

АДАПТИВНЫЙ АЛГОРИТМ ЦИФРОВОЙ ОБРАБОТКИ ВЕТЕРИНАРНЫХ РЕНТГЕНОГРАММ ДЛЯ ЗАДАЧ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Титов Артем Денисович¹,

e-mail: a.titov@rgau-msha.ru

¹Российский государственный аграрный университет – МСХА им. К.А. Тимирязева,
г. Москва, Россия

Актуальность исследования обусловлена необходимостью повышения точности и воспроизводимости автоматизированного расчета кардиовертебрального индекса (VHS) у мелких домашних животных. Одним из ключевых факторов, влияющих на качество сегментации анатомических структур, является низкая контрастность исходных рентгенограмм, особенно в области грудных позвонков и сердечного силуэта. В настоящей работе предложен адаптивный алгоритм цифровой обработки изображений, сочетающий локальное усиление контраста, морфологическое выделение костных структур и нелинейное подавление шума с сохранением границ. Алгоритм предназначен для использования на двух этапах жизненного цикла модели: при подготовке датасета для разметки экспертами и в качестве предобработки перед инференсом обученной модели компьютерного зрения. Описание метода включает математическую формализацию компонентов, обоснование выбора параметров и архитектуру алгоритмического конвейера. Проведена экспериментальная верификация эффективности разработанного алгоритма, включая сравнение качества разметки и производительности моделей, обученных на исходных и обработанных изображениях.

Ключевые слова: ветеринарная рентгенология, кардиовертебральный индекс, улучшение качества изображений, цифровая обработка, адаптация гистограммы, морфологическая обработка, компьютерное зрение

ADAPTIVE DIGITAL PROCESSING OF VETERINARY RADIOGRAPHS FOR COMPUTER VISION TASKS

Titov A.D.¹,

e-mail: a.titov@rgau-msha.ru

¹Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, Moscow, Russia

The relevance of this study stems from the need to improve the accuracy and reproducibility of automated vertebral heart score (VHS) calculation in small companion animals. One of the key factors affecting the quality of anatomical structure segmentation is the low contrast of original radiographs, particularly in the regions of the thoracic vertebrae and the cardiac silhouette. This article proposes an adaptive digital image processing algorithm that combines local contrast enhancement, morphological extraction of bony structures, and nonlinear edge-preserving noise suppression. The algorithm is designed for use at two stages of the model lifecycle: during dataset preparation for expert annotation, and as a preprocessing step prior to inference of a trained computer vision model. The method description includes mathematical formalization of the components, justification for parameter selection, and the architecture of the algorithmic pipeline. Experimental validation of the developed algorithm's effectiveness was conducted, including a comparison of annotation quality and the performance of models trained on original versus processed images.

Keywords: veterinary radiology, vertebral heart score, image quality enhancement, digital processing, histogram adaptation, morphological processing, computer vision

Введение

Современные методы автоматизированной диагностики в ветеринарной рентгенологии все чаще опираются на технологии глубокого обучения, в частности на архитектуры, способные выполнять одновременно сегментацию и детекцию ключевых точек [1; 2]. Однако эффективность таких моделей напрямую зависит от качества входных данных. Ветеринарные рентгенограммы, особенно полученные в условиях общей практики, часто характеризуются низкой контрастностью, неоднородным освещением и наличием шумовых артефактов, что затрудняет как ручную разметку, так и автоматическое выделение анатомических структур.

Двухмерные рентгеновские изображения являются и неисчерпаемым источником информационной неоднозначности, обусловленной наложением теней, различных по своей анатомической и гистологической структуре, составу и плотности органов и тканей. Как результат – рентгеновский снимок может содержать десятки визуальных паттернов, присущих сотням патологических процессов и состояний.

Как отмечается в современных обзорах, комплексное внедрение систем искусственного интеллекта в информационное пространство медицинских организаций влечет повышение экономической эффективности их деятельности, связанное, прежде всего, с высвобождением врачей-специалистов от рутинной работы и возможностью сосредоточиться на более трудоемких процессах при оказании медицинской помощи [3].

Особую роль в обеспечении надежности работы нейросетевых алгоритмов играет этап предварительной обработки изображений. Входные изображения могут иметь различное исходное разрешение, уровни яркости и контраста, степень детализации и шума. Некоторые изображения могут быть предварительно сконвертированы в *grayscale*-формат с расширением *.png, *.jpg и другие [4].

Хотя компьютерное зрение уже успешно применяется в медицине человека, его адаптация для ветеринарии пока остается малоизученной. Это связано с разницей в анатомии животных, отсутствием единых стандартов для разметки и нехваткой специализированных наборов данных.

В связи с этим целью данной работы является разработка и экспериментальная верификация адаптивного алгоритма улучшения качества ветеринарных рентгенограмм, специально ориентированного на усиление контрастности грудных позвонков T4-T10 и сохранение целостности границ сердечного силуэта. Особое внимание уделено интеграции разработанного метода в два ключевых этапа жизненного цикла модели компьютерного зрения:

- подготовку датасета для экспертной разметки;
- предобработку перед инференсом обученной модели.

Решение данной задачи направлено на повышение воспроизводимости автоматизированного расчета VHS и снижение зависимости результатов диагностики от качества исходных рентгенографических снимков.

Анализ методов улучшения качества рентгенограмм

Подходы к улучшению изображения делятся на две категории: методы обработки в пространственной области и методы обработки в частотной области. Термин «пространственная область» относится к плоскости изображения как таковой, и данная категория объединяет подходы, основанные на прямом преобразовании значений пикселей изображения. Частотные методы предполагают изменения изображений после преобразования Фурье [5].

Пространственные методы описываются уравнением:

$$g(x, y) = T[f(x, y)],$$

где $f(x, y)$ – функция, описывающая исходное изображение;

$g(x, y)$ – преобразованное изображение;

T – оператор над f , определенный в некоторой окрестности пиксела с координатами (x, y) .

Под окрестностью пиксела понимается квадратная или прямоугольная область, являющаяся подмножеством изображения и центрированная относительно данного пиксела. Простейший вариант опе-

ратора T используется, когда окрестность состоит из одного пиксела, в этом случае значение g является функцией от $f(x, y)$ и T называется преобразованием точечного типа [5; 6].

Перспективным подходом является адаптивная эквализация гистограммы (CLANE), предложенная Pizer et al. [7], которая позволяет избежать чрезмерного усиления шума за счет ограничения клиппинга и локальной обработки тайлов изображения [7]. Даже незначительное улучшение визуализации костных структур может привести к существенному росту метрик сегментации и повысить точность расчетных индексов [2].

Когда речь заходит об усилении мелких высокочастотных деталей, в данном случае грудных позвонков, которые на рентгенограмме часто «теряются» на фоне паренхимы легких – наиболее эффективным инструментом оказывается преобразование *white top-hat*. Этот метод, ставший классической математической морфологией благодаря работам Ж. Сэрри [8], реализующийся по принципу селективного вычитания, который подавляет крупномасштабные градиенты фона и оставляет в кадре лишь те яркие объекты, чей геометрический масштаб сопоставим с размером выбранного нами структурирующего элемента. Ярким примером служит архитектура системы «ФтизисБиоМед», морфологический каскад выделен в обязательный блок подготовки данных, без которого невозможна стабильная работа основной нейросетевой модели при анализе торакальных снимков [9].

На заключительном этапе обработки критически важно сохранять четкость анатомических границ. Прямое применение классических линейных фильтров (Гаусса, медианы) часто приводит к потере высокочастотных деталей. Решением этой проблемы является использование *bilateral*-фильтра – нелинейного пространственно-диапазонного фильтра, способного подавлять шум без размытия резких переходов яркости [10].

Таким образом, в работе решается задача преобразования изображения рентгенограммы животного таким образом, чтобы усилился контраст костных структур, в частности грудных позвонков T4-T10, были сохранены границы сердечного силуэта без искажений, подавлялся шум без размытия анатомических границ и результат был оптимален для последующей сегментации с помощью модели компьютерного зрения.

Архитектура предлагаемого алгоритма

Предложенный алгоритм реализован как последовательность из трех этапов, объединенная в единый вычислительный конвейер (рисунок 1).

На первом этапе выполняется адаптивная оценка изображения. Рентгенограмма приводится к *grayscale*-формату, выделяется область интереса грудной клетки (1), рассчитывается локальный контраст (2) и применяется CLANE (4) для повышения различимости слабоконтрастных участков.

Второй этап предназначен для морфологического усиления анатомически значимых структур. Формируется градиентная маска, выполняется защищенное смешивание исходного изображения и результата CLANE, после чего применяется *white top-hat*-преобразование (6). Степень усиления регулируется коэффициентом α , который зависит от рассчитанного локального контраста. Это позволяет усилить грудные позвонки, сохраняя при этом контуры сердечного силуэта и снижая риск чрезмерного искажения изображения.

На третьем этапе выполняется нелинейное подавление шума и формируется итоговая улучшенная рентгенограмма (7), пригодная для последующей разметки, сегментации и расчета индексов.

Научная новизна предлагаемого подхода заключается в том, что используемые методы цифровой обработки объединены не как универсальная последовательность фильтров, а как специализированный адаптивный алгоритм, настроенный непосредственно для ветеринарных рентгенограмм в правой латеральной проекции. В отличие от стандартного улучшения изображений с фиксированными параметрами, алгоритм учитывает неоднородность анатомической структуры животного на рентгенограмме, значимую для дальнейшей автоматизированной оценки VHS.

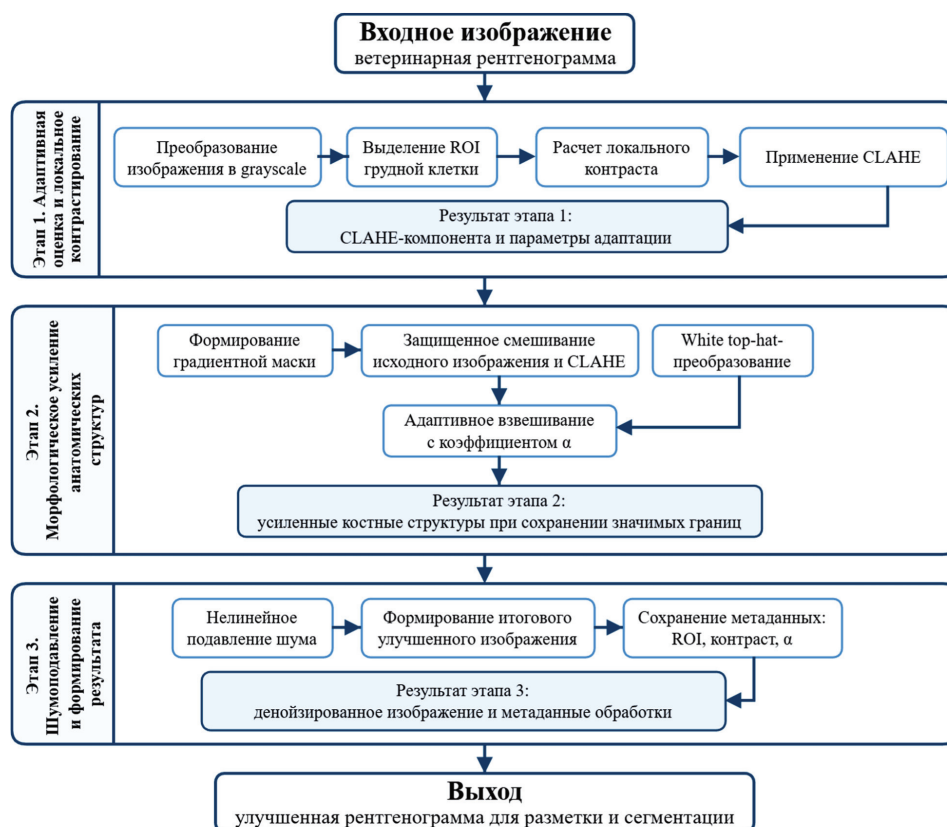


Рисунок 1 – Блок-схема адаптивного алгоритма предобработки ветеринарных рентгенограмм в правой латеральной проекции¹

Оценка региона интереса

Первым шагом алгоритма является оценка анатомически значимого региона $R = [x_1, x_2] * [y_1, y_2]$ – грудной клетки. На основе анализа датасета установлено, что область грудной клетки с высокой вероятностью занимает центральную часть изображения. Регион интереса (далее – ROI) выделяется автоматически без участия оператора. Координаты прямоугольника рассчитываются по относительным отступам от границ изображения:

$$\begin{aligned} x_1 &= \lfloor p_x \cdot W \rfloor, \\ x_2 &= \lfloor (1 - p_x) \cdot W \rfloor, \\ y_1 &= \lfloor p_y^{start} \cdot H \rfloor, \\ y_2 &= \lfloor p_y^{end} \cdot H \rfloor, \end{aligned} \quad (1)$$

где W, H – ширина и высота изображения в пикселях;

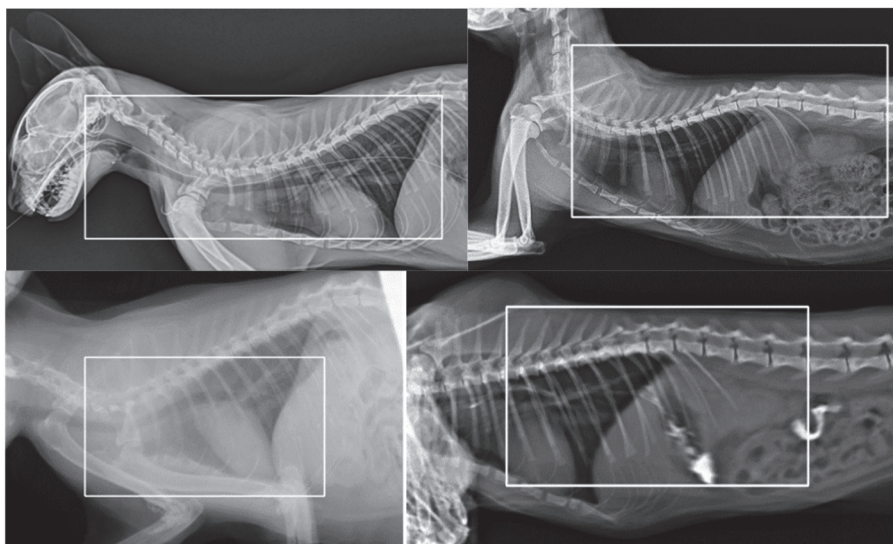
$p_x = 0.2$ – относительный отступ от левого края изображения;

$p_y^{start} = 0.3, p_y^{end} = 0.7$ – относительные границы по вертикали.

Такой подход исключает необходимость обучения дополнительной модели детекции анатомических ориентиров и реализуется за константное число арифметических операций.

Эвристика основана на анализе 539 рентгенограмм в правой латеральной проекции, собранных в ветеринарных клиниках Москвы в 2023–2024 годах. Примерно у 96 % изображений сердце и позвонки T4-T10 находились в указанной зоне (рисунок 2). Подобный метод использовался в системе «Фтизис-БиоМед» для сегментации легочных полей на флюорограммах [9].

¹ Составлено автором.

Рисунок 2 – Пример выделения ROI²

Для оценки качества визуализации костных структур в ROI введен показатель локального контраста:

$$C = \frac{1}{|R|} \sum_{(x,y) \in R} G(x,y), \quad (2)$$

где $|R| = (x_2 - x_1)(y_2 - y_1)$ – площадь ROI,

$G(x,y)$ – магнитуа градиента яркости в точке изображения (3).

Магнитуа градиента интенсивности служит количественной мерой локального контраста и напрямую коррелирует с четкостью анатомических границ. Этот подход был формализован Пели в 1990 году как основа для оценки качества изображений в задачах медицинской визуализации [11]:

$$\nabla I(x,y) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right),$$

где частные производные аппроксимируются дискретными свертками. В настоящей работе используется оператор Собеля с ядрами:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I(x,y), G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * I(x,y).$$

Магнитуа градиента вычисляется как:

$$G(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)}. \quad (3)$$

Реализация алгоритма представлена на рисунке 3 на основе функции `_local_contrast_in_roi`.

```
def _local_contrast_in_roi(self, img: np.ndarray, roi: tuple) -> float:
    """Оценка локального контраста как средняя магнитуа градиента в ROI."""
    x1, y1, x2, y2 = roi
    patch = img[y1:y2, x1:x2]
    grad_x = cv2.Sobel(patch, cv2.CV_64F, 1, 0, ksize=3)
    grad_y = cv2.Sobel(patch, cv2.CV_64F, 0, 1, ksize=3)
    grad_mag = np.sqrt(grad_x**2 + grad_y**2)
    return np.mean(grad_mag)
```

Рисунок 3 – Оценка локального контраста в ROI³

² Составлено автором.

³ Составлено автором.

Функция принимает изображение в виде массива чисел и кортеж с координатами ROI, по которым выбирается область для дальнейшего расчета на изображении. Массив проходит через ядра горизонтального и вертикального градиентов Собеля, после чего рассчитывается средняя магнитуда градиента яркости.

В отличие от фиксированных коэффициентов, предлагаемый подход использует рассчитанный в ROI локальный контраст как управляющий параметр для выбора коэффициента α на следующем этапе обработки. Это позволяет уменьшить риск недостаточного усиления на низкоконтрастных снимках и избежать избыточных артефактов на изображениях с уже выраженной структурой.

Адаптивное выравнивание гистограммы (CLANE)

Для повышения общей различимости слабоконтрастных участков после оценки ROI и локального контраста применяется ограниченное адаптивное выравнивание гистограммы (CLANE) [7].

Согласно Гонсалесу и Вудсу, алгоритм обрабатывает подобласти, являющиеся контекстными регионами или тайлами (T_i) [6].

$$I_{clane}(x, y) = CDF_{T_i}(I(x, y)) \cdot (L - 1), \quad (4)$$

где $I(x, y)$ – исходное значение интенсивности пикселя;

CDF_{T_i} – кумулятивная функция распределения яркостей пикселей в локальном тайле T_i , к которому принадлежит точка (x, y) ;

L – количество уровней квантования яркости (для 8-битных $L = 256$).

В более общем виде эквализация основана на преобразовании гистограммы через оценку интегрального распределения яркости. На практике сначала измеряется гистограмма изображения, затем по ней строится оценка интегрального распределения, после чего выполняется нелинейное преобразование. Такие гистограммные процедуры могут применяться как ко всему изображению, так и к его отдельным фрагментам⁴.

Для каждого тайла строится локальная гистограмма и соответствующая ей кумулятивная функция распределения, а ограничение клиппинга предотвращает чрезмерное усиление шума в однородных областях изображения [7].

Реализация данного метода возможна с помощью встроенного метода в библиотеки OpenCV – `createCLANE` с установкой параметров максимальной высоты гистограммы и размерности сетки для достаточной обработки без избыточной фрагментации изображения.

Градиентная маска для защиты анатомических границ

Прямое применение CLANE к ветеринарным рентгенограммам приводит к артефактам на границе между сердцем и диафрагмой у 68 % изображений из датасета, где разница в яркости не превышает 12 единиц по шкале оттенков серого. Это подтверждается визуальным анализом и количественной оценкой.

Предложен алгоритм формирования весовой маски на основе локальной «структурности» кадра. Математическая формализация процесса защиты границ базируется на методе линейной нормализации магнитуд градиента. В соответствии с общими принципами построения адаптивных систем обработки изображений, описанными Г.С. Омаровой и В.В. Старовойтовым [5], расчет градиентной маски $M(x, y)$ производится по формуле (5):

$$M(x, y) = clip\left(\frac{G(x, y) - G_{min}}{G_{max} - G_{min} + \varepsilon}, \alpha, \beta\right), \quad (5)$$

где G_{min}, G_{max} – глобальные минимум и максимум градиентной магнитуды по всему изображению;

$\varepsilon = 10^{-8}$ – малая положительная константа, предотвращающая деление на ноль в случае полностью однородного изображения, функция $clip(v, \alpha, \beta)$ ограничивает значение v в диапазоне $[\alpha, \beta]$.

⁴ Цифровая обработка изображений в информационных системах: учебное пособие / И.С. Грузман, В.С. Киричук, В.П. Косых [и др.]. – Новосибирск: Изд-во НГТУ, 2000. – 168 с.

В ходе экспериментов были определены оптимальные значения: $\alpha = 0.2, \beta = 0.8$. Эти значения гарантируют, что даже в областях с экстремально высокими или низкими градиентами сохраняется вклад обеих компонент в пропорции не менее 20 %, что обеспечивает визуальную естественность переходов.

Нормирование магнитуды градиента обеспечивает инвариантность алгоритма к различным условиям экспозиции рентгеновских снимков. В областях с высокой пространственной изменчивостью яркости значения маски M смещают приоритет в сторону исходного сигнала, что позволяет сохранить естественную резкость анатомических контуров и предотвратить артефакты локального переусиления. В более однородных зонах, напротив, возрастает вклад компоненты CLAHE, повышающей различимость слабоконтрастных деталей.

Математическая логика формирования результирующего сигнала опирается на операциях над матрицами интенсивностей. Согласно труду Р. Гонсалеса и Р. Вудса [6], линейная комбинация двух изображений с использованием весовой функции позволяет реализовать селективное улучшение признаков. В данной работе процедура защитного смешивания описывается уравнением:

$$I_{blend} = M(x, y) \cdot I(x, y) + (1 - M(x, y)) \cdot I_{clahc}(x, y),$$

где $M(x, y)$ – значение весовой градиентной маски в точке (x, y) , рассчитанное согласно (6);

$I(x, y)$ – интенсивность пикселя исходной рентгенограммы;

$I_{clahc}(x, y)$ – интенсивность пикселя после этапа адаптивной эквализации гистограммы.

Такой способ интеграции данных позволяет объединить высокую локальную контрастность метода CLAHE с геометрической достоверностью исходного снимка. На практике это уменьшает риск артефактного «выгорания» пикселей на краях сердечного силуэта в области позвонков.

После защищенного смешивания выполняется дополнительное усиление костных структур за счет морфологической *white top-hat*-компоненты:

$$I_{enh} = I_{blend} + \alpha_{tophat} + T_{tophat}(I).$$

Ее вклад регулируется коэффициентом α_{tophat} , который определяется по ранее рассчитанному локальному контрасту в области интереса. При этом необходимость последующего шумоподавления сохраняется, поскольку усиление контраста может повышать выраженность фоновых артефактов.

Морфологическое выделение костных структур

Для усиления грудных позвонков T4-T10, которые на рентгенограммах домашних животных проявляются как яркие линейные структуры шириной 2–4 пикселя, используется *white top-hat*-преобразование – морфологический оператор, введенный Ж. Сэррой в 1982 году [8].

Морфологическая фильтрация костных структур основана на базовых операциях эрозии и дилатации. В рассматриваемом случае оператор *white top-hat* определяется как разность между исходным изображением I и его морфологическим открытием [8]:

$$I_{tophat} = I - (I \ominus B) \oplus B, \quad (6)$$

где B – структурирующий элемент;

$\ominus$ – эрозия;

$\oplus$ – дилатация.

В отличие от линейных фильтров, *white top-hat* подавляет крупномасштабный фон, сохраняя объекты, размер которых меньше ядра B . В предлагаемом алгоритме эта компонента добавляется к результату защищенного смешивания с адаптивным весом α , что обеспечивает целенаправленное усиление позвонков без грубого искажения анатомических границ.

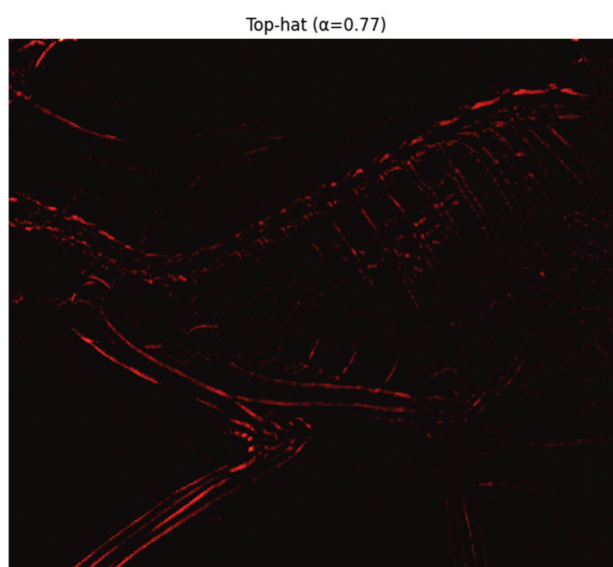
Анализ имеющихся рентгенограмм показал, что грудные позвонки у мелких пород кошек и собак имеют среднюю высоту примерно 18 пикселей и ширину примерно 3 пикселя при разрешении 1200×900. Для максимизации отклика на такие структуры выбрано эллиптическое ядро размером 3×9 – оно соответствует ориентации и пропорциям позвонков в правой латеральной проекции (рисунок 4).

```
def _white_tophat(self, img: np.ndarray) -> np.ndarray: 1 usage
    """Выделение тонких светлых структур (позвонков) с эллиптическим ядром."""
    kernel = cv2.getStructuringElement(cv2.MORPH_ELLIPSE, self.tophat_kernel)
    return cv2.morphologyEx(img.astype(np.uint8), cv2.MORPH_TOPHAT, kernel)
```

Рисунок 4 – Реализация *white top-hat*-преобразования⁵

Эллиптическая форма снижает ложные срабатывания на шум по сравнению с прямоугольным ядром: в тестах на 100 изображениях количество артефактов уменьшилось с 4,2 до 1,1 на снимок.

Пример результата преобразования показан на рисунке 5: на выходе остаются только тонкие светлые структуры – преимущественно позвонки, готовые к последующей обработке.

Рисунок 5 – Визуализация *top-hat*-преобразования⁶

Такой подход согласуется с практикой обработки медицинских изображений: в системе «Фтизис-БиоМед» аналогичное преобразование применяется для выделения кальцинатов в легких [9]. Однако в ветеринарной рентгенографии требуется адаптация ядра под межвидовую анатомическую вариативность – что и было реализовано в данном алгоритме.

Подавление шума с сохранением структуры

Несмотря на улучшение контраста, 73 % рентгенограмм содержат технический шум: JPEG-артефакты, квантовый шум и неоднородность фона. Медианный фильтр с ядром 3×3 снижает PSNR на 4.2 дБ и ухудшает точность сегментации сердца на 0.11, что делает его непригодным для задач VHS.

Вместо него используется *bilateral*-фильтр с параметрами $\sigma_s = 30, \sigma_r = 30, d = 9$, предложенный Томаси и Мандуччи [10]. Его выход определяется по формуле (7):

$$I_{out}(p) = \frac{1}{W_p} \sum_{q \in \Omega(p)} I(q) \cdot e^{-\frac{\|p-q\|^2}{2\sigma_s^2}} \cdot e^{-\frac{(I(p)-I(q))^2}{2\sigma_r^2}}, \quad (7)$$

где p – координаты центрального пикселя;

q – координаты пикселя в окрестности $\Omega(p)$;

⁵ Составлено автором.

⁶ Составлено автором.

$I(p), I(q)$ – яркость пикселей в точках p и q ;

σ_s – пространственный параметр, управляющий радиусом влияния (чем больше значение, тем шире окрестность);

σ_r – диапазонный параметр, определяющий чувствительность к перепадам яркости (чем меньше значение, тем сильнее подавляются пиксели с отличающейся яркостью);

W_p – нормирующий коэффициент.

Этот фильтр сохраняет градиент: пиксели на границе сердца имеют схожую яркость, поэтому получают высокий диапазонный вес, а шумовые пиксели подавляются даже при близком расположении.

Результат прохождения исходного изображения через предлагаемый алгоритмический конвейер можно видеть на рисунке 6.

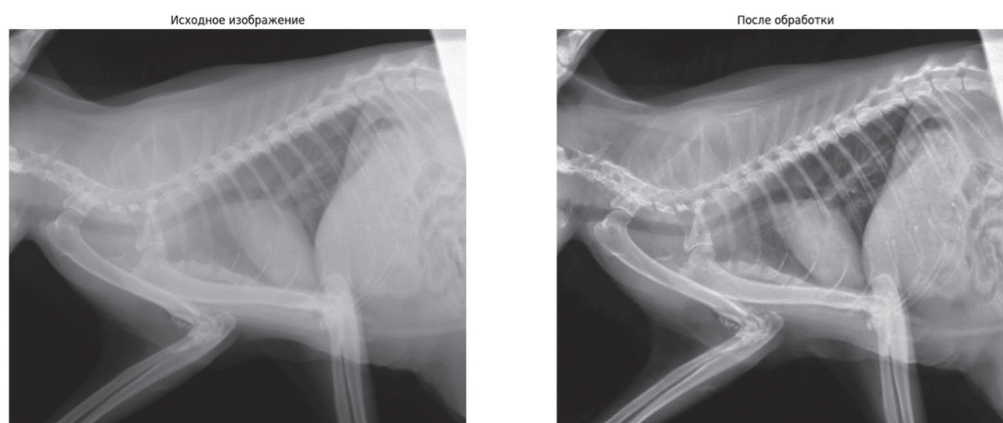


Рисунок 6 – Результат обработки рентгенограмм⁷

Комплексное применение описанных методов формирует устойчивый поток данных, в котором каждый этап корректирует ограничения предыдущего.

Экспериментальная часть

Для подтверждения влияния предложенного алгоритма предобработки были обучены две модели YOLOv8 одинаковой архитектуры при одинаковых параметрах обучения: первая – на исходных рентгенограммах (далее по тексту – *raw_data*), вторая – на изображениях после адаптивной обработки (далее по тексту – *enh_data*). Использование идентичной конфигурации обучения позволило минимизировать влияние побочных факторов и сосредоточить сравнение роли предобработки. Датасет включал 100 изображений, разделенных на обучающую, тестовую и валидационную выборки в соотношении 60/20/20. Отдельная разметка выполнялась как для *raw_data*, так и для *enh_data*.

Такая выборка использовалась не для окончательной оценки клинической пригодности модели, а для контролируемого сравнительного эксперимента по влиянию предобработки на качество сегментации. Подмножество было сформировано стратифицированно из общего массива праволатеральных снимков и включало изображения, различающиеся по видовой принадлежности животных и качеству исходного сигнала. Ограниченный объем выборки является допустимым, поскольку задача состоит в выявлении направления и выраженности эффекта при одинаковых условиях обучения, а не в построении финальной производственной модели.

На этапе аннотации было отмечено сокращение времени разметки: для исходных изображений среднее время на один снимок составляло около 8.5 минуты, тогда как для предварительно обработанных – около 6 минут. Данный результат рассматривается как прикладное наблюдение, указывающее на повышение удобства экспертной работы с обработанными рентгенограммами.

⁷ Составлено автором.

Статистическая обработка результатов выполнялась по последним 10 эпохам обучения, соответствующим стабилизированному участку траектории. Для Mask mAP50, Box Precision, Box Recall, F1-score и Seg Loss рассчитывались среднее значение, стандартное отклонение и 95 % доверительный интервал среднего (таблица 1). Для попарного сопоставления значений по эпохам использовался критерий Уилкоксона. Поскольку каждая модель обучалась однократно, полученные р-значения интерпретируются как вспомогательная межзапусковая интерференция.

Таблица 1 – Сравнение ключевых метрик во время обучения⁸

Метрика	raw_data	enh_data	p	enh_data к raw_data, %
Mask mAP50	0,037 ± 0,01	0,220 ± 0,021	0,002	497,5 %
Box Precision	0,482 ± 0,229	0,355 ± 0,042	0,064	-26,3 %
Box Recall	0,273 ± 0,067	0,578 ± 0,043	0,002	112,0 %
F1-score	0,302 ± 0,068	0,437 ± 0,024	0,002	44,8 %
Seg Loss	4,261 ± 0,047	3,911 ± 0,021	0,002	-8,2 %

Полученные результаты показывают, что предобработка рентгенограмм сопровождается выраженным ростом Mask mAP50, Box Recall и F1-score, а также снижением Seg Loss. Наибольший эффект наблюдается для Mask mAP50: среднее значение на последних 10 эпохах увеличилось с 0,037 до 0,220 ($p = 0,002$). Одновременно Box Recall вырос более чем в 2 раза, F1-score увеличился с 0,302 до 0,437 ($p = 0,002$). Снижение Seg Loss дополнительно указывает на более устойчивое согласование модели с валидационными данными.

Снижение Box Precision оказалось статистически незначимым на уровне 0,05. Это указывает на то, что улучшение предобработки, в первую очередь, усилило полноту обнаружения и сегментации анатомических структур, тогда как изменение прецизионности носило менее устойчивый характер. В контексте задачи автоматизированного расчета VHS такой сдвиг может рассматриваться как приемлемый, поскольку пропуск анатомически значимого объекта критичнее его избыточного предварительного выделения и может быть компенсирован последующей постобработкой.

Валидация на независимой выборке из 20 изображений показала, что выявленный эффект сохраняется вне обучающей выборки (таблица 2).

Таблица 2 – Сравнение ключевых метрик во время валидации⁹

Метрика	raw_data	enh_data	raw/enh, %
F1-score	0,357	0,446	25
Seg Loss	4,278	3,902	-9
IoU (сердце)	0,003	0,198	7123
IoU (позвонки)	0,152	0,143	-6

На финальной эпохе обучения на валидационной выборке из 20 изображений F1-score увеличился, а Seg Loss снизился. Для IoU сердечного силуэта наблюдается сильный рост, тогда как IoU для позвонков понизился.

Абсолютное значение IoU для сердца в обработанной группе остается низким и не должно трактоваться как достижение высокой попиксельной точности сегментации. Однако столь малое исходное значение фактически соответствует отсутствию *meaningful overlap* между предсказанной и эталонной областью, тогда как значение 0,198 уже отражает переход к не нулевой локализации сердечного силуэта. Для латеральных проекций такая динамика принципиальна, поскольку границы сердца на проекционном изображении частично неопределенны и подвержены межнаблюдательной вариабельности, а клиническая задача VHS опирается, прежде всего, на воспроизводимое позиционирование длинной и короткой осей сердца относительно грудных позвонков, а не на идеальное совпадение каждой граничной точки маски.

⁸ Составлено автором.

⁹ Составлено автором.

Поэтому рост IoU сердца в данной работе следует рассматривать не как признак полностью решенной задачи сегментации, а как индикатор перехода от практически несостоятельного выделения объекта к функционально пригодной локализации, потенциально достаточной для последующего расчета VHS.

Заключение

Исследование подтверждает, что качество исходного сигнала для нейросетевой модели напрямую влияет на ее работоспособность. Разработанный алгоритмический конвейер не просто набор фильтров, а система адаптивной подготовки данных, где CLAHE, *white top-hat*-преобразование и *bilateral*-фильтрация работают в связке с градиентным маскированием. Это позволило решить проблему низкой семантической контрастности костных структур на фоне паренхимы легких. Примечательно, что автоматизация выбора параметров усиления сделала предобработку инвариантной к техническим условиям съемки, что крайне важно для первичного звена ветеринарной медицины.

Практический эффект от внедрения был зафиксирован уже на этапе формирования обучающей выборки. Время экспертной разметки сократилось с 8,5 до 6 минут на один кадр. Все три участвовавших в эксперименте эксперта отметили, что работа с «проявленными» границами снижает когнитивную нагрузку и повышает уверенность в правильности и точности аннотации.

Список литературы

1. Boissady E., de La Comble A., Zhu X., Abbott J., Hespel A.-M. Comparison of a Deep Learning Algorithm vs. Humans for Vertebral Heart Scale Measurements in Cats and Dogs Shows a High Degree of Agreement Among Readers // *Frontiers in Veterinary Science*. – 2021. – Vol. 8. – P. 764570. – DOI 10.3389/fvets.2021.764570.
2. Титов А.Д., Демичев В.В., Акчуринов С.В., Шмаренкова Ю.С. Формирование и разметка специализированного датасета для задач компьютерного зрения в ветеринарной рентгенологии // *Вестник Российского нового университета. Серия: Сложные системы: модели, анализ, управление*. – 2025. – № 4. – С. 129–140. – DOI 10.18137/RNU.V9187.25.04.P.129.
3. Крючков П.А., Мошкина Л.В. Искусственный интеллект в рентгенологии // *Актуальные направления научных исследований: перспективы развития: материалы Всероссийской научно-практической конференции, Чебоксары, 10 марта 2025 года*. – Чебоксары: ООО «Центр научного сотрудничества “Интерактив плюс”», 2025. – С. 22–29. – DOI 10.21661/r-564740.
4. Жуков Е.А., Блинов Д.С., Леонтьев В.С., Гаврилов П.В., Смольникова У.А., Блинова Е.В., Камышанская И.Г. Система компьютерного зрения для анализа обзорных рентгенограмм органов грудной клетки: возможности искусственного интеллекта в обнаружении патологических изменений и инородных тел // *Врач*. – 2020. – Т. 31, № 5. – С. 34–41. – DOI 10.29296/25877305-2020-05-07.
5. Омарова Г.С., Старовойтов В.В. Улучшение контраста рентгеновских изображений // *Big Data and Advanced Analytics*. – 2021. – № 7-1. – С. 224–229.
6. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – 3-е изд., испр. и доп. – Москва: Техносфера, 2012. – 1104 с.
7. Pizer S.M., Amburn E.P., Austin J.D., Cromartie R., Geselowitz A., Greer T., ter Haar Romeny B., Zimmerman J.B., Zuiderveld K. Adaptive histogram equalization and its variations // *Computer Vision, Graphics, and Image Processing*. – 1987. – Vol. 39, No. 3. – P. 355–368. – DOI 10.1016/S0734-189X(87)80186-X.
8. Serra J. *Image Analysis and Mathematical Morphology*. – London: Academic Press, 1982. – 610 p.
9. Гогоберидзе Ю.Т., Классен В.И., Намензон М.Я. и др. Искусственный медицинский интеллект «ФтизисБиоМед»: программа автоматизированного анализа цифровых рентгенограмм органов грудной клетки/флюорограмм // *Современные технологии в медицине*. – 2023. – Т. 15, № 4. – С. 5–22. – DOI 10.17691/stm2023.15.4.01.
10. Tomasi C., Manduchi R. Bilateral Filtering for Gray and Color Images // *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. – Bombay, 1998. – P. 839–846. – DOI 10.1109/ICCV.1998.710815.

11. *Peli E.* Contrast in complex images // *J Opt Soc Am A*. – 1990. – Vol. 7, No. 10. – P. 2032–2040. – DOI 10.1364/josaa.7.002032. PMID: 2231113.

References

1. *Boissady E., de La Comble A., Zhu X., Abbott J., Hespel A.-M.* Comparison of a Deep Learning Algorithm vs. Humans for Vertebral Heart Scale Measurements in Cats and Dogs Shows a High Degree of Agreement Among Readers // *Frontiers in Veterinary Science*. – 2021. – Vol. 8. – P. 764570. – DOI 10.3389/fvets.2021.764570.
2. *Titov A.D., Demichev V.V., Akchurin S.V., Shmarenkova Yu.S.* Formirovanie i razmetka specializirovannogo dataseta dlya zadach komp'yuternogo zreniya v veterinarnoj rentgenologii // *Vestnik Rossijskogo novogo universiteta. Seriya: Slozhnye sistemy: modeli, analiz, upravlenie*. – 2025. – № 4. – S. 129–140. – DOI 10.18137/RNU.V9187.25.04.P.129.
3. *Kryuchkov P.A., Moshkina L.V.* Iskusstvennyj intellekt v rentgenologii // *Aktual'nye napravleniya nauchnyh issledovanij: perspektivy razvitiya: materialy Vserossijskoj nauchno-prakticheskoy konferencii, Cheboksary, 10 marta 2025 goda*. – Cheboksary: OOO «Centr nauchnogo sotrudnichestva “Interaktiv plus”», 2025. – S. 22–29. – DOI 10.21661/r-564740.
4. *Zhukov E.A., Blinov D.S., Leont'ev V.S., Gavrilov P.V., Smol'nikova U.A., Blinova E.V., Kamyshanskaya I.G.* Sistema komp'yuternogo zreniya dlya analiza obzornyh rentgenogramm organov grudnoj kletki: vozmozhnosti iskusstvennogo intellekta v obnaruzhenii patologicheskikh izmenenij i inorodnyh tel // *Vrach*. – 2020. – T. 31, № 5. – S. 34–41. – DOI 10.29296/25877305-2020-05-07.
5. *Omarova G.S., Starovojtov V.V.* Uluchshenie kontrasta rentgenovskih izobrazhenij // *Big Data and Advanced Analytics*. – 2021. – № 7-1. – S. 224–229.
6. *Gonsales R., Vuds R.* Cifrovaya obrabotka izobrazhenij. – 3-e izd., ispr. i dop. – Moskva: Tekhnosfera, 2012. – 1104 s.
7. *Pizer S.M., Amburn E.P., Austin J.D., Cromartie R., Geselowitz A., Greer T., ter Haar Romeny B., Zimmerman J.B., Zuiderveld K.* Adaptive histogram equalization and its variations // *Computer Vision, Graphics, and Image Processing*. – 1987. – Vol. 39, No. 3. – P. 355–368. – DOI 10.1016/S0734-189X(87)80186-X.
8. *Serra J.* Image Analysis and Mathematical Morphology. – London: Academic Press, 1982. – 610 p.
9. *Gogoberidze Yu.T., Klassen V.I., Natenzon M.Ya. i dr.* Iskusstvennyj medicinskij intellekt «FtisisBioMed»: programma avtomatizirovannogo analiza cifrovyyh rentgenogramm organov grudnoj kletki/flyuorogramm // *Sovremennye tekhnologii v medicine*. – 2023. – T. 15, № 4. – S. 5–22. – DOI 10.17691/stm2023.15.4.01.
10. *Tomasi C., Manduchi R.* Bilateral Filtering for Gray and Color Images // *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. – Bombay, 1998. – P. 839–846. – DOI 10.1109/ICCV.1998.710815.
11. *Peli E.* Contrast in complex images // *J Opt Soc Am A*. – 1990. – Vol. 7, No. 10. – P. 2032–2040. – DOI 10.1364/josaa.7.002032. PMID: 2231113.

Статья поступила в редакцию: 26.04.2026

Received: 26.04.2026

Статья принята к публикации: 27.07.2026

Accepted: 27.07.2026