

УДК 004.032.26

DOI: [10.26102/2310-6018/2026.60.9.005](https://doi.org/10.26102/2310-6018/2026.60.9.005)

Адаптивная модель предобработки изображений при классификации чертежей строительных металлоконструкций

В.Д. Колесников✉, П.С. Кабальянц

*Белгородский государственный технологический университет им. В.Г. Шухова,
Белгород, Российская Федерация*

Резюме. Постоянный рост рынка инженерных услуг и наличие значительного объема архивной бумажной документации актуализируют задачу перевода технических чертежей и спецификаций к ним в цифровой формат. Анализ состояния исследований демонстрирует активное применение распознающих и генеративных моделей для извлечения ключевых характеристик чертежей, что упрощает их интеграцию в системы компьютерного моделирования. Для повышения эффективности нейросетевых моделей применяется предобработка изображения, однако ввиду особенностей структуры и возможных источников чертежей, подобный подход несет с собой риск потери информации или искажения данных. Авторами предлагается адаптивная модель предобработки, ориентированная на системы классификации чертежей строительных металлоконструкций. Данная модель включает в себя операции масштабирования, фильтрации шумов, бинаризации и восстановления цвета информативных пикселей, где каждая операция применяется в зависимости от значения критериев, определяющих необходимость ее вызова. В рамках исследований определяются алгоритмы и методы, подходящие для реализации вышеописанных операций: метод интерполяции Ланцоша, взвешенный медианный фильтр с переключением и метод бинаризации Вольфа. Авторами предлагается набор критериев оценки эффективности предобработки при классификации чертежей строительных металлоконструкций. Ряд проведенных экспериментов показывает, что разработанная адаптивная модель предобработки дает больший прирост точности классификации по сравнению со стандартной предобработкой на основе таких же алгоритмов или полным ее отсутствием. Делаются выводы о возможности дальнейшего внедрения модели в системы классификации и распознавания деталей чертежей строительных металлоконструкций.

Ключевые слова: цифровые изображения, технические чертежи, адаптивная предварительная обработка, масштабирование, фильтрация шумов, бинаризация, классификация чертежей.

Ссылка на статью: Колесников В.Д., Кабальянц П.С. Адаптивная модель предобработки изображений при классификации чертежей строительных металлоконструкций. *Моделирование, оптимизация и информационные технологии*. 2026;14(9). URL: <https://moitvvt.ru/ru/journal/article?id=2452> DOI: 10.26102/2310-6018/2026.60.9.005

An adaptive image preprocessing model for classifying drawings of building metal structures

V.D. Kolesnikov✉, P.S. Kabalyants

*Belgorod State Technological University named after V.G. Shukhov, Belgorod,
the Russian Federation*

Abstract. The constant growth of the engineering services market and the significant volume of archival paper documentation make the task of digitizing technical drawings and their specifications urgent. An analysis of the current state of research demonstrates the active use of recognition and generative models to extract key drawing characteristics, which facilitates their integration into computer modeling systems. Image preprocessing is used to improve the efficiency of neural network models; however, due to the specific structure and potential sources of drawings, this approach carries the risk of information

loss or data corruption. The authors propose an adaptive preprocessing model designed for classification systems for drawings of building metal structures. This model includes scaling, noise filtering, binarization, and color restoration of informative pixels, where each operation is applied depending on the values of criteria that determine whether it is required. The research identifies algorithms and methods suitable for implementing the above-described operations: the Lanczos interpolation method, the switching weighted median filter, and the Wolf binarization method. The authors propose a set of criteria for evaluating the effectiveness of preprocessing in classifying drawings of building metal structures. A series of experiments shows that the developed adaptive preprocessing model yields a greater increase in classification accuracy compared to standard preprocessing based on the same algorithms or no preprocessing at all. Conclusions are drawn regarding the potential for further implementation of the model in systems for classification and recognition of parts in drawings of building metal structures.

Keywords: digital images, technical drawings, adaptive preprocessing, scaling, noise filtering, binarization, classification of drawings.

For citation: Kolesnikov V.D., Kabalyants P.S. An adaptive image preprocessing model for classifying drawings of building metal structures. *Modeling, Optimization and Information Technology*. 2026;14(9). (In Russ.). URL: <https://moitvvt.ru/ru/journal/article?id=2452> DOI: 10.26102/2310-6018/2026.60.9.005

Введение

Аналитические исследования современных рынков инженерных услуг предсказывают ежегодный прирост капитала крупнейших компаний на уровне 6,3 % в период с 2023 по 2033 год¹, что свидетельствует о стабильном спросе на товары и услуги инженерной сферы. В связи с этим сохраняет актуальность область разработки, анализа и сопровождения технической документации инженерных проектов. Основным форматом представления такой документации остаются технические чертежи и спецификации к ним, однако их хранение и обработка активно переходят в цифровую форму. Данный переход обуславливает необходимость преобразования существующих аналоговых чертежей в цифровой формат, что включает задачи анализа изображений, их классификации, распознавания деталей и текста, перевода в формат, с которым взаимодействуют автоматизированные системы, и др. Для решения вышеописанных задач активно ведутся соответствующие исследования и разработки. Например, в работе [1] предложен метод генерации трехмерных моделей по изображениям планов этажей с использованием векторизации и распознавания графических примитивов. В исследовании [2] представлена многоэтапная модель преобразования изображений чертежей в форматы САД-систем. В работах [3, 4] также рассматриваются методы распознавания и векторизации.

Для повышения эффективности подобных алгоритмов во многих из рассмотренных систем к изображениям применяются методы предварительной обработки [2, 3]. Подобный подход позволяет повысить эффективность обработки цифровых изображений, однако также несет риски потери полезной информации в результате возникновения артефактов, сглаживания и других модификаций исходного изображения.

В рамках данной научной работы предлагается адаптивная модель предобработки изображений технических чертежей строительных металлоконструкций, включающая этапы масштабирования, фильтрации шумов, изменения цветовой модели, бинаризации и восстановление цвета информативных пикселей. Выбор вышеописанных операций

¹ *Engineering Services Market Size & Share Report 2026–2033*. Grand View Research. URL: <https://www.grandviewresearch.com/industry-analysis/engineering-services-market-report> (дата обращения: 19.06.2026).

обуславливается характеристиками обрабатываемых изображений и требованиями к высокой точности классификации для используемых моделей.

Материалы и методы

Входящее изображение чертежа можно представить в виде трехмерного тензора пикселей – цветовых точек, формирующих изображение на устройствах отображения:

$$I \in \mathbb{R}^{H \times W \times C}, \quad (1)$$

где W и H – высота и ширина изображения соответственно, C – количество цветовых каналов, определяемое используемой цветовой моделью (например, RGB, CMYK, и др.). Цифровое изображение I в данном случае можно рассматривать как элемент множества всех возможных изображений заданной размерности $\mathbb{R}^{H \times W \times C}$. Математически изображение также можно представить в виде функции:

$$I(x, y, c) = p, \quad x \in [0, H - 1], \quad y \in [0, W - 1], \quad c \in [0, C - 1], \quad x, y, c \in \mathbb{Z}, \quad (2)$$

где значение $p \in \mathbb{R}$ определяет интенсивность (или значение цвета) пикселя с координатами (x, y) в канале c . Большинство операций обработки изображений могут быть формализованы как отображения между пространствами изображений:

$$f: \mathbb{R}^{H \times W \times C} \rightarrow \mathbb{R}^{H' \times W' \times C'}, \quad (3)$$

где $I' = f(I)$ – результат преобразования. Параметры H' , W' и C' могут как совпадать с исходными значениями H , W и C , так и отличаться от них в зависимости от типа выполняемой операции. Обобщенно предварительную обработку изображения можно описать функцией:

$$I' = f_p(I), \quad f_p = f_1 \circ f_2 \circ \dots \circ f_n(I), \quad (4)$$

где f_i – i -я по счету операция преобразования изображения, а оператор $\circ$ представляет собой композицию функций.

Особенность классификации чертежей строительных металлоконструкций заключается в комплексном подходе к анализу входящих изображений, предполагающем использования нескольких компонентов обработки различных элементов чертежа. В частности, высокую значимость при определении класса изделия на чертеже вносят графические детали, текстовые надписи и идентификационная информация, указываемая в основной надписи чертежа. Корректное извлечение вышеописанной информации, ее преобразование в вектор информативных признаков и дальнейшее определение класса изделия предполагают применение нейросетевых моделей, решающих задачи распознавания объектов на изображении, оптического распознавания текста и классификации по признаковому пространству.

Подобные модели чувствительны к качеству входящих данных, что в свою очередь предполагает применение на этапе предобработки таких операций, которые минимизируют негативное влияние на качество изображения, одновременно подводя последнее к формату, эффективно обрабатываемому нейросетевыми моделями.

Поскольку источниками изображений чертежей могут быть как аналоговые, так и цифровые системы, структура и характеристики входящих чертежей могут различаться, что в свою очередь актуализирует адаптивный подход к их предварительной обработке, уменьшающий вероятность искажения или полной потери информации по сравнению со стандартной предобработкой, срабатывающей всегда.

Авторами предлагается адаптивная многоэтапная модель предобработки, описываемая следующей функцией:

$$f_p(I) = F_{bin} \circ F_{noise} \circ f_{HSV} \circ F_{res}(I), \quad (5)$$

где F_{res} – адаптивное масштабирование изображения, f_{HSV} – перевод изображения в цветное пространство HSV, F_{noise} – адаптивная фильтрация шумов, F_{bin} – адаптивная бинаризация с возможным восстановлением цвета информативных пикселей. Последовательность преобразований изображения данной моделью может быть записана как:

$$I \xrightarrow{f_{res}} I_{res} \xrightarrow{f_{HSV}} I_{HSV} \xrightarrow{f_{noise}} I_{noise} \xrightarrow{f_{bin}} I', \quad (6)$$

$$I_{res} \in \mathbb{R}^{H' \times W' \times C}, \quad I_{HSV}, I_{noise} \in \mathbb{R}^{H' \times W' \times 3}, \quad I' \in \mathbb{R}^{H' \times W' \times C'}.$$

Первой из рассматриваемых операций модели предобработки является адаптивное масштабирование. Применение данной операции обосновывается наличием строгих или рекомендованных размеров изображения для многих распознающих алгоритмов. Несмотря на то, что более современные модели (например, YOLO²) выполняют масштабирование изображения самостоятельно, применение данной операции на этапе предобработки позволяет применять большее количество алгоритмов на этапе анализа чертежа. Помимо этого, приведение обрабатываемых изображений к формату, наиболее близкому к данным, на которых обучаются распознающие модели, позволяет уменьшить вариативность данных и повысить эффективность обучения для заданного размера датасета.

Пусть $W^* \times H^*$ – рекомендованные размеры входящего изображения для используемой комплексной модели классификации, $W \times H$ – размеры входящего изображения. Тогда функция адаптивного масштабирования описывается как:

$$F_{res}(I(x, y, c)) = I_{res}(x, y, c) = \begin{cases} f_{res}(I(x, y, c)); & \text{если } W^* \neq W \vee H^* \neq H, \\ I(x, y, c); & \text{если } W^* = W \wedge H^* = H. \end{cases} \quad (7)$$

Здесь f_{res} – используемый алгоритм масштабирования. Поскольку ключевым требованием к разрабатываемой модели является ее точность, а время работы не является высоко значимым параметром, было проведено исследование по оценке алгоритмов масштабирования на предмет сохранения качества изображения при обработке. Рассматривались методы усреднения, наибольших, наименьших и медианных значений [5], а также билинейная, бикубическая [6, 7], В-сплайн интерполяции [8] и интерполяция Ланцоша [9], среди которых метод Ланцоша показал наилучшее сохранение общего качества и детализации, в связи с чем было принято решение об его использовании в адаптивной модели предобработки. В таком случае функцию f_{res} можно записать следующим образом:

$$f_{res}(I(x, y, c)) = \sum_{i=\lfloor x^* \rfloor - a + 1}^{\lfloor x^* \rfloor + a} \sum_{j=\lfloor y^* \rfloor - a + 1}^{\lfloor y^* \rfloor + a} I(i, j, c) \cdot L(x^* - i) \cdot L(y^* - j), \quad (8)$$

$$x^* = x' \cdot \frac{H}{H'}, \quad y^* = y' \cdot \frac{W}{W'}.$$

$L(t)$ представляет собой функцию интерполяции Ланцоша:

$$L(t) = \begin{cases} \text{sinc } t \cdot \text{sinc} \left(\frac{t}{a} \right); & \text{если } |t| < a, \\ 0; & \text{если } |t| \geq a, \end{cases} \quad \text{sinc } t = \frac{\sin(\pi t)}{\pi t}. \quad (9)$$

Второй операцией в модели предобработки является перевод изображения в цветное пространство HSV. Поскольку исходные изображения технических чертежей

² YOLO Object Detection & Segmentation. Ultralytics. URL: <https://docs.ultralytics.com> (дата обращения: 19.06.2026).

могут содержать не только монохромные, но и цветные информативные признаки, перед выполнением этапов фильтрации и бинаризации целесообразно представить изображение в этом цветовом пространстве. В данной модели канал V может использоваться как представление интенсивности пикселей изображения при переходе к оттенкам серого, тогда как каналы H и S позволяют сохранить информацию об их оттенке и насыщенности.

Третьей рассматриваемой операцией в модели предобработки является адаптивная фильтрация шумов. Для определения подходящих алгоритмов фильтрации было проведено исследование, направленное на оценку наиболее часто встречающихся шумовых моделей [10] на изображениях технических чертежей строительных металлоконструкций. Дополнительно была выполнена симуляция передачи изображений по каналам связи с искусственными помехами для моделирования искажений при ухудшении качества сигнала. Оценка выполнялась на двух датасетах: аналоговом, который включал в себя фотографии и сканы бумажных чертежей, и CAD-датасете, который содержал изображения, экспортированные из CAD-систем. Полученные результаты позволили сделать следующие выводы:

1. Цифровые изображения, созданные в CAD-системах, как правило, не содержат шумов и дефектов.
2. Для изображений, полученных путем сканирования или фотографирования чертежей, уровень шума определяется состоянием оригинала и качеством оборудования.
3. Наиболее распространенным является фоновый шум (около 45 % рассмотренных изображений), проявляющийся в виде пятен, царапин, замятий и т. п.
4. Гауссовский шум может возникать вследствие аппаратных дефектов, однако при использовании современных устройств встречается крайне редко (менее 1 %).
5. При передаче данных в условиях помех, дефекты проявляются преимущественно в виде импульсных (около 98 %) и реже – в виде гауссовских шумов (менее 2 %).
6. В разрабатываемой адаптивной модели предобработки необходимо акцентировать внимание на подавлении преимущественно фоновых и импульсных шумов.

Для выбора подходящего фильтра для подавления импульсных шумов оценены наиболее распространенные алгоритмы фильтрации. В качестве критериев оценки выступали точность подавления искаженных шумом пикселей, описываемая как:

$$\eta_{filter} = \frac{k_{filtered}}{k_{noise}} \quad (10)$$

и минимизация обработки фильтром корректных пикселей, не искаженных шумом, описываемая как:

$$\eta_{error} = \frac{k_{error}}{H \cdot W - k_{noise}}. \quad (11)$$

Здесь $k_{filtered}$ – количество искаженных пикселей, обработанных шумом, k_{error} – количество корректных пикселей, ложно определенных как искаженные, и обработанных фильтром, k_{noise} – общее количество искаженных пикселей на изображении, $W \times H$ – размеры изображения.

В рамках исследования рассматривались гармонический, контргармонический [11], морфологический, α -усеченный фильтры и модификации медианного фильтра [12, 13]. Наилучшие результаты для описанных выше критериев показал взвешенный медианный фильтр с переключением, который и был выбран в качестве алгоритма фильтрации в разрабатываемой модели.

Поскольку ранее был осуществлен переход к цветовому пространству HSV, возможно применение векторного представления пикселя вместо вычисления медианы независимо для каждого канала. Исходное представление изображения в пространстве HSV описывается как:

$$I_{HSV} = (H(x, y), S(x, y), V(x, y)). \quad (12)$$

Для каждого пикселя формируется локальная окрестность:

$$\Omega_{x,y} = \{p_1, p_2, \dots, p_n\}, \quad p_i = (H_i, S_i, V_i). \quad (13)$$

Вычисляется векторное значение медианы локальной окрестности:

$$I_M(x, y) = \arg \min_{p_i \in \Omega_{x,y}} \sum_{p_j \in \Omega_{x,y}} d(p_i, p_j), \quad (14)$$

$$d(p_i, p_j) = \sqrt{\alpha_H \cdot d_H^2 + \beta_S \cdot (S_i - S_j)^2 + \gamma_V \cdot (V_i - V_j)^2}.$$

Здесь α_H , β_S и γ_V – весовые коэффициенты соответствующих каналов, d_H – циклическое расстояние между значениями оттенков H_i и H_j :

$$d_H = \min(|H_i - H_j|, 360 - |H_i - H_j|). \quad (15)$$

Затем вычисляется отклонение обрабатываемого пикселя от полученного векторного значения медианы:

$$\Delta I_M(x, y) = d(I(x, y), I_M(x, y)). \quad (16)$$

Итоговый медианный фильтр с переключением может быть задан следующим образом:

$$f_{noise}(I_{HSV}(x, y)) = \begin{cases} I_M(x, y); & \text{если } \Delta I_M(x, y) > \tau(x, y), \\ I_{HSV}(x, y); & \text{если } \Delta I_M(x, y) \leq \tau(x, y), \end{cases} \quad (17)$$

$$\tau(x, y) = k \cdot \sigma_V(x, y).$$

Здесь $\tau(x, y)$ – порог обнаружения импульсного шума, σ_V – стандартное отклонение значения V в окрестности пикселя, k – коэффициент чувствительности.

Идея взвешенности медианного фильтра предполагает использование весовых коэффициентов w_{ij} для каждого элемента окрестности обрабатываемого пикселя. Пусть p_c – центральный пиксель окрестности. Тогда значение веса пикселя p_{ij} определяется как:

$$w_{ij} = e^{\frac{d^2(p_c, p_{ij})}{2 \cdot \sigma_w^2}}, \quad (18)$$

где σ_w – коэффициент чувствительности. Медиана окрестности в таком случае определяется на расширенном векторе окрестности, где количество повторений каждого пикселя задается значением $\lceil w_{ij} \rceil$.

В разрабатываемой модели предобработки используется адаптивная фильтрация шумов, предполагающая наличие критерия оценки необходимости в фильтрации изображения. Поскольку основным типом подавляемого шума является импульсный, вышеописанный критерий может быть определен следующим отношением:

$$\eta_{noise} = \frac{|\{I_{HSV}(x, y) : \Delta I_M(x, y) > \tau(x, y)\}|}{W \cdot H}. \quad (19)$$

Иными словами, определяется отношение количества пикселей, медиана которых превышает порог, к общему количеству пикселей. Функция адаптивной фильтрации в таком случае описывается как:

$$F_{noise}(I_{HSV}(x, y)) = I_{noise}(x, y) = \begin{cases} f_{noise}(I_{HSV}(x, y)); & \text{если } \eta_{noise} > T_{noise}, \\ I_{HSV}(x, y); & \text{если } \eta_{noise} \leq T_{noise}. \end{cases} \quad (20)$$

Здесь T_{noise} – порог фильтрации. Подобный адаптивный подход позволяет опускать применение фильтрации к изображению, слабо искаженному импульсными шумами, тем самым минимизируя возможное его сглаживание и потерю информации. Ниже представлен краткий фрагмент реализации данного фильтра в виде псевдокода:

```
def adaptive_weighted_median(pixel, neighborhood, T, k):
    median = weighted_vector_median(neighborhood)
    sigma = local_std(neighborhood.v)
    hue = cyclic_distance(pixel.h, median.h)
    sat = abs(pixel.s - median.s)
    val = abs(pixel.v - median.v)
    deviation = w_h * hue + w_s * sat + w_v * val
    return median if deviation > T + k * sigma else pixel
```

Четвертым и завершающим этапом предлагаемой модели предобработки является адаптивная бинаризация с возможным восстановлением цвета информативных пикселей. Необходимость в применении бинаризации обуславливается высокой вероятностью наличия фоновых шумов на изображении. Возможность восстановления цвета информативных пикселей связана с особенностями обрабатываемых изображений.

Для рассматриваемых чертежей строительных металлоконструкций допускается наличие текстовых пометок и линий, цвет которых отличается, от стандартного для чертежей черно-белого формата. Бинаризация изображения переводит его в двухуровневое представление, наиболее подходящее для последующих этапов распознавания объектов, текста и идентификационной информации, однако при этом возможна потеря полезной информации, если цвет некоторых пикселей был отличным от черного или его оттенков.

Функция адаптивной бинаризации может быть описана следующим образом:

$$F_{bin}(I_{noise}(x, y)) = I_{bin}(x, y) = \begin{cases} F_{col} \circ f_{bin}(I_{noise}(x, y)); & \text{если } \eta_{bin} > T_{bin}, \\ I_{noise}(x, y); & \text{если } \eta_{bin} \leq T_{bin}. \end{cases} \quad (21)$$

Здесь F_{col} – функция восстановления цвета информативных пикселей, f_{bin} – функция бинаризации изображения, T_{bin} – порог бинаризации, η_{bin} – критерий оценки необходимости бинаризации, определяемый следующим образом:

$$\eta_{bin} = \alpha \cdot \eta_b - \beta \cdot \eta_c - \gamma \cdot \eta_e. \quad (22)$$

Здесь η_b – оценка степени неоднородности фона, η_c – оценка цветовой информативности изображения, η_e – оценка выраженности границ объектов, α , β , γ – весовые коэффициенты.

Оценка неоднородности вычисляется как:

$$\eta_b = \frac{1}{W \cdot H} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} \sigma_{loc}^2(x, y), \quad (23)$$

$$\sigma_{loc}^2(x, y) = \frac{1}{|\Omega_{x,y}|} \sum_{i,j \in \Omega_{x,y}} (V(i, j) - \mu_{\Omega})^2.$$

Здесь $\Omega_{x,y}$ – локальная окрестность пикселя, μ_{Ω} – среднее значение яркости окрестности.

Оценка цветовой информативности изображения определяется как:

$$\eta_c = \frac{|\{I_{HSV}(x,y): S(x,y) > T_s\}|}{W \cdot H}. \quad (24)$$

Оценка выраженности границ объектов определяется как:

$$\eta_e = \frac{1}{W \cdot H} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} |\nabla V(x, y)|, \quad (25)$$

где $\nabla V(x, y)$ – градиент яркости, вычисляемый оператором Собеля:

$$\nabla V(x, y) = \sqrt{(V(x, y) * K_x)^2 + (V(x, y) * K_y)^2}, \quad (26)$$

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad K_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

Оператор $*$ означает свёртку. Вышеописанные критерии позволяют точно оценивать необходимость в бинаризации входящего изображения чертежа.

Аналогично предыдущим этапам адаптивной модели предобработки, для бинаризации также было выполнено исследование по оценке эффективности существующих алгоритмов бинаризации [14] в подавлении фоновых шумов. Так же, как и для фильтрации импульсных шумов, использовались критерии η_{filter} и η_{error} . В рамках исследования рассматривались методы Оцу [15], Капура [16], Ниблака, Сауволя [17] и Вольфа [18], среди которых последний показал наилучшие результаты и был взят в качестве метода бинаризации в предлагаемой модели. Являясь локальным алгоритмом бинаризации, метод Вольфа вычисляет порог бинаризации для каждого пикселя с учетом его окрестности:

$$T(x, y) = (1 - k) \cdot \mu(x, y) + k \cdot M + k \cdot \frac{\sigma(x, y)}{\sigma_{max}} \cdot (\mu(x, y) - M), \quad (27)$$

где $\mu(x, y)$ – локальное среднее значение интенсивности, $\sigma(x, y)$ – локальное стандартное отклонение интенсивности, $k \in [0, 1]$ – коэффициент чувствительности, значения которого обычно находятся в диапазоне $[0, 2; 0, 5]$, M – глобальный минимум интенсивности на изображении, σ_{max} – глобальный максимум стандартного отклонения интенсивности на изображении. Итоговая бинаризация описывается функцией вида:

$$f_{bin}(I_{noise}(x, y)) = I'_{bin}(x, y) = \begin{cases} 255; & \text{если } V(x, y) > T(x, y), \\ 0; & \text{если } V(x, y) \leq T(x, y). \end{cases} \quad (28)$$

После выполнения бинаризации для изображения строится маска информативных пикселей:

$$\Lambda(x, y) = \begin{cases} 1; & \text{если } I'_{bin}(x, y) = 0, \\ 0; & \text{если } I'_{bin}(x, y) = 255. \end{cases} \quad (29)$$

Для всех пикселей в $I_{noise}(x, y)$, у которых $\Lambda(x, y) = 1$, анализируется значение цветового тона $H(x, y)$. Два значения цветового тона считаются принадлежащими одному кластеру, если:

$$|H_i - H_j| < T_h, \quad (30)$$

где T_h – порог различия цветовых тонов. В результате формируется множество цветовых кластеров:

$$C_h = \{c_1, c_2, \dots, c_n\}. \quad (31)$$

Адаптивная функция восстановления определяется мощностью данного множества кластеров:

$$F_{col}(I'_{bin}(x, y)) = I_{bin}(x, y) = \begin{cases} I'_{bin}(x, y); & \text{если } |C_h| = 1, \\ f_{col}(I'_{bin}(x, y)); & \text{если } |C_h| > 1. \end{cases} \quad (32)$$

Сама функция восстановления цвета f_{col} описывается как:

$$f_{col}(I'_{bin}(x, y)) = \begin{cases} I'_{bin}(x, y); & \text{если } M(x, y) = 0, \\ I_{noise}(x, y); & \text{если } M(x, y) = 1. \end{cases} \quad (33)$$

Реализация предложенной адаптивной бинаризации с последующим восстановлением цвета представлена ниже в виде краткой функции, написанной на псевдокоде:

```
def adaptive_binarization(image, window_size, k):
    binary = wolf_binarization(image, window_size, k)
    for component in detect_color_components(image):
        metrics = compute_metrics(component)
        score = weighted_score(metrics)
        if score >= RESTORE_THRESHOLD:
            binary = restore_component(binary, component)
    return binary
```

В рассматриваемой предметной области использование большого количества различных цветов на одном чертеже не встречается. В связи с этим применение анализа цветовых тонов в пространстве HSV является достаточным для решения поставленной задачи. В случае наличия большого числа близких оттенков могут использоваться методы кластеризации цветов или сегментации в пространствах CIELAB и CIECAM02 и др., однако в рамках данной модели подобные подходы не являются необходимыми.

Наиболее вычислительно затратными этапами предложенной модели являются адаптивная фильтрация шумов и локальная бинаризация методом Вольфа. При использовании окна фиксированного размера их вычислительная сложность линейно зависит от числа пикселей изображения и составляет $O(N)$, где N – количество пикселей, однако коэффициент пропорциональности существенно выше по сравнению с глобальными методами бинаризации вследствие необходимости вычисления локальных статистик.

Для сокращения времени обработки локальные средние значения и стандартные отклонения могут вычисляться с использованием интегральных изображений, позволяющих получать суммы интенсивностей внутри локальных окон за постоянное время. Такой подход уменьшает практические временные затраты без изменения результатов бинаризации и делает модель пригодной для обработки изображений большого разрешения.

Предложенная адаптивная модель позволяет эффективно подстраивать параметры изображения под формат, наиболее подходящий для используемых комплексных моделей анализа и классификации. Для ее сравнения с применением стандартных моделей предобработки или полного отсутствия данного этапа был сформирован набор критериев оценки информативности изображения.

Первая группа критериев направлена на оценку качества итогового изображения перед его обработкой комплексной моделью классификации. Здесь используются определенные ранее критерии оценки уровня импульсного шума η_{noise} и выраженности границ объектов η_e . Также применяется оценка потери информативных пикселей:

$$\vartheta_{inf} = 1 - \frac{|\{(x,y): \Lambda_I(x,y)=1 \wedge \Lambda'_I(x,y)=1\}|}{|\{(x,y): \Lambda_I(x,y)=1\}|},$$

$$\Lambda_I(x,y) = \begin{cases} 1; & \text{если } I(x,y) < T_{inf}, \\ 0; & \text{если } I(x,y) \geq T_{inf}, \end{cases} \quad \Lambda'_I(x,y) = \begin{cases} 1; & \text{если } I'(x,y) < T_{inf}, \\ 0; & \text{если } I'(x,y) \geq T_{inf}. \end{cases} \quad (34)$$

Здесь I – изображение до предварительной обработки, I' – изображения после предварительной обработки, T_{inf} – порог, определяющий, является ли пиксель информативным, или принадлежащим фону. Последним критерием в данной группе является оценка сохранения цветовой информации:

$$\vartheta_{col} = \frac{|C_h(I')|}{|C_h(I)|}, \quad (35)$$

где $C_h(I)$ – множество цветowych кластеров до предобработки, а $C_h(I')$ – множество цветowych кластеров после предобработки. Итоговый критерий качества изображения можно записать как:

$$\vartheta_{qual} = \alpha_q \cdot \eta_e + \beta_q \cdot \vartheta_{col} - \gamma_q \cdot \vartheta_{noise} - \delta_q \cdot \vartheta_{inf}, \quad (36)$$

где $\alpha_q, \beta_q, \gamma_q, \delta_q \in [0,1]$ – весовые коэффициенты значимости отдельных критериев, для которых соблюдается условие $\alpha_q + \beta_q + \gamma_q + \delta_q = 1$.

Вторая группа критериев направлена на оценку точности распознавания трех основных составляющих обрабатываемых элементов чертежей: графических объектов, текстовых строк и полей штампа:

$$Acc_{obj} = \frac{|O^* \cap O|}{|O^*|}, \quad Acc_{text} = \frac{|L^* \cap L|}{|L^*|}, \quad Acc_{stamm} = \frac{|S^* \cap S|}{|S^*|}, \quad (37)$$

где O^*, L^* и S^* – истинные множества объектов, текстовых строк и полей штампа соответственно. O, L и S – множества распознанных в процессе анализа объектов, строк и полей штампа. Итоговый критерий комплексной информативности в таком случае можно описать как:

$$\vartheta_{acc} = \alpha_a \cdot Acc_{obj} + \beta_a \cdot Acc_{text} + \gamma_a \cdot Acc_{stamm}, \quad (38)$$

где $\alpha_a, \beta_a, \gamma_a \in [0,1]$ – весовые коэффициенты значимости информативных компонент, для которых соблюдается условие $\alpha_a + \beta_a + \gamma_a = 1$.

Критерий влияния предобработки на классификацию нацелен на оценку влияния модели предобработки на точность классификации. В общем виде точность классификации на наборе данных определяется как:

$$Acc_{class} = \frac{N_{correct}}{N}, \quad (39)$$

где $N_{correct}$ – количество корректно классифицированных изображений чертежей, N – общее количество изображений чертежей в датасете. Пусть $Acc_{class}(I)$ – точность классификации датасета до предобработки, $Acc_{class}(I')$ – после предобработки. Тогда критерий влияния предобработки можно описать следующим образом:

$$\vartheta_{class} = Acc_{class}(I') - Acc_{class}(I). \quad (40)$$

Итоговый интегральный критерий эффективности предобработки можно записать как:

$$\Theta_p = \lambda_1 \cdot \vartheta_{acc} + \lambda_2 \cdot \vartheta_{class} + \lambda_3 \cdot \vartheta_{qual} - \lambda_4 \cdot \vartheta_{inf}, \quad (41)$$

где $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in [0,1]$ – весовые коэффициенты значимости используемых критериев, для которых также соблюдается условие $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$. Увеличение значения

данного критерия означает рост эффективности модели классификации в случае применения заданной модели.

Результаты

Для оценки эффективности разработанной адаптивной модели предобработки были проведены экспериментальные исследования, в рамках которых набор данных из около 500 изображений технических чертежей строительных металлоконструкций обрабатывался комплексной моделью классификации из нейросетевых компонентов на основе архитектур YOLO и CRNN.

Для снижения влияния случайного разбиения выборки оценка эффективности модели проводилась с использованием k -блочной кросс-валидации с $k = 5$. Исходный набор изображений был разделен на пять приблизительно равных непересекающихся подмножеств. На каждой итерации четыре подмножества использовались для настройки компонентов комплексной модели классификации, а оставшееся – для тестирования. Итоговые значения критериев определялись усреднением результатов всех пяти запусков. Полученные значения характеризовались небольшим разбросом (не более 2–3 %), что свидетельствует о репрезентативности используемой выборки и устойчивости предложенной модели предобработки.

В рамках исследования использовались три состояния вышеописанного набора: без предварительной обработки, после обработки тремя стандартными операциями (масштабирование, фильтрация шумов и бинаризация) и после обработки адаптивной моделью.

Оценка классического подхода к предобработке проводилась для всех возможных комбинаций алгоритмов, рассмотренных ранее. То есть для операции масштабирования применялись методы усреднения, наибольших, наименьших и медианных значений, билинейная, бикубическая, B-сплайн интерполяции и метод Ланцоша; для фильтрации импульсных шумов использовались гармонический, контргармонический, морфологический, α -усеченный и модификации медианного фильтра; для бинаризации – методы Оцу, Капура, Ниблака, Саувола и Вольфа.

В связи с наличием в наборе данных, как аналоговых изображений, так и изображений, полученных с помощью CAD-систем, допускалась значительная вариация значений критериев в зависимости от параметров чертежей. В случае классического подхода к предобработке ввиду использования разных комбинаций методов, разброс значений был очень большим. Отдельно были выделены результаты при использовании тех же алгоритмов, что и в адаптивном подходе. Все эксперименты выполнялись на ЭВМ со следующими характеристиками:

- Процессор: AMD Ryzen 7 7800X3D 8-Core Processor;
- Графическая карта: NVIDIA GeForce RTX 5070 Ti;
- Оперативная память: DDR5, 64 ГБ;
- Операционная система: Windows 11 Home.

Настройка значений весовых коэффициентов напрямую зависит от используемой модели классификации и влияния отдельных параметров изображения на итоговое значение класса изделия. Для используемой комплексной модели классификации были подобраны следующие значения: $\alpha_q = 0,175$, $\beta_q = 0,025$, $\gamma_q = 0,5$, $\delta_q = 0,3$, $\alpha_a = 0,4$, $\beta_a = 0,25$, $\gamma_a = 0,35$, $\lambda_1 = 0,25$, $\lambda_2 = 0,65$, $\lambda_3 = 0,09$, $\lambda_4 = 0,01$. При данных коэффициентах значение $\vartheta_{qual} < 0$ подразумевает негативное влияние предобработки на общее качество изображения, $\vartheta_{qual} > 0$ – позитивное влияние. Стоит отметить, что даже при отрицательном значении критерия ϑ_{qual} допускается прирост точности

распознавания объектов и операций классификации, поскольку не все элементы изображения используются в данных процессах.

Настройка выполнялась на обучающей части выборки посредством последовательного перебора параметров в заданных диапазонах. В качестве целевой функции использовалось максимальное значение интегрального критерия эффективности (41). Для каждого набора коэффициентов проводилась оценка на всех фрагментах k -блочной кросс-валидации, после чего выбирались параметры, обеспечивающие наибольшее среднее значение критерия при минимальной дисперсии результатов.

Таблица 1 – Оценка эффективности подходов к предобработке при классификации изделий
Table 1 – Evaluation of the effectiveness of pre-processing approaches in product classification

Критерий	Тип обработки			
	Без предобработки	Классическая предобработка (все методы)	Классическая предобработка (оптимальные методы)	Адаптивная предобработка
η_{noise}	0,35 – 0,45	0,09 – 0,32	0,09 – 0,15	0,09 – 0,17
η_e	1,00	0,44 – 0,80	0,65 – 0,80	0,88 – 0,96
ϑ_{inf}	0,00	0,18 – 0,58	0,18 – 0,35	0,04 – 0,12
ϑ_{col}	1,00	0,00 – 0,15	0,00 – 0,15	0,75 – 0,95
ϑ_{qual}	0,00	~ -0,103	~ -0,013	~ 0,093
Acc_{obj}	0,78 – 0,86	0,76 – 0,90	0,82 – 0,90	0,89 – 0,95
Acc_{text}	0,70 – 0,82	0,67 – 0,92	0,84 – 0,92	0,90 – 0,96
Acc_{stamm}	0,72 – 0,84	0,71 – 0,93	0,86 – 0,93	0,92 – 0,97
ϑ_{acc}	~ 0,791	~ 0,818	~ 0,877	~ 0,931
ϑ_{class}	0,76 – 0,85	0,74 – 0,92	0,86 – 0,92	0,91 – 0,96
Θ_p	~ 0,7210	~ 0,7737	~ 0,7966	~ 0,8490

Обсуждение

Результаты исследования, представленные в Таблице 1, фиксируют больший прирост точности классификации при использовании предложенной в данной статье адаптивной модели предобработки по сравнению со стандартной предобработкой или ее полным отсутствием. Модель также показывает наивысшие значения критериев качества обработанного изображения и прироста точности распознавания всех трех рассматриваемых групп информативных объектов на чертежах.

В отличие от современных нейросетевых методов восстановления документов, основанных на архитектурах U-Net, GAN и Diffusion-моделях, предложенный подход не требует формирования больших размеченных выборок для обучения этапа предобработки. Все принимаемые решения определяются интерпретируемыми критериями качества изображения и математически формализованными правилами выбора операций обработки. Это позволяет адаптировать модель к новым типам технической документации без повторного обучения нейросетевых компонентов и обеспечивает воспроизводимость результатов обработки.

Важно отметить, что предложенная адаптивная модель не является универсальным решением для предобработки изображений чертежей. Используемые в модели алгоритмы были подобраны с учетом особенностей структуры чертежей строительных металлоконструкций. При обработке изображений других классов допускается наличие новых типов шумов, большего или меньшего количества

используемых цветов, а также совершенно других групп информативных объектов, влияющих на классификацию изображений.

В таких случаях применение как других комбинаций алгоритмов, так и совершенно иных операций предварительной обработки может давать лучший результат по сравнению с адаптивной моделью. Однако в рамках данного исследования акцент делался на чертежи строительных металлоконструкций.

Заключение

В работе предложена адаптивная модель предварительной обработки изображений чертежей строительных металлоконструкций, направленная на повышение точности их классификации. Данная модель представляет собой последовательность операций масштабирования, фильтрации шумов, бинаризации и восстановления цвета информативных пикселей, где необходимость применения каждой из которых определяется соответствующими критериями.

В рамках работы был также составлен набор критериев оценки качества предварительной обработки при взаимодействии с комплексными моделями классификации чертежей строительных металлоконструкций. Было проведено исследование, в рамках которого выявлен прирост точности классификации при применении предложенной модели предобработки по сравнению с классической предобработкой или ее отсутствием.

Результаты работы дают почву для проведения экспериментальных исследований на больших наборах данных с вариацией значений предложенных критериев. Также в дальнейшем планируется внедрение предложенной модели в комплексные системы классификации изображений чертежей строительных металлоконструкций.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Lu T., Yang H., Yang R., et al. Automatic analysis and integration of architectural drawings. *International Journal of Document Analysis and Recognition*. 2007;9(1):31–47. <https://doi.org/10.1007/s10032-006-0029-6>
2. Zhang L.Y., Zhu J.Y., Zhu Q.Y. A New System for Automatic Understanding Engineering Drawings. *CIRP Annals*. 1992;41(1):477–479. [https://doi.org/10.1016/S0007-8506\(07\)61248-7](https://doi.org/10.1016/S0007-8506(07)61248-7)
3. Fan N. Application of CAD Combined with Computer Image Processing Technology in Mechanical Drawing. In: *Application of Intelligent Systems in Multi-modal Information Analytics: Proceedings of the 2020 International Conference on Multi-model Information Analytics (MMIA2020): Volume 2, 05–06 March 2020, Changzhou, China*. Cham: Springer; 2021. P. 319–324. https://doi.org/10.1007/978-3-030-51556-0_46
4. Bai Y.B., Xu X.W. Object Boundary Encoding – A New Vectorisation Algorithm for Engineering Drawings. *Computers in Industry*. 2001;46(1):65–74. [https://doi.org/10.1016/S0166-3615\(01\)00115-4](https://doi.org/10.1016/S0166-3615(01)00115-4)
5. Lin H.-Y., Lin Ch.-Y., Lin Ch.-J., et al. A Study of Digital Image Enlargement and Enhancement. *Mathematical Problems in Engineering*. 2014;2014:825169. <https://doi.org/10.1155/2014/825169>
6. Keys R. Cubic Convolution Interpolation for Digital Image Processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1981;29(6):1153–1160. <https://doi.org/10.1109/TASSP.1981.1163711>
7. Zhu Y., Dai Y., Han K., et al. An Efficient Bicubic Interpolation Implementation for Real-Time Image Processing Using Hybrid Computing. *Journal of Real-Time Image Processing*. 2022;19(6):1211–1223. <https://doi.org/10.1007/s11554-022-01254-8>

8. Chunshi L., Yu L., Yimin Zh. Study of B-Spline Interpolation, Correction and Inverse Algorithm. In: *Advances in Automation and Robotics: Selected papers from the 2011 International Conference on Automation and Robotics (ICAR 2011): Volume 2, 01–02 December 2011, Dubai, UAE*. Berlin, Heidelberg: Springer; 2012. P. 215–221. https://doi.org/10.1007/978-3-642-25646-2_30
9. Simon H.D. The Lanczos Algorithm with Partial Reorthogonalization. *Mathematics of Computation*. 1984;42(165):115–142. <https://doi.org/10.2307/2007563>
10. Boyat A.K., Joshi B.K. A Review Paper: Noise Models in Digital Image Processing. *Signal & Image Processing: An International Journal*. 2015;6(2):63–75. <https://doi.org/10.5121/sipij.2015.6206>
11. Fatemi M., Amindavar H., Ritcey J.A. Noise reduction via harmonic estimation in Gaussian and non-Gaussian environments. *Signal Processing*. 2010;90(5):1554–1561. <https://doi.org/10.1016/j.sigpro.2009.11.002>
12. Haralick R.M., Sternberg S.R., Zhuang X. Image Analysis Using Mathematical Morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1987;PAMI-9(4):532–550. <https://doi.org/10.1109/TPAMI.1987.4767941>
13. Sun M. Comparison of Processing Results of Median Filter and Mean Filter on Gaussian Noise. *Applied and Computational Engineering*. 2023;5(1):779–785. <https://doi.org/10.54254/2755-2721/5/20230702>
14. Vo G.D., Park Ch. Robust Regression for Image Binarization under Heavy Noise and Nonuniform Background. *Pattern Recognition*. 2018;81(4):224–239. <https://doi.org/10.1016/j.patcog.2018.04.005>
15. Otsu N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*. 1979;9(1):62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
16. Kapur J.N., Sahoo P.K., Wong A.K.C. A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram. *Computer Vision, Graphics, and Image Processing*. 1985;29(3):273–285. [https://doi.org/10.1016/0734-189X\(85\)90125-2](https://doi.org/10.1016/0734-189X(85)90125-2)
17. Sauvola J.J., Pietikäinen M. Adaptive Document Image Binarization. *Pattern Recognition*. 2000;33(2):225–236. [https://doi.org/10.1016/S0031-3203\(99\)00055-2](https://doi.org/10.1016/S0031-3203(99)00055-2)
18. Wolf C., Jolion J.-M. Extraction and recognition of artificial text in multimedia documents. *Formal Pattern Analysis & Applications*. 2004;6(4):309–326. <https://doi.org/10.1007/s10044-003-0197-7>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Колесников Владимир Дмитриевич, аспирант, Белгородский государственный технологический университет им. В.Г. Шухова, Белгород, Российская Федерация.

email: burwestnikfagrex@gmail.com

ORCID: [0009-0004-5767-0424](https://orcid.org/0009-0004-5767-0424)

Vladimir D. Kolesnikov, Postgraduate, Belgorod State Technological University named after V.G. Shukhov, Belgorod, the Russian Federation.

Кабалянц Петр Степанович, кандидат технических наук, доцент, Белгородский государственный технологический университет им. В.Г. Шухова, Белгород, Российская Федерация.

email: p.s.k@list.ru

ORCID: [0000-0003-0598-0880](https://orcid.org/0000-0003-0598-0880)

Petr S. Kabalyants, Candidate of Engineering Sciences, Associate Professor of Belgorod State Technological University named after V.G. Shukhov, Belgorod, the Russian Federation.

*Статья поступила в редакцию 11.06.2026; одобрена после рецензирования 25.08.2026;
принята к публикации 11.09.2026.*

*The article was submitted 11.06.2026; approved after reviewing 25.08.2026;
accepted for publication 11.09.2026.*