УДК 004.93
DOI: 10.26102/2310-6018/2026.62.11.006

Алгоритмическое обеспечение нейросетевого обнаружения малоразмерных объектов в условиях высокой априорной неопределенности оптико-электронных данных

Е.Ю. Кирюшин[1]✉, А.В. Кошелев[2]

[1]Уральский оптико-механический завод имени Э.С. Яламова, Екатеринбург, Российская Федерация
[2]Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация

Резюме. В статье рассматривается новый алгоритм обнаружения малоразмерных (точечных) целей на изображениях средневолнового инфракрасного диапазона, функционирующий в условиях высокой априорной неопределенности оптико-электронных данных и низкого отношения сигнал/шум. Предложенная стратегия обнаружения основана на каскадной многоэтапной обработке, последовательно снижающей неопределенность исходных данных. Алгоритм включает в себя поиск локальных максимумов с последующим пороговым отсеиванием заведомо ложных кандидатов по амплитудным признакам, аппроксимацию окрестности методом Гаусса-Ньютона и финальную классификацию целеподобных объектов с помощью компактной четырехслойной сверточной нейронной сети. При этом пороговая фильтрация значительно сокращает число ложных тревог еще до применения вычислительно затратной нейросетевой классификации. Для обеспечения возможности интеграции алгоритма во встраиваемое программное обеспечение на базе программируемых логических интегральных схем применено статическое квантование параметров нейросети до формата INT8. Экспериментальные исследования на реальных изображениях средневолнового инфракрасного диапазона подтвердили высокую эффективность разработанного подхода: модель обеспечивает правильность обнаружения не менее 0,95 при интегральной метрике качества AUC = 0,96, а сокращение объема памяти под веса сети в 3,14 раза (до 632 байт) позволяет использовать алгоритм в ресурсоограниченных бортовых системах реального времени.

Ключевые слова: точечная цель, инфракрасные изображения, компьютерное зрение, каскадный алгоритм обнаружения, метод Гаусса-Ньютона, сверточная нейронная сеть, статическое квантование, INT8, встраиваемые системы, ПЛИС.

Для цитирования: Кирюшин Е.Ю., Кошелев А.В. Алгоритмическое обеспечение нейросетевого обнаружения малоразмерных объектов в условиях высокой априорной неопределенности оптико-электронных данных. *Моделирование, оптимизация и информационные технологии.* 2026;14(11). URL: https://moitvivt.ru/ru/journal/article?id=2597 DOI: 10.26102/2310-6018/2026.62.11.006

Algorithmic support for neural network detection of small objects in conditions of high a priori uncertainty of optoelectronic data

E.Y. Kiryushin[1]✉, A.V. Koshelev[2]

[1]Ural Optical and Mechanical Plant named after Mr. E.S. Yalamov, Ekaterinburg, the Russian Federation
[2]Peter the Great St. Petersburg Polytechnic University, St. Petersburg, the Russian Federation

© Кирюшин Е.Ю., Кошелев А.В., 2026

Abstract. The article discusses a new algorithm for detecting small-size (point) targets in images of the mid-wavelength infrared range, operating under conditions of high a priori uncertainty of optoelectronic data and low signal-to-noise ratio. The proposed detection strategy is based on cascading multi-stage processing, which consistently reduces the uncertainty of the initial data. The algorithm includes a search for local maxima followed by threshold screening of deliberately false candidates based on amplitude characteristics, neighborhood approximation using the Gauss-Newton method, and the final classification of target-like objects using a compact four-layer convolutional neural network. At the same time, threshold filtering significantly reduces the number of false alarms even before applying computationally expensive neural network classification. To enable the integration of the algorithm into embedded software based on FPGA-based embedded software, static quantization of neural network parameters up to INT8 format has been applied. Experimental studies on real images of the mid-wave infrared range have confirmed the high efficiency of the developed approach: the model ensures the accuracy of detection of at least 0.95 with an integral quality metric of AUC = 0.96, and reducing the amount of memory for network weights by 3.14 times (up to 632 bytes) allows the algorithm to be used in resource-limited on-board real-time systems.

Keywords: point target, infrared images, computer vision, cascade detection algorithm, Gauss-Newton method, convolutional neural network, static quantization, INT8, embedded systems, FPGA.

For citation: Kiryushin E.Y., Koshelev A.V. Algorithmic support for neural network detection of small objects in conditions of high a priori uncertainty of optoelectronic data. *Modeling, Optimization and Information Technology.* 2026;14(11). (In Russ.). URL: https://moitvivt.ru/ru/journal/article?id=2597 DOI: 10.26102/2310-6018/2026.62.11.006

Введение

Обнаружение малоразмерных инфракрасных объектов на изображениях, формируемых оптико-электронными системами, является одной из ключевых научно-технических задач в области компьютерного зрения и обработки сигналов. Практическая реализация подобных алгоритмов особенно критична при создании программно-аппаратных комплексов для беспилотных авиационных систем, систем мониторинга воздушного пространства и автономной навигации, где требуется комплексное обеспечение функциональной надежности бортового оборудования и защиты каналов управления [1]. Под точечной целью в данном контексте подразумевается объект, чья проекция на светочувствительную матрицу приемника меньше площади одного пикселя. Оптическая система формирует изображение такой цели в виде непрерывной функции рассеивания точки (ФРТ), которая подвергается деформации в процессе пространственно-яркостной дискретизации аналого-цифровым преобразователем. Дополнительные факторы, такие как тепловой шум, дефекты пикселей фотоприемной матрицы, динамическая солнечная засветка и блики, создают экстремальную высокую априорную неопределенность исходных данных, существенно снижая отношение сигнал/шум.

При решении указанной проблемы в современной литературе применяются различные подходы, основанные на многоэтапной фильтрации и каскадной обработке информации. Как правило, реализуемая стратегия сочетает в себе детекцию локальных экстремумов с последующим предварительным отсеиванием ложных кандидатов методами пространственной фильтрации [2, 3]. Для оценки параметров формы функции рассеяния точки и подавления фоноцелевой обстановки проводится аппроксимация окрестности каждого пика аналитическими функциями [4]. Завершающим этапом современных высокоточных систем является интеллектуальная классификация с применением глубоких сверточных нейронных сетей, обладающих высокой обобщающей способностью [5, 6]. Такой каскад позволяет поэтапно снижать неопределенность данных. Однако развертывание нейросетевых моделей в бортовых

вычислителях реального времени жестко ограничено вычислительной емкостью целевых платформ. В связи с этим в последнее время особое внимание уделяется методам статистического и динамического квантования параметров нейросетей, а также специализированной пороговой фильтрации [7, 8]. Такие подходы позволяют значительно снизить аппаратные и вычислительные затраты процессоров без критической потери чувствительности всей системы [9, 10].

Несмотря на обилие существующих решений, сохраняется научно-практическое противоречие между высокой точностью обнаружения, обеспечиваемой глубокими моделями, и жесткими ресурсными ограничениями (память, энергопотребление), накладываемыми встраиваемым программным обеспечением микроэлектронных устройств.

Целью данного исследования является разработка и верификация высокоэффективного и аппаратно-оптимизированного каскадного алгоритма нейросетевого обнаружения точечных целей в средневолновом инфракрасном диапазоне, адаптированного для реализации во встраиваемом программном обеспечении (ПО) на базе программируемых логических интегральных схем.

Материалы и методы

Математическая формализация. На вход алгоритма поступает один кадр видео, который представляет собой одноканальное изображение с глубиной цвета 16 бит и размером 640 пикселей по горизонтали (координата x) и 512 пикселей по вертикали (координата y). Примем ФРТ за гауссову – тогда изображение точечной цели, формируемое на приемнике, имеет распределение в форме двумерной функции Гаусса:

$$g(x,y) = B + A\,exp\left(-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}\right), \quad (1)$$

где B – фон; A – амплитуда; x_0, y_0 – положение пика относительно центра окрестности; σ_x, σ_y – среднеквадратическое отклонение.

Для прохода по массиву организуются циклы с допущением, что в крайней рамке толщиной 3 пикселя целей нет. Также, в связи с попаданием в поле зрения элементов конструкции камеры, в расчет не берутся пиксели, выходящие за границы ограничивающей эффективное поле зрения окружности. Для каждого кадра параметры окружности (координаты центра (x_c, y_c) и радиус r) должны подбираться индивидуально.

Пиксель находится внутри окружности, если выполняется условие:

$$(x - x_c)^2 + (y - y_c)^2 < r^2. \quad (2)$$

Пример изображения с визуализированной окружностью показан на Рисунке 1. Эффективное поле зрения составляет 91 % от размера кадра.

Рассмотрим k-й пиксель с координатами (x, y) на изображении, где $k = y \cdot 640 + x$ – индекс пикселя в массиве данных, и его окрестность 7×7 (Рисунок 2). Проверим пиксель на строгий максимум в серой области и нестрогий максимум в оранжевой области. Прошедший проверку пиксель обозначим $\max_{7\times7}$.

В среднем в одном кадре содержится около 2000 локальных максимумов. Перед аппроксимацией функции (1) в целях снижения вычислительной нагрузки необходимо отсеять как можно больше заведомо ложных локальных максимумов.

Для этого в окрестности 5×5 найденного локального максимума вычисляется также минимальное значение $\min_{5\times5}$ и амплитуда $A_0 = \max_{7\times7} - \min_{5\times5}$.

На Рисунке 3 приведены графики функций распределения плотности вероятности амплитуды локальных максимумов и целей.

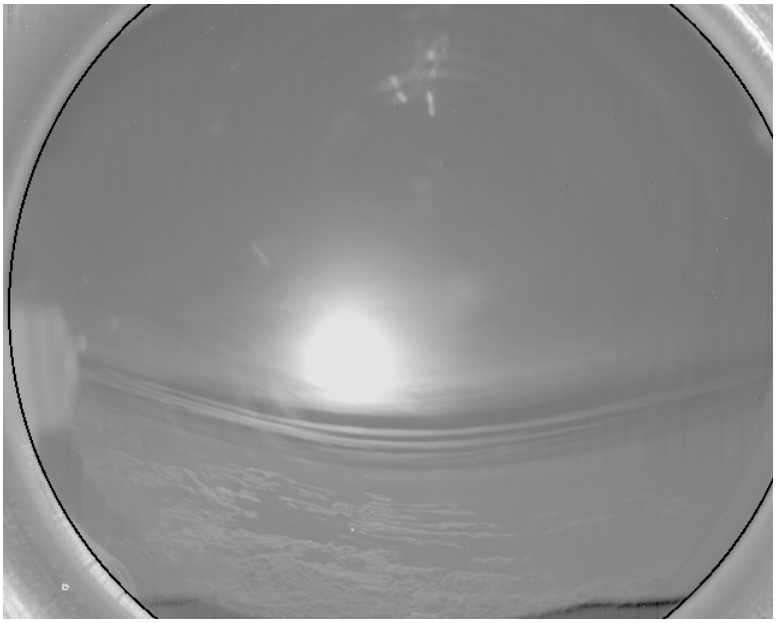


Рисунок 1 – Пример изображения
Figure 1 – Sample image

	x-4	x-3	x-2	x-1	x	x+1	x+2	x+3	x+4
y-4	-2564	-2563	-2562	-2561	-2560	-2559	-2558	-2557	-2556
y-3	-1924	-1923	-1922	-1921	-1920	-1919	-1918	-1917	-1916
y-2	-1284	-1283	-1282	-1281	-1280	-1279	-1278	-1277	-1276
y-1	-644	-643	-642	-641	-640	-639	-638	-637	-636
y	-4	-3	-2	-1	k	+1	+2	+3	+4
y+1	+636	+637	+638	+639	+640	+641	+642	+643	+644
y+2	+1276	+1277	+1278	+1279	+1280	+1281	+1282	+1283	+1284
y+3	+1916	+1917	+1918	+1919	+1920	+1921	+1922	+1923	+1924
y+4	+2556	+2557	+2558	+2559	+2560	+2561	+2562	+2563	+2564

Рисунок 2 – Окрестность 7×7 *k*-го пикселя
Figure 2 – The neighborhood of the 7×7 k^{th} pixel

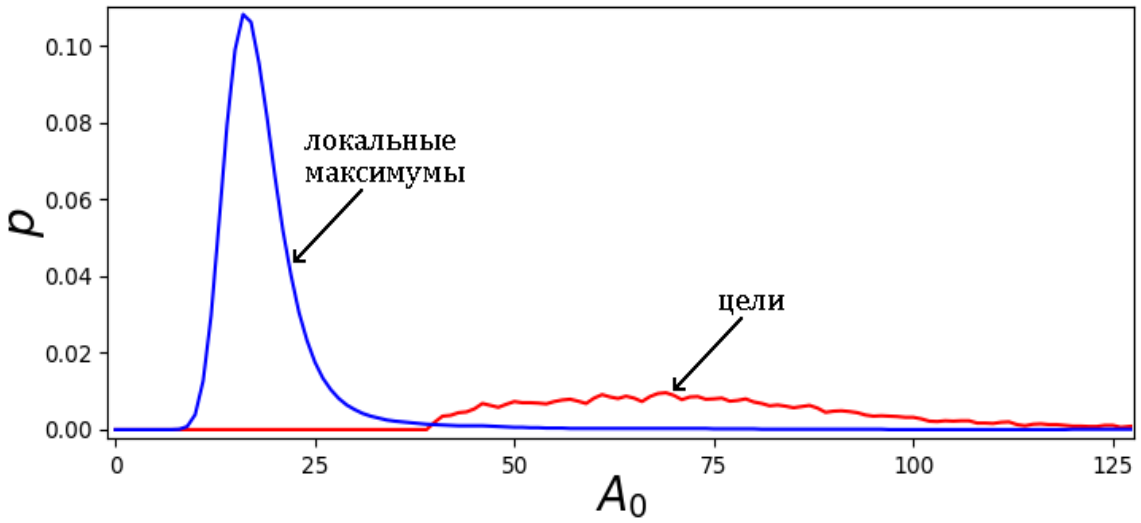


Рисунок 3 – Графики функций распределения плотности вероятности амплитуды
Figure 3 – Graphs of probability density distribution functions of amplitude

Из графиков видно, что большую часть локальных максимумов можно отсечь, поставив условие $A_0 > 40$.

Для следующего отсеивания рассчитываются средние амплитуды A_1 и A_2 в рамках 3×3 и 5×5 соответственно (зеленая и оранжевая области на Рисунке 4). Эти величины коррелируют с параметрами σ_x и σ_y функции (1).

	x-4	x-3	x-2	x-1	x	x+1	x+2	x+3	x+4
y-4	-2564	-2563	-2562	-2561	-2560	-2559	-2558	-2557	-2556
y-3	-1924	-1923	-1922	-1921	-1920	-1919	-1918	-1917	-1916
y-2	-1284	-1283	-1282	-1281	-1280	-1279	-1278	-1277	-1276
y-1	-644	-643	-642	-641	-640	-639	-638	-637	-636
y	-4	-3	-2	-1	k	+1	+2	+3	+4
y+1	+636	+637	+638	+639	+640	+641	+642	+643	+644
y+2	+1276	+1277	+1278	+1279	+1280	+1281	+1282	+1283	+1284
y+3	+1916	+1917	+1918	+1919	+1920	+1921	+1922	+1923	+1924
y+4	+2556	+2557	+2558	+2559	+2560	+2561	+2562	+2563	+2564

Рисунок 4 – Рамки 3×3 и 5×5 вокруг локального максимума
Figure 4 – 3×3 and 5×5 frames around the local maximum

На Рисунке 5 приведена точечная диаграмма зависимости значений A_1 и A_2. Красными крестиками отмечены цели, синими точками – локальные максимумы. Заведомо ложные локальные максимумы отсекаются плоскостями, задаваемыми двумя точками (x_1, y_1) и (x_2, y_2), где $x_1, x_2 \in A_1$, $y_1, y_2 \in A_2$.

Положение точки относительно прямой определяется по знаку расстояния между данной точкой и прямой. Расстояние определяется по формуле:

$$D = A_1 + \frac{y_2 - y_1}{x_1 - x_2} A_2 + \frac{x_2 y_1 - x_1 y_2}{x_1 - x_2}. \quad (3)$$

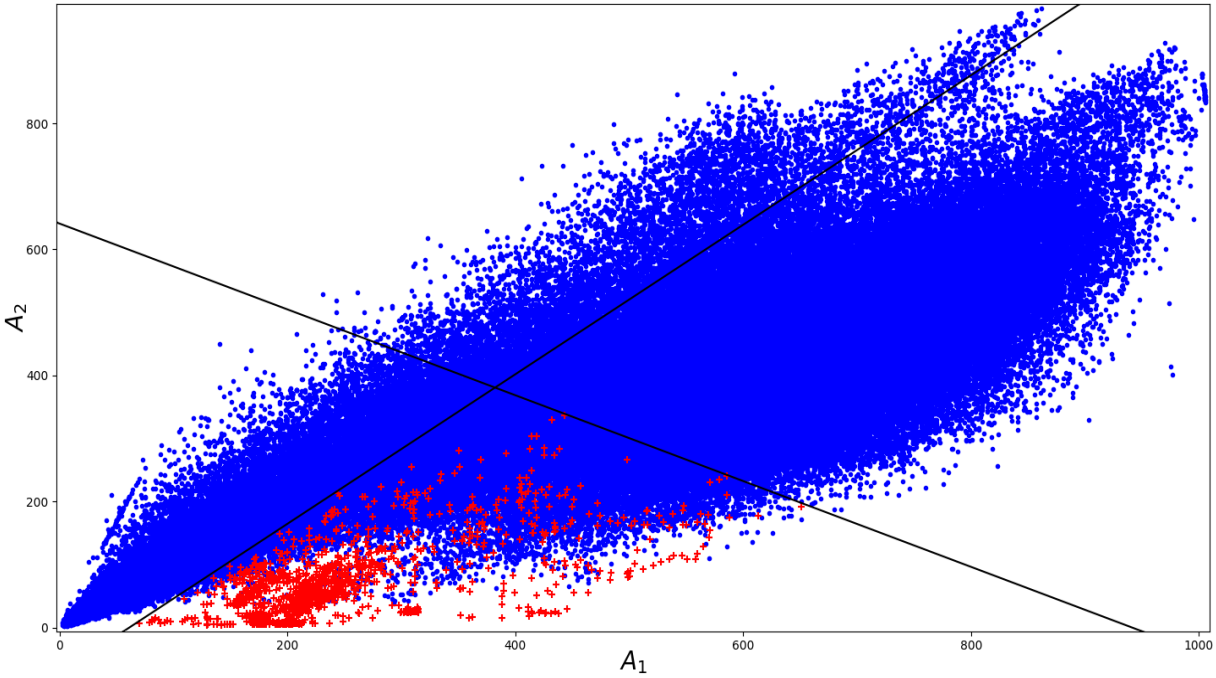


Рисунок 5 – Отсеивание по A_1 и A_2
Figure 5 – Screening by A_1 and A_2

Аппроксимация ФРТ. Функция (1) аппроксимируется в окрестности 5×5 локального максимума.

В целях оптимизации преобразуем функцию (1): чтобы избавиться от фона B примем его равным $min_{5\times5}$ и заранее будем вычитать из данных, величины x_0 и y_0 примем равными нулю и произведем замену:

$$\alpha = -\frac{1}{2\sigma_x^2}, \quad (4)$$

$$\beta = -\frac{1}{2\sigma_y^2}. \quad (5)$$

В итоге функция (1) приняла вид:

$$g(x, y) = A\, exp(\alpha x^2 + \beta y^2). \quad (6)$$

Для аппроксимации используется алгоритм Гаусса-Ньютона. Зададим начальные значения параметров A, α, β функции (6):

$$A = max_{7\times7} - min_{5\times5}, \quad (7)$$

$$\alpha = \beta = -\frac{1}{2 \cdot 0{,}9 \cdot 0{,}9}. \quad (8)$$

Вектор начального приближения:

$$p^0 = [A \quad \alpha \quad \beta]^T. \tag{9}$$

Осуществим итерации:

$$p^{s+1} = p^s + (J^T J)^{-1} J^T r(p^s), \tag{10}$$

где J – матрица Якоби, составленная из частных производных функции (6) по переменным A, α, β; $r(p^s)$ – остаточные ошибки.

$$J = \begin{bmatrix} \frac{\partial g_1}{\partial A} & \frac{\partial g_1}{\partial \alpha} & \frac{\partial g_1}{\partial \beta} \\ \frac{\partial g_2}{\partial A} & \frac{\partial g_2}{\partial \alpha} & \frac{\partial g_2}{\partial \beta} \\ \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial A} & \frac{\partial g_n}{\partial \alpha} & \frac{\partial g_n}{\partial \beta} \end{bmatrix}, \tag{11}$$

$$\frac{\partial g}{\partial A} = exp(\alpha x^2 + \beta y^2), \tag{12}$$

$$\frac{\partial g}{\partial \alpha} = A\, exp(\alpha x^2 + \beta y^2)\, x^2, \tag{13}$$

$$\frac{\partial g}{\partial \beta} = A\, exp(\alpha x^2 + \beta y^2)\, y^2. \tag{14}$$

После аппроксимации происходит отсеивание максимумов с параметрами (6), статистически несоответствующими целям (Рисунок 6). Вычисления аналогичны формуле (3).

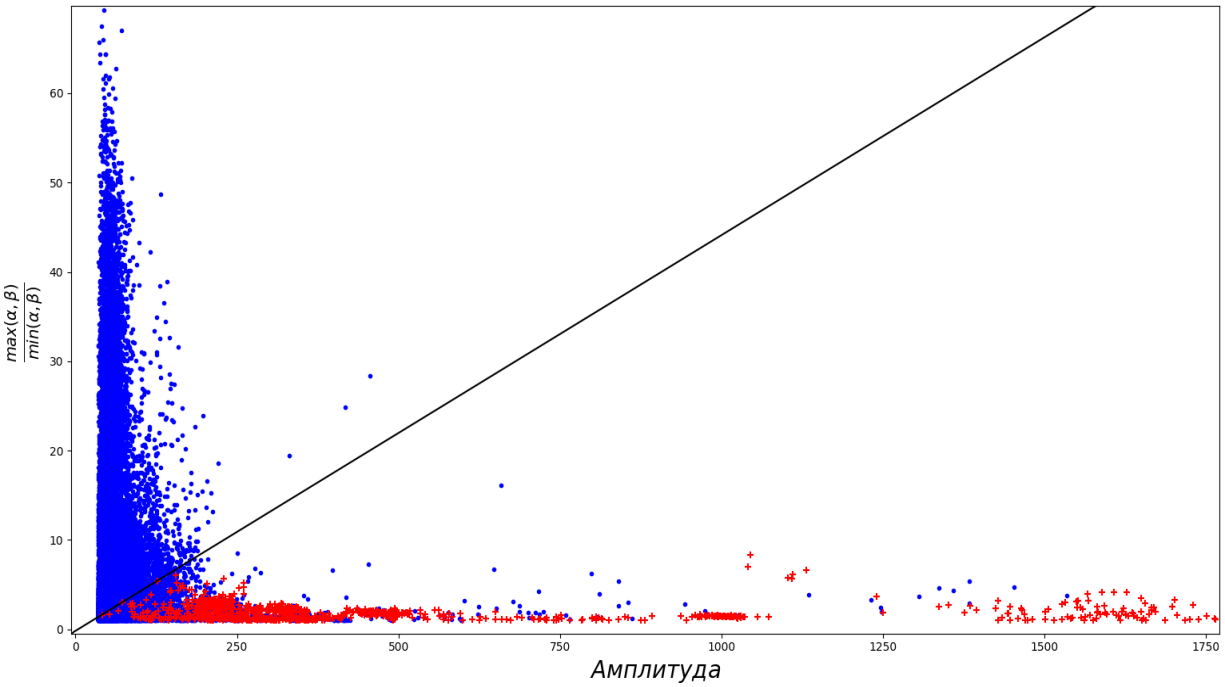


Рисунок 6 – Отсеивание по параметрам функции
Figure 6 – Screening by function parameters

Результаты отсеивания приведены в Таблице 1.

Таблица 1 – Предварительное отсеивание
Table 1 – Preliminary screening

Файл данных	MW0007.dat	dump.dat	MW0010.dat
Среднее количество локальных максимумов на 1 кадр	1796	2493	1431
После отсеивания по амплитуде	658	1166	170
После отсеивания по avg_1 и avg_2	77	220	13
После аппроксимации	65	130	9

Нейросетевая классификация. На вход четырехслойной сверточной нейронной сети после масштабирования подается фрагмент изображения, охватываемый окрестностью 7×7 пикселей с центром в локальном максимуме, содержащий целеподобный объект. На выходе получаем вероятность наличия на изображении цели. Архитектура разработана с учетом небольшого размера изображения и особенностей операции свертки. Описание архитектуры представлено ниже в виде кода на PyTorch и на Рисунке 7.

```
class CNN(nn.Module):
    def __init__(self):
        super(CNN, self).__init__()
        self.conv1 = nn.Conv2d(1, 4, 3)
        self.conv2 = nn.Conv2d(4, 4, 3)
        self.conv3 = nn.Conv2d(4, 8, 3)
        self.fc = nn.Linear(8, 1)
    def forward(self, x):
        x = nn.ReLU() (self.conv1(x))
        x = nn.ReLU() (self.conv2(x))
        x = nn.ReLU() (self.conv3(x))
        x = x.view(-1, 8)
        x = nn.Sigmoid() (self.fc(x))
        return x
```

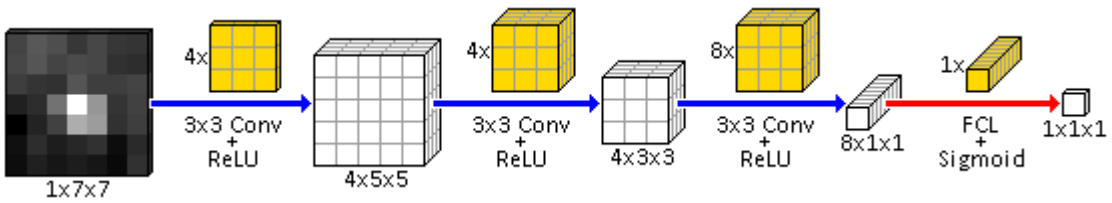


Рисунок 7 – Архитектура сверточной нейронной сети
Figure 7 – Architecture of a convolutional neural network

Сверточный слой выделяет признаки на входном изображении и формирует карту признаков. Карта признаков – это массив матриц, в котором каждая матрица (канал) отвечает за какой-нибудь выделенный признак.

Для того чтобы слой мог выделять признаки, в нем имеются фильтры (ядра). Фильтры имеют один и тот же размер, например 4×3×3 (показаны над стрелками на Рисунке 7), а их количество определяет глубину выходного 3D массива. При этом глубина самих фильтров совпадает с количеством каналов входного изображения. Так, если на вход сверточному слою подается одноканальное изображение и требуется карта

признаков, состоящая из 4 каналов, то сверточный слой будет содержать в себе 4 фильтра глубиной 1.

Квантование параметров нейронной сети. Квантование данных в цифровых устройствах, как правило, ассоциируется с потерей точности, однако в ряде практических задач оно позволяет существенно сократить аппаратные затраты без ущерба для быстродействия. В представленном случае анализ временных характеристик показывает, что время выполнения операции на ПЛИС как до, так и после применения квантования остается неизменным и составляет 1 мкс. При этом основная цель введения квантования – не повышение скорости, а экономия ресурсов кристалла. Уменьшение разрядности обрабатываемых данных за счет квантования приводит к сокращению числа требуемых регистров и упрощению комбинаторных схем, что особенно важно при ограниченной логической емкости ПЛИС.

Для квантования сверточной нейронной сети применяется метод статического квантования после обучения. Модель предварительно обучается в формате FP32, а затем преобразуется в формат INT8 (диапазон значений [−128,127]), что позволяет теоретически в 4 раза уменьшить размер модели и в 4 раза снизить требования к пропускной способности памяти. Применяется симметричное квантование для весов и несимметричное для активаций, смещения не квантуются. Активации объединяются с предыдущими слоями.

Формулы квантования для весов и активаций:

$$w_q = round\left(\frac{w}{s_w}\right), \tag{15}$$

$$x_q = round\left(\frac{x}{s_x} + z_x\right). \tag{16}$$

Формулы деквантования для весов и активаций:

$$w_{dq} = s_w w_q \approx w, \tag{17}$$

$$x_{dq} = s_x\left(x_q - z_x\right) \approx x, \tag{18}$$

где w, x – значения весов и активаций в формате FP32; s_w, s_x – положительное число с плавающей запятой; z_x – целое число.

Значения s и z определяются по формулам:

$$s = \frac{v_{max} - v_{min}}{\beta_q - \alpha_q}, \tag{19}$$

$$z = round\left(\frac{\alpha_q v_{max} - \beta_q v_{min}}{v_{max} - v_{min}}\right), \tag{20}$$

где v_{min}, v_{max} – минимальное и максимальное значения квантуемых параметров в формате FP32; α_q, β_q – минимальное и максимальное значения диапазона целочисленного типа данных.

Операцию свертки можно представить в виде матричного умножения:

$$Y_{dq} = b + \sum X_{dq} W_{dq} = b + \sum s_x\left(X_q - z_x\right) s_w W_q = b + s_x s_w\left[\sum X_q W_q - z_x \sum W_q\right]. \tag{21}$$

Также:

$$Y_{dq} = s_y\left(Y_q - z_y\right). \tag{22}$$

При слиянии свертки с функцией активации ReLU для получения выходного значения используется формула:

$$Y_q = \begin{cases} z_y, & Y_{dq} \le 0 \\ z_y + \frac{Y_{dq}}{s_y}, & Y_{dq} > 0 \end{cases}. \quad (23)$$

Подставим выражение для Y_{dq} (21) в формулу (23), после преобразований получим:

$$Y_q = \begin{cases} z_y, & \sum X_q W_q \le z_x \sum W_q - \frac{b}{s_x s_w} \\ \frac{s_x s_w}{s_y} \sum X_q W_q + z_y - \frac{1}{s_y}\left(s_x s_w z_x \sum W_q - b\right), & \sum X_q W_q > z_x \sum W_q - \frac{b}{s_x s_w} \end{cases}. \quad (24)$$

Введем переменные:

$$ReLU_0 = round\left(z_x \sum W_q - \frac{b}{s_x s_w}\right), \quad (25)$$

$$ReLU_a = 2^n \frac{s_x s_w}{s_y}, \quad (26)$$

$$ReLU_b = 2^n \left(z_y - \frac{1}{s_y}\left(s_x s_w z_x \sum W_q - b\right)\right). \quad (27)$$

Выражения (25), (26) и (27) зависят только от параметров, жестко заданных после обучения, поэтому могут быть рассчитаны заранее и приведены к формату INT32 округлением или умножением на масштабный коэффициент 2^n.

Итоговая формула для первых трех слоев:

$$Y_q = \begin{cases} z_y, & \sum X_q W_q \le ReLU_0 \\ \frac{1}{2^n}\left(ReLU_a \sum X_q W_q + ReLU_b\right), & \sum X_q W_q > ReLU_0 \end{cases}. \quad (28)$$

Выражение для последнего полносвязного слоя с сигмоидальной функцией активации имеет вид:

$$\frac{1}{1+e^{-Y_{dq}}} > p, \quad (29)$$

где p – вероятность, выше которой локальный максимум считается целью, в данной работе имеет значение 0,75.

Подставим выражение для Y_{dq} (21) в формулу (29), после преобразований получим:

$$\sum X_q W_q > \frac{\ln\frac{p}{1-p} - b}{s_x s_w} + z_x \sum W_q. \quad (30)$$

Введем переменную:

$$p_{new} = round\left(\frac{\ln\frac{p}{1-p} - b}{s_x s_w} + z_x \sum W_q\right). \quad (31)$$

Выражение (31) зависит только от параметров, жестко заданных после обучения, поэтому может быть рассчитано заранее и приведено к формату INT32 округлением.

Итоговая формула для последнего четвертого слоя:

$$\sum X_q W_q > p_{new}. \quad (32)$$

После применения статического метода квантования сети, модель стала весить 632 байта (Таблица 2), что в ≈3,14 раза меньше исходной модели с плавающей запятой. Потеря точности при переходе от FP32 к INT8 составила не более 1 %.

Таблица 2 – Количество параметров нейронной сети
Table 2 – Number of neural network parameters

Параметры	До квантования			После квантования		
	Кол-во	Тип	Размер, байт	Кол-во	Тип	Размер, байт
Коэффициенты z-нормализации	2	FP32	8	2	INT32	8
Веса	476	FP32	1904	476	INT8	476
Смещения	17	FP32	68			
ReLU0				16	INT32	64
ReLUa				3	INT32	12
ReLUb				16	INT32	64
n				4	INT8	4
p	1	FP32	4	1	INT32	4
Итого:			1984			632

Результаты

Разработанный в ходе выполнения данной работы алгоритм был протестирован на персональном компьютере на различных данных. Алгоритм эффективно отсеивает шум, локальные засветки, блики и некорректно работающие пиксели. Результаты сведены в Таблицу 3.

Таблица 3 – Точностные характеристики
Table 3 – Precision characteristics

Файл данных	MW0007.dat	dump.dat	MW0010.dat
Количество кадров	900	960	736
Количество целей	437	960	0
Ошибка первого рода	123	269	92
Ошибка второго рода	23	5	–
Точность	77,1%	78,0%	–
Полнота	94,7%	99,5%	–

Для наглядной оценки эффективности работы предложенного алгоритма была построена и проанализирована зависимость вероятности правильного обнаружения от вероятности ложной тревоги, известная в теории статистических решений как ROC-кривая (Receiver Operating Characteristic) (Рисунок 8). Данная характеристика позволяет наглядно оценить компромисс между чувствительностью алгоритма и его специфичностью. Величина AUC (Area Under Curve) представляет собой площадь под ROC-кривой и служит интегральной метрикой качества работы бинарного классификатора, не зависящей от выбора конкретного порога принятия решения. Полученное в ходе экспериментов значение AUC = 0,96 является крайне высоким и свидетельствует об отличной различительной способности разработанного алгоритма.

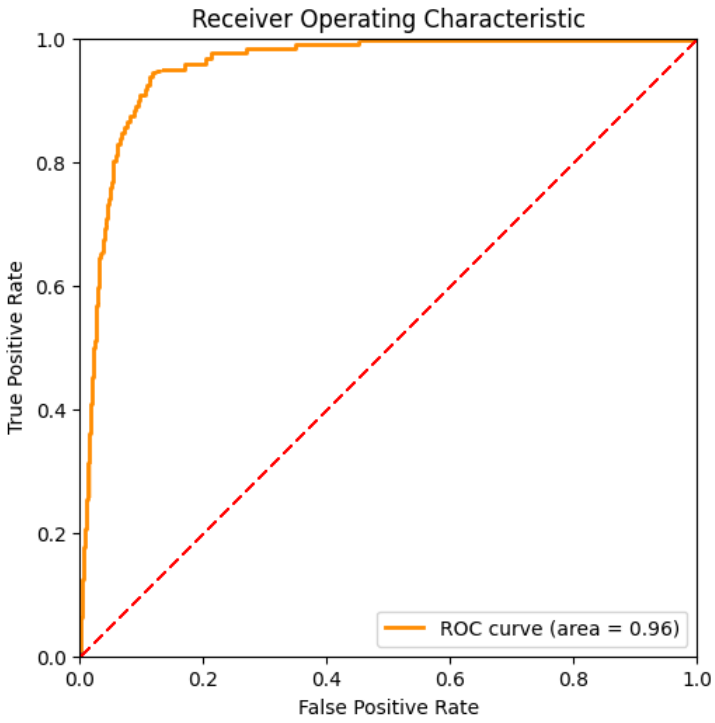


Рисунок 8 – ROC-кривая
Figure 8 – The ROC curve

Пример объекта, классифицированного как «цель», приведен на Рисунке 9.

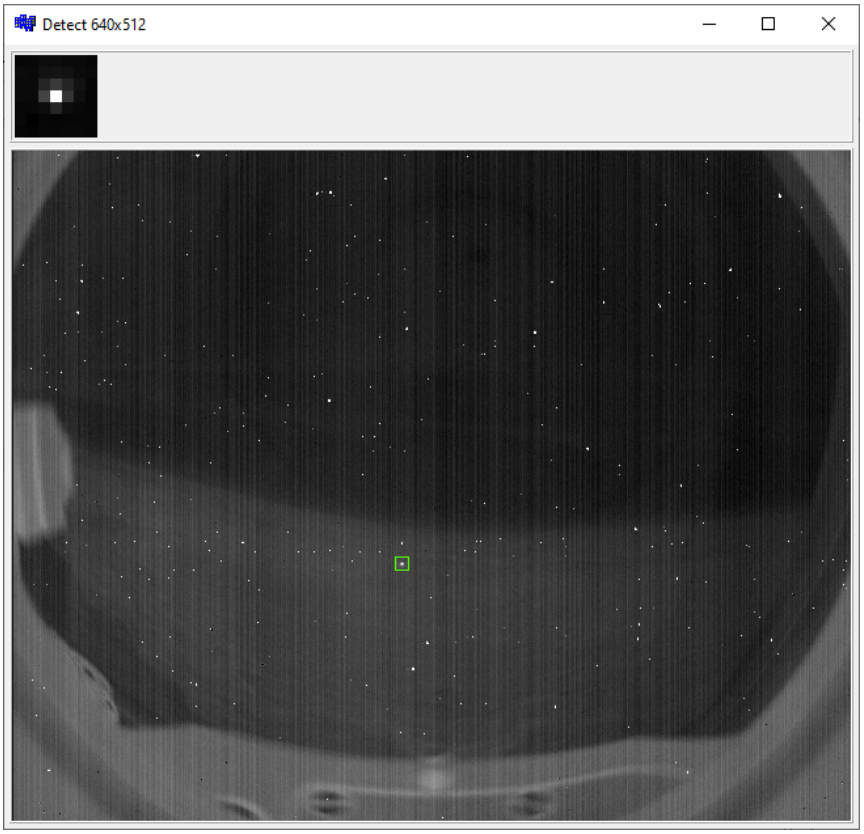


Рисунок 9 – Обнаруженный объект
Figure 9 – Detected object

Заключение

В ходе исследования характеристик алгоритма обнаружения точечных целей в средневолновом инфракрасном диапазоне было установлено:

– Использование примененного в алгоритме подхода по предварительной фильтрации целей на основе локальных максимумов позволяет отсеивать большую часть ложных целей. В представленной выборке не произошло ни одного отсеивания истинной цели до поступления на вход нейросетевого модуля алгоритма.

– Алгоритм имеет высокие степени вероятности обнаружения точечных целей (Таблица 3) в условиях наличия большого количества ложных целей на изображении (различные шумы, солнечные блики). На рассматриваемой тестовой выборке полнота обнаружения составляет 94,7 %, а точность распознавания составляет 77,1 %. Данные параметры достигнуты за счет применения сверточной нейронной сети и проведенного дообучения по имеющимся исходным данным.

– Модель нейронной сети занимает 632 байта, что позволяет ее применить во встраиваемом ПО на ПЛИС для параллельной обработки кадров изображений из разных информационных каналов.

– Все еще имеется значительная вероятность получения ложных целей на выходе модуля сопровождения целей при определенных условиях (например, перемещающийся по фотоприемной области солнечный блик). В дальнейшем необходимо рассмотреть возможность использования рекуррентных нейронных сетей с целью дальнейшего совершенствования отсеивания ложных целей.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Wang Y., Gao K., Gebel E.S. End-To-End Mapless Navigation in Environments with Mixed Obstacles. In: *2026 12th International Conference on Automation, Robotics and Applications (ICARA), 05–07 February 2026, Istanbul, Turkiye.* IEEE; 2026. P. 123–127. https://doi.org/10.1109/ICARA69401.2026.11480297
2. Костров В.В., Ракитин А.В., Богатов А.Д. Алгоритмическое обеспечение обнаружения и координатометрии точечных объектов в оптико-электронных системах с болометрическим матричным приемником. *Радиотехнические и телекоммуникационные системы.* 2025;(4):44–54. https://doi.org/10.24412/2221-2574-2025-4-44-54
   Kostrov V.V., Rakitin A.V., Bogatov A.D. Algorithmic support for the detection and coordination of point objects in optoelectronic systems with a bolometric matrix receiver. *Radio Engineering and Telecommunication Systems.* 2025;(4):44–54. (In Russ.). https://doi.org/10.24412/2221-2574-2025-4-44-54
3. Знаменский И.В. Зависимость величины сигнала от смещения функции рассеяния точки относительно центра пиксела фоточувствительной матрицы. *Оптический журнал.* 2023;90(12):96–110. https://doi.org/10.17586/1023-5086-2023-90-12-96-110
   Znamenskiy I.V. Dependence of the signal magnitude on displacement of the point spread function relative to the center of the photosensitive matrix pixel. *Opticheskii Zhurnal.* 2023;90(12):96–110. (In Russ.). https://doi.org/10.17586/1023-5086-2023-90-12-96-110
4. McKay M.B., Lawrence N.P., Wetton B., Gopaluni R.B. *Error whitening: Why Gauss-Newton outperforms Newton.* arXiv. URL: https://arxiv.org/abs/2605.11316 [Accessed 10th July 2026].
5. Fan M., Tian Sh., Liu K., Zhao J., Li Y. Infrared small target detection based on region proposal and CNN classifier. *Signal, Image and Video Processing.* 2021;15(4):1927–1936. https://doi.org/10.1007/s11760-021-01936-z

6. Moran J., Qing H. Brain-inspired filtering Network for small infrared target detection. *Multimedia Tools and Applications*. 2023;82(18):28405–28426. https://doi.org/10.1007/s11042-023-14762-x
7. Wei L., Ma Zh., Yang Ch., Yao Q. Advances in the Neural Network Quantization: A Comprehensive Review. *Applied Sciences*. 2024;14(17):7445. https://doi.org/10.3390/app14177445
8. Gholami A., Kim S., Dong Zh., Yao Zh., Mahoney M.W., Keutzer K. *A Survey of Quantization Methods for Efficient Neural Network Inference.* arXiv. URL: https://arxiv.org/abs/2103.13630 [Accessed 10th July 2026].
9. Sali S.M., Meribout M., Majeed A.A. *Real Time FPGA Based CNNs for Detection, Classification, and Tracking in Autonomous Systems: State of the Art Designs and Optimizations.* arXiv. URL: https://arxiv.org/abs/2509.04153 [Accessed 10th July 2026].
10. Hozhabr S.H., Giorgi R. A Survey on Real-Time Object Detection on FPGAs. *IEEE Access*. 2025;13:38195–38238. https://doi.org/10.1109/ACCESS.2025.3544515

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Кирюшин Евгений Юрьевич, главный специалист, Уральский оптико-механический завод имени Э.С. Яламова, Екатеринбург, Российская Федерация.
e-mail: 6c3726@mail.ru

Evgeniy Y. Kiryushin, Main Specialist, Ural Optical and Mechanical Plant named after Mr. E.S. Yalamov, Ekaterinburg, the Russian Federation.

Кошелев Александр Викторович, аспирант, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация.
e-mail: 9122622678@gmail.com

Alexander V. Koshelev, Postgraduate, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, the Russian Federation.

Статья поступила в редакцию 03.08.2026; одобрена после рецензирования 19.10.2026; принята к публикации 11.11.2026.

The article was submitted 03.08.2026; approved after reviewing 19.10.2026; accepted for publication 11.11.2026.