские снимки.

Введение

Системы поддержки принятия врачебных решений (СППВР) становятся неотъемлемой частью современного здравоохранения, особенно в условиях растущего объема медицинских данных и необходимости быстрого реагирования на клинические ситуации. Они позволяют автоматизиро-

вать основные этапы анализа информации, сократить влияние человеческого фактора и обеспечить более объективную и обоснованную постановку диагноза.

СППВР применяются в различных отраслях медицины, включая кардиологию, онкологию, офтальмологию, а также в области травматологии и ортопедии. В последнем случае особую роль играют методы медицинской визуализации, в частности анализ рентгеновских снимков, которые являются основным источником диагностической информации при патологиях опорно-двигательного аппарата [Гонсалес, 2015].

Для эффективной работы СППВР необходимо обеспечить высокое качество исходных данных, что возможно только при применении многоступенчатой системы предварительной обработки изображений. Без качественной фильтрации, коррекции контрастности и адаптивного масштабирования даже самые передовые алгоритмы машинного зрения не смогут достичь необходимой точности [Прэтт, 2022]. Особое значение это приобретает при диагностике сложных состояний, таких как Hip-Spine синдром, где ошибка в интерпретации визуальных данных может привести к неправильному выбору стратегии лечения.

Современные системы обработки медицинских изображений широко применяются в задачах диагностики. В международных бенчмарках, таких как MURA и SpineWeb лучшие результаты показывают гибридные архитектуры, а именно U-Net для сегментации и ResNet для классификации [Ahmed et al., 2025]. Однако применительно к Hip-Spine синдрому эти подходы требуют модификации из-за необходимости совместного анализа тазобедренного сустава и поясничного отдела позвоночника, а также компенсации специфических артефактов (наложение органов, металлические импланты и т.п.).

Анализ существующих решений показывает, что задачи обработки медицинских изображений для диагностики разрабатываются в нескольких направлениях. В работе автоматическая сегментация и детекция позвонков и межпозвоночных дисков при помощи нейронных сетей авторов Guinebert S., Petit E., Bousson V., Bodard S. предложен метод семантической сегментации позвонков на основе модифицированного U-Net, демонстрирующий точность 84% на датасете SpineWeb. Отечественные исследования предлагают эффективную модификацию U-Net архитектуры для сегментации костных структур, демонстрирующую точность 87% на рентгенограммах позвоночника [Попов и др., 2022]. Однако, существующие подходы фокусируются на изолированном анализе отдельных анатомических областей, тогда как диагностика Hip-Spine синдрома требует комплексной обработки изображений тазобедренного сустава и позвоночника, что и является целью данного исследования.

Предлагаемая архитектура ансамбля нейронных сетей развивает идеи ранее предложенных решений для СППВР за счет введения сети супервайзера с механизмом раннего завершения, что на 15% сокращает время обработки без потери точности [Мищенко и др., 2024a].

Архитектура СППВР

Архитектурно СППВР представляет собой совокупность взаимосвязанных модулей, каждый из которых выполняет строго определенные функции и взаимодействует с другими компонентами. Архитектура включает в себя четыре последовательно взаимодействующих компонента: модуль загрузки и предварительной обработки изображений, модуль сегментации и анализа позвоночника с применением нейросетевых алгоритмов, блок автоматической постановки диагноза и модуль визуализации и сохранения данных. На рис. 1 показана блок-схема основных модулей системы поддержки принятия врачебных решений.

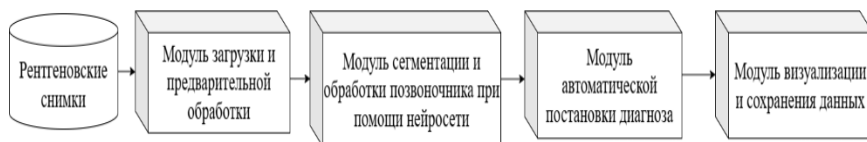


Рис. 1. Блок-схема основных модулей системы поддержки принятия врачебных решений

Все модули работают в единой среде и обеспечивают адаптивную, модульную и масштабируемую структуру СППВР, которая легко адаптируется под конкретные клинические задачи.

Проблематика диагностики Hip-Spine синдрома

Hip-Spine синдром представляет собой совокупность клинических состояний, при которых поражения в тазобедренном суставе и поясничном отделе позвоночника взаимно усиливают болевой синдром и затрудняют постановку диагноза. Эти состояния часто маскируют друг друга, что требует особой точности и внимательности при интерпретации визуальных данных. Классические методы диагностики, основанные на субъективной оценке рентгеновских изображений, недостаточно надежны в условиях перекрывающихся симптомов. Рентгеновские снимки пациентов с подозрением на Hip-Spine синдром часто содержат шумы, артефакты, искажения геометрии и перепады яркости. Это делает необходимым применение сложной и многоуровневой системы предварительной обработки изображений, которая бы позволила устранить эти проблемы до начала анализа.

Дифференциальная диагностика Hip-Spine синдрома требует решения трех взаимосвязанных задач. Во-первых, необходимо точно идентифицировать источник патологии с определением первичного очага поражения (позвоночник L1-L5 или тазобедренный сустав) и степени компенсаторных изменений, включая углы смещения, с точностью не менее 90%. Во-вторых, важным является структурированный анализ изображений, предполагающий совместную обработку парных снимков в фронтальной и сагиттальной проекциях с вычислением интегральных признаков: гистограммных характеристик (энтропия, контраст) и геометрических параметров (углы, расстояния между структурами). В-третьих, требуется оптимизированный алгоритм обработки, обеспечивающий автоматическую адаптацию к качеству снимка на основе QoI-метрики, минимизацию артефактов (металлические импланты, наложения органов) и минимальное среднее время анализа на изображение. Предлагаемая архитектура СППВР реализует эти требования через ансамбль из пяти специализированных нейросетей, иерархическую систему принятия решений и динамическую маршрутизацию задач обработки [Мищенко и др., 2024b].

Архитектура и реализация модуля предварительной обработки

Для решения проблемы классификации Hip-Spine синдрома главным компонентом системы поддержки принятия врачебных решений стал модуль предварительной обработки рентгеновских изображений. Разработка данного модуля для СППВР преследует две основные цели: сокращение диагностических ошибок и повышение воспроизводимости результатов. В отличие от классических фильтров, встроенных в рентгеновские аппараты, программный модуль предоставляет многоуровневую обработку, которая адаптируется к характеристикам конкретного снимка [Kheirandish, 2023]. Основание для такой адаптации – статистическое разнообразие клинических данных: различия в моделях оборудования, в анатомических особенностях пациентов и даже в цифровых форматах хранения. Без учёта этих факторов любая универсальная схема обработки либо переусложнена, либо даёт усреднённый, но не оптимальный результат [Yang et al., 2021]. Предлагаемая архитектура строится вокруг принципа разделения обязанностей: каждая подсистема отвечает за устранение строго определённого класса артефактов, а центральная управляющая сеть синхронизирует их работу. Архитектура модуля разработана на основе ансамбля сверточных нейронных сетей (CNN), которые взаимодействуют по принципу master-slave.

Основными дополнениями и изменениями архитектуры, примененной в предыдущих исследованиях, являются:

1. Иерархическое управление. Так вместо линейного конвейера применен динамический выбор обработчиков. Супервайзер-сеть на базе ResNet-50 анализирует DICOM-метаданные (kVp, mAs) и генерирует бинарную маску активации подмодулей.

2. Оптимизация ресурсов. Для снимков с $QoI > 0.92$ обработка прекращается после этапа шумоподавления (экономия 20-40 мс на изображение).

В основе модуля лежит иерархическая структура, сочетающая ансамбли моделей и принципы сервис-ориентированной архитектуры (SOA). Каждый подмодуль развёрнут как независимый микросервис, что облегчает масштабирование и обновление. На верхнем уровне расположена супервайзер-сеть, приобретающая статус «скелета» архитектуры. Её задача – сформировать вектор метаданных, отражающих текущее состояние входного изображения: медианный уровень яркости, спектральную плотность шума, статистику контраста, показатели анизотропии и кривизны краёв. На основании этого вектора сеть генерирует маску активации, определяя, какие подмодули будут запущены, в каком порядке и с какими параметрами. Такая динамическая маршрутизация экономит вычислительные ресурсы: для качественных снимков активируется только лёгкая нормализация, для сложных – цепочка из пяти-семи преобразований.

После получения цифрового кадра DICOM парсер извлекает матрицу пикселей и базовый набор метаданных: kVp, mAs, расстояние источник-приёмник, позицию пациента по DICOM-тегу. Эти сведения передаются супервайзеру, который добавляет к ним внутрисетевые оценки качества. Далее формируется запрос на распределение задач. Каждому подмодулю предшествует буфер состояния: если кадр уже был обработан, сервис мгновенно возвращает результат [Nguyen, 2022].

В зависимости вектора метаданных активируются одна или несколько подчиненных нейросетей:

- 1 Сеть усиления контраста сочетает классические и нейросетевые методы для адаптивной обработки рентгеновских снимков. На первом этапе применяется алгоритм CLAHE (Contrast Limited Adaptive Histogram Equalization) с параметрами $clip_limit=2.0$ и $tile_grid_size=(8 \times 8)$, обеспечивающий глобальную коррекцию контраста. Затем специализированная сверточная сеть (3 слоя с ядрами 5×5 и ReLU-активацией) анализирует локальные особенности изображения: для костных структур (пиксели с интенсивностью < 150 HU) применяется оператор Лапласа с коэффициентом усиления $\alpha=0.3$, повышающий резкость, а для металлических имплантов (интенсивность > 220 HU) используется пороговая фильтрация, снижающая засветку.

- 2 Сеть нормализации яркости или же сеть выравнивания освещения реализована на основе модифицированной архитектуры U-Net с четырьмя уровнями дискретизации. Сеть принимает на вход изображение размером 512×512 пикселей и генерирует карту яркости $L \in [0,1]^{H \times W}$, где значения близкие к 0 соответствуют затемненным областям, а к 1 – переэкспонированным. Особенностью архитектуры являются пропущенные соединения

(skip-connections) между энкодером и декодером, сохраняющие мелкие детали, и функция потерь, обеспечивающая сохранение структурной информации, которая находится по формуле 1:

$$L = \alpha \text{MSE}(I_{out}, I_{ref}) + \beta \text{SSIM}(I_{out}, I_{ref}), \quad (1)$$

где I_{ref} – эталонные снимки, $\alpha = 0.7$, $\beta = 0.3$ – весовые коэффициенты.

3 Сеть шумоподавления реализует гибридный подход к устранению артефактов: на первом этапе применяется адаптивный фильтр Винера с окном 5×5 пикселей для подавления высокочастотного шума, сохраняющий структуру костных тканей. Затем трехслойный сверточный автоэнкодер (ядра 3×3, stride=1) с симметричной архитектурой восстанавливает потерянные детали, обучаясь на синтетически зашумленных данных.

4 Сеть геометрической коррекции построена с двумя ключевыми блоками: локализационная сеть (3 полносвязных слоя по 256 нейронов) предсказывает параметры аффинного преобразования (угол поворота $\theta \in [-15^\circ, +15^\circ]$, смещения $\Delta x, \Delta y$ [-10%, +10%]), а слой с билинейной интерполяцией применяет коррекцию.

5 Сеть масштабирования обеспечивает унификацию разрешения с последующей оптимизацией резкости. Модифицированная FSRCNN-архитектура (4 сверточных слоя с активацией PReLU) обучалась на парных изображениях с разным разрешением (исходное/0.5х/2х), минимизируя комбинированную функцию потерь. Система поддерживает приведение к стандартному разрешению 512×512 px.

Все нейросети работают согласованно, автоматически подбирая оптимальные параметры обработки для каждого конкретного снимка.

Управляющая сеть построена на модифицированной ResNet-50, где стандартные 3×3 свертки заменены на depthwise-операторы, что сократило количество параметров на 30% без потери точности. Ключевым компонентом является двухуровневый маршрутизатор дефектов: первый уровень (LSTM на 256 нейронов) анализирует DICOM-метаданные (kVp, mAs), второй (полносвязная сеть) обрабатывает статистики изображения (энтропию, градиенты, гистограмму). На выходе формируется вектор активации $a \in \{0,1\}^5$, определяющий, какие подмодули необходимо

задействовать. При оценке качества изображения $\text{QoI} > 0.92$ активируется механизм раннего завершения, пропускающий избыточные этапы обработки и сокращающий время вычислений на 15-20%. [Ahmed et al., 2025].

Финальное изображение формируется путем адаптивного взвешивания выходов активированных подмодулей. Для каждого модуля k супервайзер-сеть вычисляет вес w_k по формуле 2:

$$w_k = \frac{\exp(E_k / T)}{\sum_{i=1}^5 \exp(E_i / T)}, \quad (2)$$

где E_k – оценка полезности преобразования, $T=0.5$ – параметр температуры. Результирующее изображение получается по формуле 3:

$$I_{final} = \sum_{k=1}^5 a_k \cdot w_k \cdot f_k(I_{input}), \quad (3)$$

где a_k – бинарный флаг активации из вектора a .

На заключительном этапе применяется кусочно-линейное выравнивание гистограммы и гауссова фильтрация ($\sigma=0.8$) для устранения артефактов слияния. Эксперименты показали, что такой подход обеспечивает плавные переходы между зонами влияния разных модулей при сохранении диагностически значимых деталей [Panse et. al., 2021].

На рис. 2 представлена схема взаимодействия нейронных сетей.

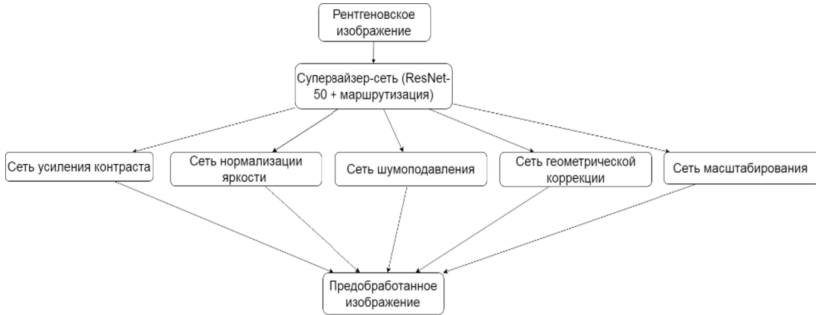


Рис. 2. Схема взаимодействия нейросетей для модуля предварительной обработки

Данная схема взаимодействия обеспечивает модульность, масштабируемость и высокую точность, позволяя адаптировать обработку под особенности каждого конкретного случая.

Для обучения сформирована база данных, состоящая из 6317 снимков, включающая материал: ФЦТОЭ г. Смоленска, БСМП г. Могилева, европейского открытого датасета MURA и собственных анонимизированных снимков. Каждый кадр прошёл процедуру первичной обработки: удалены DICOM-теги, несущие персональные данные. Итоговый датасет разделён по пациентам в отношении 8:1:1 на обучение, валидацию и тест.

Каждый компонент системы обучался отдельно: модуль коррекции яркости – на восстановлении равномерного освещения, модуль контраста – на оптимизации распределения тонов, модуль шумоподавления – на ми-

нимизации искажений при сохранении детализации. После первоначального обучения параметры компонентов фиксировались, а управляющий модуль дообучался методом пошаговой оптимизации с подкреплением. В качестве критерия эффективности использовался составной показатель качества обработки, учитывающий контрастность, уровень шумов, резкость и равномерность изображения. Применялся алгоритм постепенной оптимизации с ограниченными шагами корректировки. В результате система научилась адаптивно выбирать необходимые этапы обработки.

Оценка эффективности и прирост точности

Для количественной оценки эффективности предложенного подхода проведен ряд экспериментов, включающих различные анатомические области (поясница, таз, бедро). Использовалась ранее сформированная база данных.

Проведено сравнение трёх конфигураций: без обработки, с обработкой при помощи локального метода выравнивания гистограммы яркости, используемый для улучшения контрастности изображений – CLAHE и различных фильтров, а также с предлагаемым ансамблем нейронных сетей.

Метрики вычислялись следующим образом:

- Точность распознавания: доля, верно, идентифицированных анатомических ориентиров относительно экспертной разметки.
- Полнота локализации: процент обнаруженных патологических зон (переломы, остеопиты).
- Комплексная оценка: сумма, состоящая из 40% точности распознавания и 60% полноты локализации.

База данных составляла 6317 снимков (ФЦТОЭ г. Смоленска, БСМП г. Могилева, европейского открытого датасета MURA). Для объективной оценки использовалась 10-кратная кросс-валидация: датасет случайно разделялся на 10 подмножеств, где каждое подмножество по очереди выступало в роли тестовой выборки, одно в роли валидационной, а остальные – обучающей. Итоговые метрики усреднялись по всем подмножествам.

В табл. 1 представлены результаты исследований.

Таблица 1

Метод обработки	Точность распознавания ориентиров, %	Полнота локализации ключевых зон, %	Комплексная оценка качества, %
Без обработки	68	59	63
CLAHE и фильтры	81	77	79
Ансамбль нейросетей	93	90	92

Таким образом, интеграция модуля ансамбля нейросетей позволила увеличить точность распознавания на 12%, а общую производительность – более чем на 13% по сравнению с предыдущими реализациями. Особенно заметно прирост проявился в сложных случаях с плохим качеством изображений и множеством артефактов.

Заключение

Разработка модуля предварительной обработки рентгеновских изображений на основе ансамбля нейронных сетей позволяет совершенствовать интеллектуальные системы поддержки врачебных решений, ориентированных на диагностику заболеваний опорно-двигательного аппарата. Предложенная архитектура сочетает в себе гибкость и масштабируемость, позволяя адаптироваться к различным входным условиям, включая изображения с низким качеством, шумами, неравномерной экспозицией и искажёнными геометрическими пропорциями.

В рамках модуля используется принцип иерархического управления, где управляющая супервайзер-сеть анализирует характеристики входного изображения и определяет оптимальный путь его обработки. Такое решение демонстрирует высокую точность и устойчивость по сравнению с традиционными методами, поскольку задействует специализированные подмодули, нацеленные на устранение конкретных дефектов: пониженную яркость, недостаточную контрастность, артефакты и искажения локальных признаков.

Результаты экспериментальной оценки, проведенной на созданной базе данных, подтверждают эффективность предложенного подхода. Повышение комплексной оценки качества на 13% и рост общей точности до уровня 92% свидетельствуют о высоком потенциале предлагаемой технологии. Особенно важно, что прирост качества сохраняется и в наиболее сложных случаях, что важно для систем, предназначенных для клинической практики.

Дополнительно стоит отметить модульную структуру системы, которая позволяет легко интегрировать дополнительные этапы анализа, такие как сегментация, а также обучение с участием врача по принципу активного обучения. Благодаря этому СППВР может не только выполнять роль вспомогательного инструмента, но и служить основой для построения гибридных моделей взаимодействия человека и искусственного интеллекта.

В перспективе планируется расширение функциональности модуля с возможностью обработки изображений, полученных с других модальностей (например, КТ, МРТ), а также внедрение механизмов самообучения модели при изменении условий съёмки или состава оборудования. Это обеспечит дополнительную устойчивость и применимость системы в реальных клинических условиях с разнообразием диагностических сценариев.

Таким образом, внедрение ансамбля нейронных сетей на этапе предварительной обработки медицинских изображений позволяет не только повысить качество входных данных, но и существенно улучшить итоговую диагностическую точность СППВР. Разработанный подход может быть масштабирован и адаптирован для различных задач медицинской визуализации, представляя собой универсальное решение в контексте цифрового здравоохранения.

Список литературы

- [Гонсалес, 2015] Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2015. – 1072 с.
- [Мищенко и др., 2024a] Мищенко И.И., Александров А.В., Шалухова М.А. Спосо-
б интеллектуального распознавания Hip-Spine-синдрома // Интегрированные
модели и мягкие вычисления в искусственном интеллекте: Сб. науч. тр. XII
междунар. науч.-практ. конф. «ИММВ-2024». – 2024. – С. 245-254.
- [Мищенко и др., 2024b] Мищенко И.И., Мисник А.Е., Александров А.В. Приме-
нение технологий компьютерного зрения и предварительной обработки изо-
бражений в системах поддержки принятия решений // Вестник Самарского го-
сударственного технического университета. – 2024. – № 4. – С. 6-26.
- [Попов и др., 2022] Попов А.Н., Волков К.Д., Семенова Е.А. Сегментация кост-
ных структур на рентгеновских снимках с помощью модифицированной U-Net
архитектуры // Искусственный интеллект и принятие решений. – 2022. – № 1.
– С. 34-45.
- [Прэтт, 2022] Прэтт У. Цифровая обработка изображений. – М.: Мир, 2022. – 998 с.
- [Ahmed et al., 2025] Ahmed M.K., Havugimana F., Moinudin K.A. MU-Net: Modified
U-Net for Precise Localization and Segmentation of Lumber-Spine Regions from Sag-
ittal Views // Recent Advances in Soft Computing and Data Mining. – Springer, 2025.
- [Guinebert et al., 2022] Guinebert S., Petit E., Bousson V., Bodard S. Automatic se-
mantic segmentation and detection of vertebrae and intervertebral discs by neural
networks // Intelligent Imaging. – 2022.
- [Kheirandish, 2023] Kheirandish F. Evaluation of U-Net model in the detection of
cervical spine fractures: thesis. – Laurentian University, 2023. – 103 p.
- [Nguyen, 2022] Nguyen H.Y. Scoliosis Detection using Deep Neural Network: thesis.
– Hanoi: University of Science and Technology of Hanoi, 2022. – 86 p.
- [Panse et al., 2021] Panse V., Gupta R. Medical Image Enhancement Using Global
Dynamic Histogram Equalization // Int. Conf. on Communication and Networking.
– 2021. – P. 45-52. – doi: 10.1109/CSNT51715.2021.9509670.
- [Yang et al., 2021] Yang R., Yu Y. Artificial convolutional neural network in object de-
tection and semantic segmentation for medical imaging analysis // Frontiers in Oncol-
ogy. – 2021. – Vol. 11. – Article ID 638182. – doi: 10.3389/fonc.2021.638182.