

УДК 004.932.4

DOI: [10.26102/2310-6018/2026.61.10.016](https://doi.org/10.26102/2310-6018/2026.61.10.016)

Алгоритм пространственной линейной фильтрации изображений на основе сепарабельного биномиального ядра для потоковой реализации в программируемой логике

К.Р. Гайсин¹✉, А.А. Егорчев¹, М.С. Азов²

¹Казанский (Приволжский) федеральный университет, Казань, Российская Федерация

²Акционерное общество «Ульяновское конструкторское бюро приборостроения»,
Ульяновск, Российская Федерация

Резюме. В работе исследовано влияние размера ядра гауссова фильтра на характеристики обнаружения объектов облегченным сверточным нейросетевым детектором в видеосцене с пониженной освещенностью. Рассмотрена сепарабельная модель фильтра, ориентированная на потоковую реализацию в программируемой логике, и выполнено сравнение вариантов ее аппаратного исполнения по числу операций на пиксель. Целью работы является экспериментальная оценка влияния размера ядра на латентность и устойчивость обнаружения, а также проверка сохранения характеристик детекции при переходе к реализации без аппаратных умножителей. Обработка выполнялась покадрово для видеопоследовательности длиной 1351 кадр при двух значениях порога уверенности детектора. Установлено, что воздействие гауссовой фильтрации проявляется преимущественно в повышении устойчивости детекции, а не в улучшении раннего обнаружения: число кадров с обнаружением возрастает на 6,7 % при базовом пороге уверенности и на 25 % при строгом, тогда как момент первого обнаружения практически не смещается. При строгом пороге и размере ядра, превышающем три отсчета, обнаружена деградация раннего обнаружения, вызванная избыточным сглаживанием объекта малого углового размера; это задает верхнюю границу допустимой области параметра. Показано, что близость биномиальной аппроксимации к точному гауссову фильтру сохраняется и на уровне откликов нелинейного детектора: расхождение по моменту обнаружения не превышает одного кадра, по числу кадров с обнаружением – 0,7 %. Задача выбора параметра формализована как максимизация устойчивости детекции при ограничении на латентность обнаружения. Оптимальным оказалось ядро из трех отсчетов с коэффициентами один, два, один и нормировкой сдвигом на два разряда, требующее четырех сумматоров и двух буферов строк при нулевом расходе блоков цифровой обработки сигналов.

Ключевые слова: гауссова фильтрация, сепарабельное биномиальное ядро, шумоподавление, обнаружение объектов, латентность обнаружения, устойчивость детекции, программируемая логика, потоковая обработка, реализация без умножителей, техническое зрение.

Для цитирования: Гайсин К.Р., Егорчев А.А., Азов М.С. Алгоритм пространственной линейной фильтрации изображений на основе сепарабельного биномиального ядра для потоковой реализации в программируемой логике. *Моделирование, оптимизация и информационные технологии*. 2026;14(10). URL: <https://moitvvt.ru/ru/journal/article?id=2600> DOI: 10.26102/2310-6018/2026.61.10.016

A spatial linear image filtering algorithm based on a separable binomial kernel for streaming implementation in programmable logic

K.R. Gaisin¹✉, A.A. Egorchev¹, M.S. Azov²

¹Kazan (Volga Region) Federal University, Kazan, the Russian Federation

²Joint-Stock Company Ulyanovsk Instrument Manufacturing Design Bureau, Ulyanovsk,
the Russian Federation

Abstract. The effect of the Gaussian filter kernel size on the object detection performance of a lightweight convolutional neural network detector in a low-light video scene is studied. A separable filter model oriented towards streaming implementation in programmable logic is considered, and the hardware implementation options are compared in terms of the number of operations per pixel. The aim of the work is an experimental assessment of the influence of the kernel size on detection latency and detection stability, together with a verification that the detection performance is preserved when moving to an implementation without hardware multipliers. Processing was carried out frame by frame for a video sequence of 1351 frames at two values of the detector confidence threshold. It is established that Gaussian filtering acts mainly by increasing detection stability rather than by improving early detection: the number of frames with detection grows by 6.7 % at the baseline confidence threshold and by 25 % at the strict one, whereas the moment of first detection is practically unchanged. At the strict threshold and a kernel size exceeding three samples, a degradation of early detection is observed, caused by excessive smoothing of an object of small angular size; this sets the upper bound of the admissible parameter region. It is shown that the closeness of the binomial approximation to the exact Gaussian filter is preserved at the level of the responses of a nonlinear detector as well: the discrepancy in the detection moment does not exceed one frame, and in the number of frames with detection it is 0.7 %. The parameter selection problem is formalised as maximisation of detection stability under a constraint on detection latency. The optimum is a three-sample kernel with coefficients one, two, one and normalisation by a two-bit shift, requiring four adders and two line buffers with zero consumption of digital signal processing blocks.

Keywords: Gaussian filtering, separable binomial kernel, noise suppression, object detection, detection latency, detection stability, programmable logic, streaming processing, multiplierless implementation, machine vision.

For citation: Gaisin K.R., Egorchev A.A., Azov M.S. A spatial linear image filtering algorithm based on a separable binomial kernel for streaming implementation in programmable logic. *Modeling, Optimization and Information Technology*. 2026;14(10). (In Russ.). URL: <https://moitvvt.ru/journal/article?id=2600> DOI: 10.26102/2310-6018/2026.61.10.016

Введение

Задача обнаружения объектов в системах технического зрения предъявляет высокие требования к качеству предварительной обработки видеопотока, особенно в условиях пониженной освещенности. Высокочастотный и дробовой шум матрицы, характерный для ночной съемки, снижает различимость объектов малого углового размера и приводит к нестабильности срабатывания детектора на ранних кадрах наблюдения [1]. Для обнаружения в режиме реального времени во встраиваемых системах применяются легкие нейросетевые архитектуры, в частности, YOLOv3-tiny [2, 3], точность которых напрямую зависит от качества входного изображения.

Гауссова фильтрация – классический метод пространственного линейного сглаживания и подавления высокочастотного шума [4]. В отличие от прямоугольного усредняющего фильтра гауссово ядро имеет монотонную частотную характеристику: оно не вносит переколебаний (звона) и ложных контуров, способных создать у детектора паразитные градиенты. Это оконный оператор с локальной областью зависимости – выходной пиксель определяется только своей окрестностью. Отсюда возможность полной конвейеризации и пространственного параллелизма, что позволяет встроить фильтр в потоковый тракт с детерминированной латентностью [5].

Для аппаратной реализации существенны еще два свойства гауссова ядра. Первое – сепарабельность: двумерное гауссово ядро представимо в виде произведения двух одномерных ядер, и свертка ядром $k \times k$ сводится к двум последовательным одномерным сверткам. Разложение точное и не вносит погрешности, а число операций на пиксель снижается с k^2 до $2k$ [6]. Второе – близость дискретизированной функции Гаусса к биномиальным коэффициентам (строкам треугольника Паскаля). Такие

коэффициенты целочисленные, а их сумма равна степени двойки, поэтому умножения и нормировка выполняются сдвигами, и фильтр перестает требовать блоки цифровой обработки сигналов (DSP) [7].

Влияние предварительной обработки на качество нейросетевого обнаружения исследовалось неоднократно. Показано, что подбор параметров сглаживающих операторов – медианной фильтрации, гауссова сглаживания, пороговой обработки – повышает среднюю точность обнаружения на несколько процентов [8]. Отдельное направление составляют исследования устойчивости детекторов семейств YOLO и R-CNN к деградациям входного изображения: шуму, размытию, снижению контраста [9]. Аппаратная сторона задачи проработана не менее подробно: структуры фильтров на биномиальных коэффициентах, не требующие умножителей и реализуемые сдвигами и сложениями, описаны применительно к программируемой логике [7], предложены схемы на основе многослойной свертки базового биномиального ядра [10], а сама биномиальная фильтрация входит в промышленные библиотеки машинного зрения. Реализации детекторов семейства YOLO на программируемых логических интегральных схемах рассмотрены в [3, 11].

Вместе с тем два вопроса остаются вне рассмотрения. Первый касается критерия оценки приближенной реализации. Качество биномиальной аппроксимации оценивалось по мере численного отклонения от точного гауссова фильтра, тогда как для системы обнаружения существенно иное – сохраняются ли характеристики детекции. Детектор нелинеен, и малое отклонение входного изображения не гарантирует малого изменения отклика. Второй вопрос касается самих метрик. Средняя точность, полнота и точность обнаружения описывают качество на наборах статических изображений, но не характеризуют латентность: номер кадра, начиная с которого объект обнаруживается в видеопоследовательности. Для систем видеонаблюдения реального времени именно этот показатель определяет время реакции, и зависимость его от параметров предобработки не изучена.

Целью работы является экспериментальная оценка влияния размера ядра гауссова фильтра на латентность и устойчивость обнаружения объектов в видеопоследовательности, а также проверка сохранения характеристик детекции при переходе к реализации без аппаратных умножителей. Для этого в работе решаются три задачи: экспериментально оценивается влияние размера ядра на латентность и устойчивость обнаружения; точная и биномиальная реализации сопоставляются по метрикам детекции; задача выбора размера ядра формализуется в виде, пригодном для адаптивного контура.

Материалы и методы

Параллелизуемая модель алгоритма гауссовой фильтрации. Рассматривается оператор оконной линейной фильтрации, применяемый к каждому пикселю входного кадра. Для пикселя с координатами (x, y) гауссов фильтр с ядром размером $k \times k$ вычисляет выходное значение яркости как взвешенную сумму яркостей пикселей локальной окрестности:

$$Y_{out}(x, y) = \sum_{i=-r}^r \sum_{j=-r}^r G(i, j) \cdot Y_{in}(x + i, y + j), \quad r = \frac{k-1}{2}. \quad (1)$$

Весовые коэффициенты задаются дискретизированной функцией Гаусса:

$$G(i, j) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{i^2+j^2}{2\sigma^2}\right), \quad \sum_{i=-r}^r \sum_{j=-r}^r G(i, j) = 1. \quad (2)$$

где $Y_{in}(x, y)$ – яркость входного пикселя, σ – параметр ширины ядра. Конечный размер окна усекает функцию Гаусса, поэтому нормировка суммы весов к единице не является формальностью: она компенсирует отброшенные хвосты и сохраняет среднюю яркость изображения.

Параметры σ и k связаны между собой. Их отношение определяет степень усечения: полуширина окна r должна охватывать основную часть массы функции Гаусса, иначе нормировка вносит заметное искажение формы ядра, но при чрезмерно большом окне периферийные отсчеты получают пренебрежимо малые веса и увеличивают вычислительные затраты без эффекта. В работе принято соотношение, используемое в библиотеке компьютерного зрения OpenCV:

$$\sigma(k) = 0,3 \cdot \left(\frac{k-1}{2} - 1 \right) + 0,8, \quad (3)$$

которое дает значения σ от 0,80 при $k = 3$ до 1,70 при $k = 9$ (Таблица 2). Отношение r/σ при этом остается в диапазоне от 1,25 до 2,35, а доля массы функции Гаусса, попадающая в окно, возрастает с 79 % при $k = 3$ до 98 % при $k = 9$. Размер ядра k выступает, таким образом, единственным управляющим параметром модели.

При $k = 1$ ядро вырождается в единичное, и оператор реализует тождественное преобразование $Y_{out}(x, y) = Y_{in}(x, y)$. Это значение принимается базовым.

На границах кадра окно выходит за пределы изображения, что требует правила доопределения отсчетов. В программной модели используется зеркальное отражение относительно границы, в аппаратной реализации – дублирование краевого пикселя, не требующее дополнительной адресной логики. Различие затрагивает полосу шириной r пикселей вдоль границ кадра и на результаты детекции в центральной области не влияет.

Ключевое свойство оператора – локальность области зависимости. Выходное значение каждого пикселя определяется только его $k \times k$ окрестностью и не зависит от значений других выходных пикселей:

$$Y_{out}(x_1, y_1) \perp Y_{out}(x_2, y_2), \quad \forall (x_1, y_1) \neq (x_2, y_2), \quad (4)$$

где символ $\perp$ обозначает вычислительную независимость: значения могут быть получены в произвольном порядке или одновременно. Отсюда следует возможность полной конвейеризации и пространственного параллелизма. При прямой реализации оператор строится на скользящем окне с $k - 1$ буферами строк, формирующими выборку из k^2 отсчетов для каждой пространственной позиции [5, 12].

Характер воздействия фильтра задается размером ядра. Малые k подавляют высокочастотную составляющую шума, почти не затрагивая структуру изображения. С ростом k область сглаживания расширяется – мелкие детали объекта размываются все сильнее, вплоть до слияния с фоном. Для объектов малого углового размера это означает конкуренцию двух эффектов: подавление шума повышает различимость объекта, а размытие снижает его текстурную характерность. Соотношение этих эффектов и определяет выбор параметра, что требует экспериментальной оценки.

Отдельного рассмотрения заслуживает сепарабельность гауссова ядра. Двумерное ядро представимо в виде произведения двух одномерных:

$$G(i, j) = g(i) \cdot g(j), \quad (5)$$

где g – одномерное гауссово ядро длины k . Двумерная свертка ядром $k \times k$ тождественно эквивалентна двум последовательным одномерным сверткам: горизонтальной по строке и вертикальной по столбцу. Разложение точное, вычислительной погрешности не вносит, а число операций умножения-накопления на пиксель снижается с k^2 до $2k$.

Изменение размера ядра не затрагивает структуру вычислительной логики: переключаются только глубина буферов строк и набор коэффициентов. Замкнутый адаптивный контур в такой схеме реализуется без реконфигурации программируемой логики.

Реализация в программируемой логике и сепарабельное биномиальное ядро. Реализация гауссовой фильтрации в потоковом тракте на базе программируемой логики допускает несколько схем, существенно различающихся по затратам ресурсов.

Вариант 1. Двумерная свертка на блоках цифровой обработки сигналов. Прямое вычисление по формуле (1) требует k^2 операций умножения-накопления на пиксель и $k - 1$ буферов строк. Каждое умножение реализуется отдельным DSP-блоком либо мультиплексируется во времени с соответствующим снижением пропускной способности. Затраты растут как $O(k^2)$: для ядра 9×9 это 81 умножение на каждый пиксель. При потоковой обработке видео высокого разрешения такая схема упирается в число доступных DSP-блоков [13].

Вариант 2. Двумерная свертка без DSP-блоков. Те же k^2 умножений выполняются сдвигами и сложениями на логических ресурсах. Нагрузка с DSP-блоков снимается, но переносится на логику: число сумматоров по-прежнему составляет $O(k^2)$. Для ядер большого размера бюджет логических таблиц истинности растет быстрее, чем в первом варианте растет число умножителей, поэтому схема оправдана лишь при малых k .

Вариант 3. Сепарабельная реализация. В силу разложения (5) двумерная свертка выполняется двумя одномерными проходами: горизонтальным по строке и вертикальным по столбцу с использованием буфера на $k - 1$ строк. Число операций умножения-накопления составляет $2k$ вместо k^2 , затраты растут линейно. Сопоставление вариантов приведено в Таблице 1.

Таблица 1 – Число операций умножения-накопления на пиксель для прямой и сепарабельной реализаций

Table 1 – Number of multiply-accumulate operations per pixel for the direct and separable implementations

k	Двумерная свертка (k^2)	Сепарабельная ($2k$)	Сокращение, %
3	9	6	33
5	25	10	60
7	49	14	71
9	81	18	78

Дальнейшее снижение затрат достигается выбором самих коэффициентов. Дискретизированная функция Гаусса с высокой точностью аппроксимируется биномиальными коэффициентами – строками треугольника Паскаля; это следствие центральной предельной теоремы, поскольку многократная свертка прямоугольного окна сходится к гауссову виду. Одномерное биномиальное ядро длины k имеет вид:

$$g_k = \frac{1}{2^{k-1}} \left[\binom{k-1}{0}, \binom{k-1}{1}, \dots, \binom{k-1}{k-1} \right], \quad (6)$$

что дает $g_3 = [1, 2, 1]/4$, $g_5 = [1, 4, 6, 4, 1]/16$, $g_7 = [1, 6, 15, 20, 15, 6, 1]/64$, $g_9 = [1, 8, 28, 56, 70, 56, 28, 8, 1]/256$.

Практическую ценность придают два обстоятельства. Во-первых, сумма коэффициентов каждого ядра равна 2^{k-1} , поэтому нормировка выполняется сдвигом вправо на $k - 1$ разрядов – делитель не нужен. Во-вторых, сами коэффициенты являются малыми целыми числами, и умножение на них раскладывается в сдвиги и сложения. Для

ядра $g_3 = [1, 2, 1]$ умножение на 2 сводится к сдвигу на один разряд, а умножение на 1 не требует операций вовсе: фильтр реализуется без единого аппаратного умножителя. Структуры фильтров на биномиальных коэффициентах и варианты их организации в программируемой логике систематически исследованы в [7].

Известным структурным свойством биномиального ядра является его представимость в виде каскада одинаковых двухотводных звеньев [7, 10]: поскольку $[1, 2, 1] = [1, 1] * [1, 1]$, ядро g_k в общем случае получается $(k - 1)$ -кратной сверткой ядра $[1, 1]$. Такая структура удобна для конвейеризации: тракт собирается из идентичных ступеней, а переход к большему размеру ядра означает добавление ступени, а не перепроектирование схемы.

Ядро (6) представляет собой биномиальное распределение с параметром $1/2$, дисперсия которого равна $(k - 1)/4$. Эффективная ширина биномиального ядра составляет

$$\sigma_{\text{бин}} = \frac{\sqrt{k-1}}{2} \quad (7)$$

и связана с размером окна жестко, без свободного параметра – в отличие от точного гауссова ядра, где σ задается соотношением (3) независимо. Численное сопоставление приведено в Таблице 2.

Сепарабельность вычислительной погрешности не вносит: разложение (5) точное. Единственным приближением остается замена точных гауссовых весов биномиальными. Численные характеристики этого приближения известны [7]; вопрос, исследуемый в настоящей работе, состоит в том, сохраняется ли близость на уровне характеристик детекции, и проверяется экспериментально в разделе «Результаты». Управляющим параметром тракта остается размер ядра k , переключаемый через регистры управления и глубину буферов строк без реконфигурации программируемой логики.

Методика экспериментального исследования. Эксперимент проводился с целью оценить влияние размера ядра гауссова фильтра k на характеристики детекции объекта при неизменной структуре нейросетевого детектора. Программная реализация на данном этапе позволяет верифицировать модель, сохраняя логику обработки, соответствующую целевой аппаратной архитектуре. Схема экспериментального контура приведена на Рисунке 1.

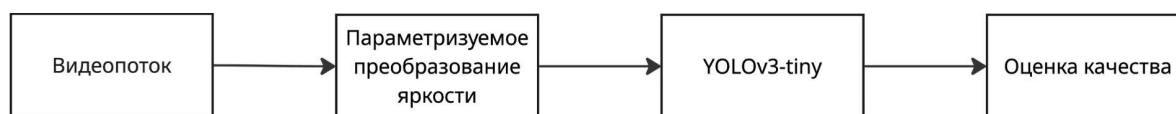


Рисунок 1 – Схема экспериментального программного контура: видеопоток – параметризуемое преобразование яркости (гауссова фильтрация) – нейросетевой детектор – оценка качества детекции

Figure 1 – Diagram of the experimental software loop: video stream – parameterised brightness transform (Gaussian filtering) – neural network detector – detection quality assessment

Тестовым материалом служила видеопоследовательность длительностью 1351 кадр с разрешением 2560×1440 пикселей. Сцена содержит одного человека, движущегося по пешеходной улице в условиях пониженной освещенности. В начальной фазе наблюдения объект удален от камеры и имеет малый угловой размер, затем постепенно приближается к локальному источнику освещения. Дополнительную сложность создают частичные заслонения объектами переднего плана и шумы, характерные для ночной съемки (Рисунок 2). Применение той же видеопоследовательности, что и в предшествующих исследованиях операторов

предобработки, обеспечивает сопоставимость результатов в рамках единой адаптивной платформы.



Рисунок 2 – Исходная сцена в начальной фазе наблюдения объекта и в момент его приближения к освещенной области

Figure 2 – Source scene in the initial phase of object observation and at the moment of its approach to the illuminated area

Детекция выполнялась моделью YOLOv3-tiny без переобучения и изменения архитектуры, что позволяет рассматривать детектор как фиксированный вычислительный модуль и относить наблюдаемые различия исключительно к предварительной обработке [14]. Класс интереса – person. Порог подавления немаксимумов принят равным 0,45. Фильтрация выполняется над кадром полного разрешения; приведение к входному размеру детектора 416×416 пикселей производится после нее, что соответствует архитектуре тракта, в которой геометрическая нормализация следует за фильтрацией.

Размер ядра варьировался по дискретному набору нечетных значений $k \in \{1, 3, 5, 7, 9\}$; значение $k = 1$ соответствует отсутствию фильтрации и принято базовым.

Для каждого размера ядра исследовались две реализации фильтра. Первая – точный гауссов фильтр с плавающей точкой, где σ вычисляется из размера ядра по соотношению (3); эта реализация принята эталонной. Вторая – сепарабельный биномиальный фильтр со степенью-двойки нормировкой (целочисленные одномерные ядра Паскаля), моделирующий реализацию без аппаратных умножителей на программируемой логике.

Ширина двух ядер совпадает не полностью. Согласно (7), эффективная ширина биномиального ядра определяется только размером окна, тогда как для точного ядра σ задается соотношением (3). Численное сопоставление приведено в Таблице 2.

Таблица 2 – Эффективная ширина точного гауссова и биномиального ядер при одинаковом размере окна

Table 2 – Effective width of the exact Gaussian and the binomial kernels for the same window size

k	σ точного ядра	σ биномиального ядра	Отношение
3	0,80	0,71	0,88
5	1,10	1,00	0,91
7	1,40	1,22	0,87
9	1,70	1,41	0,83

Биномиальное ядро систематически на 9–17 % уже точного при том же размере окна. Различие учитывается при интерпретации результатов сопоставления реализаций в разделе «Результаты».

Обработка выполнялась покадрово с полным прогоном видеопоследовательности для каждой комбинации размера ядра, метода реализации и порога уверенности. Фильтр применялся ко всем трем цветовым каналам независимо, что соответствует поканальному характеру операции; в целевой аппаратной архитектуре обработке подлежит один канал яркости, что снижает объем вычислений втрое, не меняя характера воздействия.

Качество детекции описывалось следующими показателями. Номер кадра первого обнаружения:

$$F_{\text{det}} = \min\{n \mid \text{det}(n) = 1\}, \quad (8)$$

количество кадров с детекцией:

$$N_{\text{det}} = \sum_{n=1}^N 1 [\text{det}(n) = 1], \quad (9)$$

средняя уверенность детекции по кадрам с обнаружением:

$$C_{\text{avg}} = \frac{1}{N_{\text{det}}} \sum_{n: \text{det}(n)} c_n, \quad (10)$$

средняя площадь ограничивающей рамки:

$$S_{\text{avg}} = \frac{1}{N_{\text{det}}} \sum_{n: \text{det}(n)} s_n, \quad (11)$$

где $\text{det}(n)$ – событие обнаружения объекта на кадре n ; c_n – уверенность; s_n – площадь рамки; N – общее число кадров видеопоследовательности. Показатель F_{det} характеризует латентность обнаружения, N_{det} – устойчивость детекции.

Анализ проводился при двух значениях порога уверенности: $T_c = 0,25$ – базовый порог, обеспечивающий баланс между чувствительностью и устойчивостью к ложным срабатываниям, и $T_c = 0,7$ – строгий порог для проверки воспроизводимости наблюдаемых зависимостей.

Результаты

Программная реализация экспериментального контура выполнена на языке Python с использованием библиотеки компьютерного зрения OpenCV. Точный гауссов фильтр реализован функцией двумерного гауссова сглаживания, сепарабельный биномиальный – функцией раздельной двумерной фильтрации с одномерными ядрами Паскаля; последняя схема соответствует потоковому методу реализации на программируемой логике с буферами строк. Для каждой комбинации размера ядра,

метода реализации и порога уверенности выполнялся полный прогон видеопоследовательности длиной 1351 кадр. Результаты для точного гауссова фильтра сведены в Таблицу 3.

Таблица 3 – Сводные метрики детекции при различных размерах ядра (точный гауссов фильтр)
Table 3 – Summary detection metrics for different kernel sizes (exact Gaussian filter)

k	$T_c = 0,25$				$T_c = 0,7$			
	F_{det}	N_{det}	C_{avg}	S_{avg}	F_{det}	N_{det}	C_{avg}	S_{avg}
1	197	715	0,428	10221	220	68	0,765	13342
3	197	742	0,434	10449	220	82	0,765	12955
5	197	750	0,436	10523	292	83	0,770	12660
7	197	759	0,443	10760	292	84	0,774	12680
9	198	763	0,449	10949	292	85	0,778	12891

Полужирным в Таблице 3 выделена строка оптимального значения $k^* = 3$, обоснование которого приведено в разделе «Обсуждение».

Результаты при базовом пороге уверенности. При базовом пороге уверенности $T_c = 0,25$ момент первого обнаружения практически не смещается: $F_{det} = 197$ для $k \in \{1, 3, 5, 7\}$ и $F_{det} = 198$ при $k = 9$. Устойчивость детекции, напротив, растет монотонно. Число кадров с обнаружением увеличивается с 715 до 763, что составляет 6,7 %, средняя уверенность – с 0,428 до 0,449. Средняя площадь ограничивающей рамки возрастает с 10221 до 10949 пкс²: сглаживание мелких деталей приводит к более цельному выделению силуэта.

Таким образом, при базовом пороге гауссова фильтрация не улучшает раннее обнаружение, но заметно стабилизирует детекцию на протяжении наблюдения.

Результаты при строгом пороге уверенности. Тенденция к росту устойчивости сохраняется и при строгом пороге $T_c = 0,7$, причем выражена сильнее: число кадров с обнаружением возрастает с 68 до 85, то есть на 25 %, средняя уверенность – с 0,765 до 0,778.

Показатель раннего обнаружения ведет себя иначе. При $k \in \{1, 3\}$ он составляет 220 кадров, а начиная с $k = 5$ скачкообразно возрастает до 292 кадров и далее не меняется. Причина в том, что увеличение ядра снижает текстурную характерность объекта малого углового размера на ранних кадрах. Уверенность детектора перестает достигать порога 0,7 до того момента, когда объект приблизится к камере настолько, что его различимость обеспечивается уже угловым размером, а не текстурой. Скачок на 72 кадра при переходе от $k = 3$ к $k = 5$ задает верхнюю границу допустимой области параметра.

Обратим внимание на поведение средней площади рамки: при строгом пороге она не растет монотонно, как при базовом, а имеет минимум при $k = 5$ (12660 пкс² против 13342 пкс² у базового варианта). Это согласуется со сдвигом момента первого обнаружения – при $k \geq 5$ в выборку попадают только кадры с уже приблизившимся объектом. Сводные зависимости метрик от размера ядра для обоих порогов приведены на Рисунке 3.

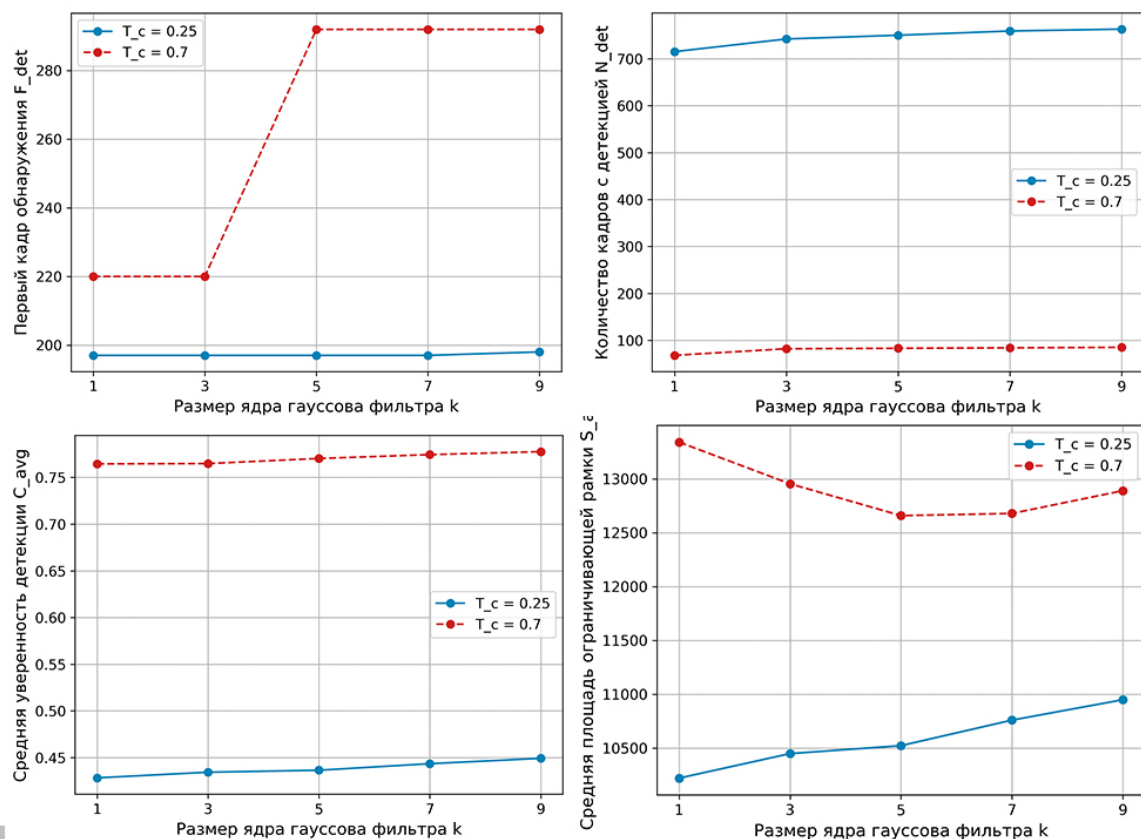


Рисунок 3 – Зависимости метрик детекции от размера ядра при базовом и строгом порогах уверенности

Figure 3 – Dependences of the detection metrics on the kernel size at the baseline and the strict confidence thresholds

Сопоставление точной и биномиальной реализации. Для обоснования реализации без аппаратных умножителей точный гауссов фильтр сопоставлялся с сепарабельным биномиальным. Результаты приведены в Таблице 4.

Таблица 4 – Сравнение точного гауссова и сепарабельного биномиального фильтров
Table 4 – Comparison of the exact Gaussian and the separable binomial filters

k	$T_c = 0,25$						$T_c = 0,7$					
	F_{det}		N_{det}		C_{avg}		F_{det}		N_{det}		C_{avg}	
	точн.	бин.	точн.	бин.	точн.	бин.	точн.	бин.	точн.	бин.	точн.	бин.
1	197	197	715	715	0,428	0,428	220	220	68	68	0,765	0,765
3	197	197	742	742	0,434	0,434	220	220	82	83	0,765	0,764
5	197	197	750	749	0,436	0,437	292	292	83	83	0,770	0,770
7	197	197	759	760	0,443	0,441	292	292	84	83	0,774	0,775
9	198	197	763	758	0,449	0,445	292	292	85	83	0,778	0,776

Расхождения оказались в пределах, не влияющих на выводы. Момент первого обнаружения при строгом пороге совпадает во всех без исключения точках; при базовом пороге максимальное расхождение составляет один кадр. По числу кадров с обнаружением наибольшее отличие зафиксировано при $k = 9$: 5 кадров из 763, то есть 0,7 %. При строгом пороге расхождение по числу кадров с обнаружением не превышает двух кадров. Средняя уверенность различается в третьем-четвертом знаке, максимально на 0,004.

Направление расхождений закономерно и объясняется различием эффективной ширины ядер (Таблица 2). Биномиальное ядро на 9–17 % уже точного, то есть сглаживает слабее. Отсюда следуют два наблюдения, согласующиеся между собой: при $k = 9$ и базовом пороге биномиальная реализация дает меньшую устойчивость (758 против 763 кадров), поскольку слабее подавляет шум, и одновременно обнаруживает объект на кадр раньше (197 против 198), поскольку меньше размывает его текстуру. Оба отклонения соответствуют более узкому ядру и не свидетельствуют о погрешности квантизации коэффициентов.

Обсуждение

Характер воздействия фильтра на детекцию. Близость биномиальной аппроксимации к точному гауссову фильтру, ранее установленная по мере численного отклонения [7], сохраняется, как показано выше, и на уровне откликов нелинейного детектора. Это обосновывает переход к реализации без аппаратных умножителей: отказ от них в пользу сдвигов и сложений не сопровождается ухудшением характеристик детекции.

Характер воздействия гауссовой фильтрации отличается от ранее исследованных операторов предобработки. Медианная фильтрация и гамма-коррекция влияли преимущественно на момент первого обнаружения; здесь же он остается неизменным, а изменяется устойчивость детекции. Это соответствует природе оператора. Линейный сглаживающий фильтр подавляет аддитивную высокочастотную составляющую шума и стабилизирует вход детектора, однако различимость объекта на самых ранних кадрах определяется его угловым размером, а не уровнем шума, и потому фильтрацией не улучшается.

Существенно, что наименьшее ядро $k = 3$ оказывается единственным значением, обеспечивающим прирост устойчивости при обоих порогах уверенности без деградации момента первого обнаружения. В варианте без аппаратных умножителей это ядро [1, 2, 1] с нормировкой сдвигом на два разряда – минимальная из возможных реализаций. Наименьшие ресурсные затраты и наилучшее качество детекции достигаются здесь одним и тем же выбором.

Отметим ограничение использованных метрик: показатель N_{det} учитывает все кадры с обнаружением объекта заданного класса и не разделяет истинные и ложные срабатывания. На исследуемой сцене, содержащей единственный объект интереса, это допущение приемлемо, однако при переносе методики на сцены с множественными объектами потребуются отдельный учет.

Постановка задачи адаптивного выбора размера ядра. Выбор целевого показателя определяется функциональным назначением оператора. В предшествующих работах, посвященных медианной фильтрации и гамма-коррекции, минимизировался момент первого обнаружения: обе операции воздействовали именно на этот показатель. Результаты показывают, что для гауссовой фильтрации такая постановка неадекватна – момент первого обнаружения от размера ядра практически не зависит, тогда как устойчивость детекции изменяется закономерно и монотонно. Целевым показателем поэтому принимается число кадров с обнаружением N_{det} .

Рост устойчивости не должен достигаться ценой ухудшения остальных характеристик. Задача формулируется как условная максимизация:

$$k^* = \arg \max_{k \in \{1, 3, 5, 7, 9\}} N_{\text{det}}(k) \quad (12)$$

при ограничениях

$$F_{\det}(k) \leq F_{\det}(1), \quad C_{\text{avg}}(k) \geq C_{\text{avg}}(1), \quad T_c \in \{0,25; 0,7\}, \quad (13)$$

где $F_{\det}(1)$ и $C_{\text{avg}}(1)$ – значения метрик базового варианта без фильтрации. Требование выполнения ограничений при обоих порогах уверенности одновременно введено для того, чтобы полученное решение не зависело от выбора рабочей точки детектора.

Ограничение на $F_{\det}$ имеет прямой физический смысл: латентность обнаружения определяет время реакции системы, и ее ухудшение недопустимо независимо от того, насколько выросла устойчивость. Ограничение на C_{avg} исключает решения, при которых число кадров с обнаружением увеличивается за счет срабатываний вблизи порога.

Проверка ограничений по данным Таблицы 3 дает следующее. Условие $F_{\det}(k) \leq F_{\det}(1)$ при базовом пороге нарушается только при $k = 9$ (198 против 197 кадров). При строгом пороге оно нарушается для всех $k \geq 5$, причем с существенным запасом – 292 кадра против 220. Ограничение на C_{avg} выполняется во всем диапазоне. Допустимое множество, таким образом, составляет $k \in \{1, 3\}$ (Рисунок 4).

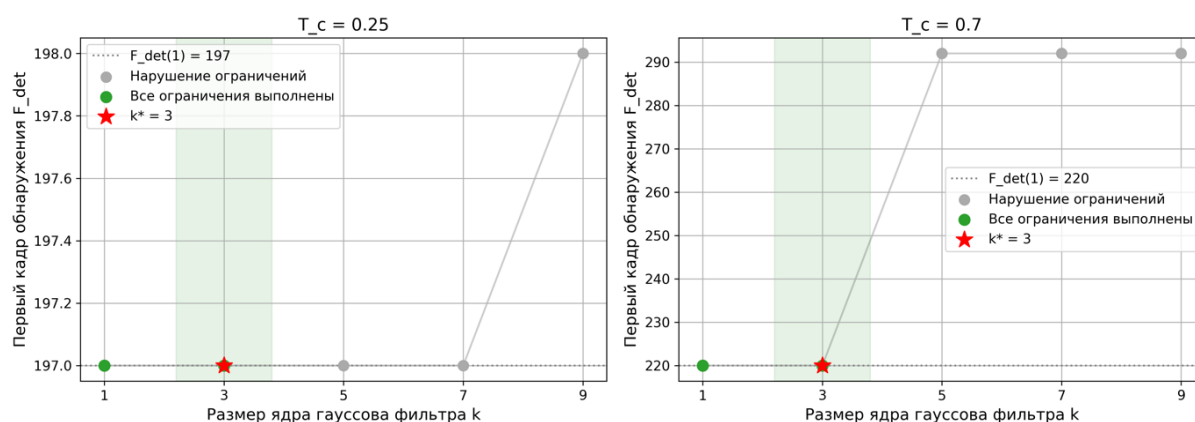


Рисунок 4 – Проверка ограничений: значения момента первого обнаружения при различных размерах ядра для обоих порогов уверенности, допустимые и недопустимые точки, отметка оптимума $k^* = 3$

Figure 4 – Verification of the constraints: first detection frame for different kernel sizes at both confidence thresholds, feasible and infeasible points, and the marked optimum $k^* = 3$

Из двух допустимых значений максимум устойчивости достигается при $k^* = 3$ (Рисунок 5). Прирост числа кадров с обнаружением составляет 27 кадров (3,8 %) при базовом пороге и 14 кадров (20,6 %) при строгом, причем момент первого обнаружения сохраняется на базовом уровне – 197 и 220 кадров соответственно. Средняя уверенность при строгом пороге превышает базовую в четвертом знаке (0,7649 против 0,7645), то есть ограничение выполняется на грани равенства; при базовом пороге запас существеннее (0,4343 против 0,4283).

Отметим, что базовое значение $k = 1$, соответствующее отсутствию фильтрации, также принадлежит допустимому множеству. Формально это означает, что отказ от гауссовой фильтрации не нарушает ни одного из ограничений и является корректным решением задачи – хотя и не оптимальным. Для адаптивной платформы это существенно: контур должен располагать возможностью не применять оператор, если условия сцены не обеспечивают выигрыша [15].

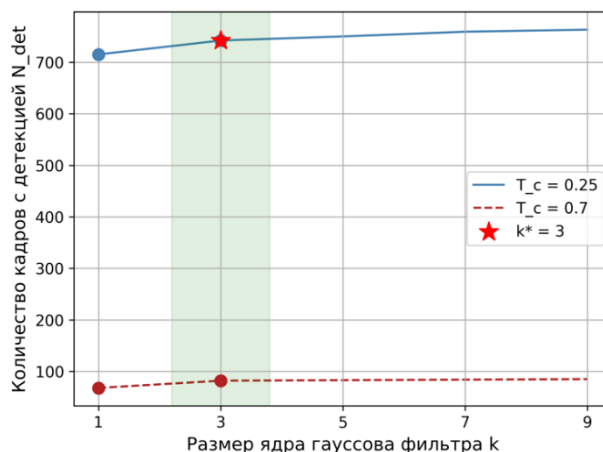


Рисунок 5 – Целевая функция (число кадров с обнаружением) в зависимости от размера ядра с выделением допустимого множества

Figure 5 – Objective function (number of frames with detection) versus the kernel size with the feasible set highlighted

Согласно полученным результатам, биномиальная реализация ядра $k^* = 3$ по метрикам детекции неотличима от точной. Следовательно, оптимальное решение задачи (12)–(13) реализуется одномерным ядром

$$g_3 = \frac{1}{4} [1, 2, 1], \quad (14)$$

применяемым последовательно по строке и по столбцу. Умножение на коэффициент 2 выполняется сдвигом на один разряд, умножение на 1 не требует вычислений, нормировка на 4 – сдвигом вправо на два разряда. Аппаратные умножители не задействуются; фильтр строится на двух буферах строк, сумматорах и сдвиговых соединениях.

Оценка ресурсных затрат оптимальной конфигурации. Изменение размера ядра в такой архитектуре сводится к записи значения k в регистры управления и переключению глубины буферов строк. Реконфигурация программируемой логики не требуется, детерминированная латентность обработки видеопотока сохраняется [16]. Это позволяет замкнуть контур адаптации без прерывания потоковой обработки.

Затраты на реализацию оптимального решения поддаются точной оценке. Одномерная свертка ядром длины k требует $k - 1$ сложений, следовательно сепарабельная схема при $k = 3$ выполняет четыре сложения на пиксель – по два в каждом проходе. Умножители отсутствуют полностью, поскольку оба коэффициента ядра $[1, 2, 1]$ тривиальны. Для сравнения, прямая двумерная свертка ядром 9×9 , соответствующая наибольшему из исследованных размеров, потребовала бы 81 умножения и 80 сложений на пиксель.

Расход блочной памяти определяется буферами строк. При ширине кадра 2560 отсчетов ядру $k = 3$ соответствуют два буфера. Разрядность их содержимого зависит от способа нормировки. Если нормировка выполняется однократно после обоих проходов, промежуточные отсчеты достигают величины $4 \times 255 = 1020$ и требуют десяти разрядов, что дает 6,25 КБ на два буфера. Если нормировка выполняется после каждого прохода, восьмизрядное представление сохраняется и объем составляет 5 КБ, ценой отбрасывания двух младших разрядов; при округлении к ближайшему вносимая погрешность не превышает половины единицы младшего разряда. Обе величины пренебрежимо малы относительно ресурсов блочной памяти современных кристаллов, поэтому выбор определяется требованиями к точности, а не дефицитом памяти.

Полная стоимость оптимального решения, таким образом, составляет четыре сумматора, два буфера строк и сдвиговые соединения при нулевом расходе блоков цифровой обработки сигналов. Наименьшая из реализуемых конфигураций фильтра оказывается одновременно наилучшей по качеству детекции – результат, не следующий из общих соображений и полученный только экспериментально.

Заключение

Установлено, что гауссова фильтрация воздействует на детекцию иначе, чем ранее исследованные операторы предобработки. Медианная фильтрация и гамма-коррекция изменяли момент первого обнаружения; здесь он остается практически неизменным во всем исследованном диапазоне, тогда как устойчивость детекции растет монотонно – число кадров с обнаружением увеличивается на 6,7 % при базовом пороге уверенности и на 25 % при строгом. Это соответствует природе линейного сглаживающего оператора: подавление высокочастотной составляющей шума стабилизирует вход детектора, но не повышает различимость объекта на ранних кадрах, где определяющим фактором остается его угловой размер.

Обнаружена верхняя граница допустимой области параметра. При строгом пороге уверенности и $k \geq 5$ момент первого обнаружения скачкообразно смещается с 220-го кадра на 292-й: увеличение ядра снижает текстурную характерность объекта малого углового размера, и уверенность детектора перестает достигать порога до момента его приближения к камере.

Показано, что близость биномиальной аппроксимации к точному гауссову фильтру, ранее установленная по мере численного отклонения, сохраняется и на уровне откликов нелинейного детектора: момент первого обнаружения при строгом пороге совпадает для обеих реализаций во всех точках, при базовом расходитсся не более чем на один кадр, число кадров с обнаружением – не более чем на 0,7 %. Наблюдаемые отклонения объясняются тем, что биномиальное ядро на 9–17 % уже точного при том же размере окна, и по направлению соответствуют этому различию. Это обосновывает переход к реализации без аппаратных умножителей: отказ от них в пользу сдвигов и сложений не сопровождается ухудшением характеристик детекции.

Задача выбора параметра формализована как максимизация устойчивости детекции при ограничениях на латентность обнаружения и среднюю уверенность, выполняемых одновременно при обоих порогах. Допустимое множество составили значения $k \in \{1, 3\}$, оптимальным оказалось $k^* = 3$. Прирост устойчивости достигает 20,6 % при строгом пороге при сохранении момента первого обнаружения на базовом уровне. Соответствующее ядро $[1, 2, 1]$ с нормировкой сдвигом на два разряда требует четырех сумматоров и двух буферов строк при нулевом расходе блоков цифровой обработки сигналов. Таким образом, наименьшая из реализуемых конфигураций фильтра оказалась одновременно наилучшей по качеству детекции – соотношение, не вытекающее из общих соображений и установленное только экспериментально.

Предложенная архитектура допускает изменение размера ядра записью в регистры управления и переключением глубины буферов строк – без реконфигурации программируемой логики и без прерывания видеопотока, что делает ее пригодной для встраивания в системы технического зрения реального времени. Направлениями дальнейшей работы являются аппаратная реализация на платформе Zynq UltraScale+ MPSoC с измерением фактических ресурсных затрат и латентности контура, а также совместная адаптация размера ядра с ранее исследованными параметрами медианной фильтрации и гамма-коррекции в составе единого тракта предобработки.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Медведева Е.В., Перевощикова А.А. Исследование нейросетевых алгоритмов распознавания объектов железнодорожной инфраструктуры на видеоизображениях. *Компьютерная оптика*. 2025;49(3):443–450. <https://doi.org/10.18287/2412-6179-CO-1563>
Medvedeva E.V., Perevoshchikova A.A. Research of neural network algorithms for recognizing railway infrastructure objects in video images. *Computer Optics*. 2025;49(3):443–450. (In Russ.). <https://doi.org/10.18287/2412-6179-CO-1563>
2. Redmon J., Farhadi A. *YOLOv3: An Incremental Improvement*. arXiv. URL: <https://arxiv.org/abs/1804.02767> [Accessed 15th May 2026].
3. Araki Y., Matsuda M., Manabe T., Ishizuka Y., Shibata Y. FPGA Implementation of an Object Recognition System with Low Power Consumption Using a YOLOv3-tiny-based CNN. In: *Complex, Intelligent and Software Intensive Systems: Proceedings of the 16th International Conference on Complex, Intelligent and Software Intensive Systems (CISIS-2022), 29 June – 01 July 2022, Virtual Event*. Cham: Springer; 2022. P. 223–233. https://doi.org/10.1007/978-3-031-08812-4_22
4. Gonzalez R.C., Woods R.E. *Digital Image Processing*. New York: Pearson; 2018. 1022 p.
5. Mostafa Sh.Sh., Sousa L.N., Ferreira N.F., et al. Full image-processing pipeline in field-programmable gate array for a small endoscopic camera. *Journal of Electronic Imaging*. 2017;26(1):013005. <https://doi.org/10.1117/1.JEI.26.1.013005>
6. Север К.О., Алексеев К.Н., Турулин И.И. Реализация эффективного разделимого двумерного цифрового фильтра на ПЛИС. *Известия ЮФУ. Технические науки*. 2024;(3):158–168. <https://doi.org/10.18522/2311-3103-2024-3-158-168>
Sever K.O., Alekseev K.N., Turulin I.I. Implementation of an efficient separable vector digital filter on FPGA. *Izvestiya SFedU. Engineering Sciences*. 2024;(3):158–168. (In Russ.). <https://doi.org/10.18522/2311-3103-2024-3-158-168>
7. Aubury M., Luk W. Binomial filters. *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*. 1996;12(1):35–50. <https://doi.org/10.1007/BF00936945>
8. Zhao M., Shi P., Xu X., Xu X., Liu W., Yang H. Improving the accuracy of an R-CNN-based crack identification system using different preprocessing algorithms. *Sensors*. 2022;22(18):7089. <https://doi.org/10.3390/s22187089>
9. Gholinavaz S., Saeedi N., Samadi Gharehveran S. Robustness analysis of YOLO and faster R-CNN for object detection in realistic weather scenarios with noise augmentation. *Scientific Reports*. 2025;15(1):44888. <https://doi.org/10.1038/s41598-025-28737-5>
10. Yano T., Kuroki Y. Fast implementation of Gaussian filter by parallel processing of binominal filter. In: *2016 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 24–27 October 2016, Phuket, Thailand*. IEEE; 2016. <https://doi.org/10.1109/ISPACS.2016.7824738>
11. Cambuim L., Barros E. FPGA-based pedestrian detection for collision prediction system. *Sensors*. 2022;22(12):4421. <https://doi.org/10.3390/s22124421>
12. Мурсаев А.Х. Аппаратная реализация медианного фильтра. *Научное приборостроение*. 2013;23(3):114–119.
Mursaev A.Kh. Hardware implementation of median filtering. *Nauchnoe priborostroenie*. 2013;23(3):114–119. (In Russ.).
13. Rodríguez-Andina J.J., Moure M.J., Valdés M.D. Features, design tools, and application domains of FPGAs. *IEEE Transactions on Industrial Electronics*. 2007;54(4):1810–1823. <https://doi.org/10.1109/TIE.2007.898279>

14. Wang J., Yang P., Liu Y., et al. Research on improved YOLOv5 for low-light environment object detection. *Electronics*. 2023;12(14):3089. <https://doi.org/10.3390/electronics12143089>
15. Гладков Л.А., Гладкова Н.В. Алгоритм синтеза комбинационных логических схем на основе эволюционного подхода. *Известия ЮФУ. Технические науки*. 2023;(6):24–34. <https://doi.org/10.18522/2311-3103-2023-6-24-34>
Gladkov L.A., Gladkova N.V. Algorithm for synthesis of combinational logic circuits based on the evolutionary approach. *Izvestiya SFedU. Engineering Sciences*. 2023;(6):24–34. (In Russ.). <https://doi.org/10.18522/2311-3103-2023-6-24-34>
16. Li X., Li Z., Zhou L., Huang Zh. FOLD: Low-level image enhancement for low-light object detection based on FPGA MPSoC. *Electronics*. 2024;13(1):230. <https://doi.org/10.3390/electronics13010230>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Гайсин Камиль Рафаэлевич, аспирант, Институт вычислительной математики и информационных технологий, Казанский (Приволжский) федеральный университет, Казань, Российская Федерация.
e-mail: gaisinkamil@yandex.ru
ORCID: [0009-0008-7003-0359](https://orcid.org/0009-0008-7003-0359)

Kamil R. Gaisin, Postgraduate, Institute of Computational Mathematics and Information Technology, Kazan (Volga Region) Federal University, Kazan, the Russian Federation.

Егорчев Антон Александрович, кандидат технических наук, доцент, директор Института вычислительной математики и информационных технологий, ведущий научный сотрудник, Казанский (Приволжский) федеральный университет, Казань, Российская Федерация.
e-mail: anton@egorchev.ru
ORCID: [0000-0001-8561-8616](https://orcid.org/0000-0001-8561-8616)

Anton A. Egorchev, Candidate of Engineering Sciences, Docent, Director of the Institute of Computational Mathematics and Information Technology, Leading Researcher, Kazan (Volga Region) Federal University, Kazan, the Russian Federation.

Азов Максим Сергеевич, кандидат технических наук, заместитель генерального директора по развитию гражданской продукции – главный конструктор, Акционерное общество «Ульяновское конструкторское бюро приборостроения», Ульяновск, Российская Федерация.
e-mail: m.azov@ukbp.ru

Maksim S. Azov, Candidate of Engineering Sciences, Deputy General Director for Civil Product Development – Chief Designer, Joint-Stock Company Ulyanovsk Instrument Manufacturing Design Bureau, Ulyanovsk, the Russian Federation.

Статья поступила в редакцию 04.08.2026; одобрена после рецензирования 13.10.2026; принята к публикации 21.10.2026.

The article was submitted 04.08.2026; approved after reviewing 13.10.2026; accepted for publication 21.10.2026.