

УДК 004.9

ОБНАРУЖЕНИЕ ОБЪЕКТОВ НА ПРОСТРАНСТВЕННО-НЕОДНОРОДНЫХ ФОНАХ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

© А. К. Шакенов

*Институт автоматизи и электрометрии СО РАН,
630090, г. Новосибирск, просп. Академика Коптюга, 1
E-mail: adil.shakenov@ngs.ru*

Рассмотрено несколько подходов к применению нейронных сетей для обнаружения объектов на пространственно-неоднородных фонах. Реализован метод построения классификатора для обнаружения объектов непосредственно по наблюдаемым фрагментам. Предложен подход, заключающийся в комбинации метода согласованной линейной фильтрации и свёрточных нейронных сетей. Показано, что применённый подход позволяет уменьшить вероятность ложной тревоги при сохранении вероятности обнаружения объекта.

Ключевые слова: обнаружение и распознавание объектов, свёрточные нейронные сети, машинное обучение, малоразмерные объекты.

DOI: 10.15372/AUT20190608

Введение. Задача обнаружения малоразмерных слабоконтрастных объектов активно исследуется в течение последних десятилетий. Из-за широкого разброса условий регистрации данных и сопутствующих помех [1] исследования в этой области продолжают оставаться актуальными, что подтверждается большим количеством работ, опубликованных в последние годы. При обнаружении малоразмерных слабоконтрастных объектов их форма и размер определяются аппаратной функцией системы и не содержат достаточной информации для надёжного обнаружения. Важной особенностью алгоритмов обнаружения таких объектов является необходимость оценить и исключить из рассмотрения подстилающий фон. Наиболее эффективным подходом для этого является пространственно-временная фильтрация серии последовательных изображений [2]. Однако в ряде случаев из-за особенностей геометрии съёмки или вычислительных ограничений системы обработки данных необходимо производить оценку и подавление фона по одному изображению. Известен подход к решению рассматриваемой задачи, позволяющий получить согласованный линейный фильтр для случая стационарного фона с известной ковариационной матрицей [3]. Активно разрабатываются и применяются различные алгоритмы оценки и фильтрации фона по наблюдаемой локальной окрестности, например билатеральная [4] и медианная [5] фильтрации, оптимальный линейный прогноз [6, 7], другие эвристические методы [8–11]. Метод согласованной линейной фильтрации разработан в предположении, что статистические свойства фона одинаковы по всему полю кадра. Это предположение может не выполняться для широкого круга наблюдаемых фонов, чем обусловлена актуальность поиска новых подходов к решению задачи обнаружения малоразмерных слабоконтрастных объектов на пространственно-неоднородных фонах.

В последнее время активно развиваются методы распознавания и обнаружения объектов с использованием обучаемых нейронных сетей [12]. Примеры использования такого подхода к обнаружению малоразмерных объектов можно найти в работах [13–16]. Вопросы применения рекуррентных нейронных сетей исследованы в [17, 18]. Нейросетевая классификация гиперспектральных изображений рассмотрена в [19].

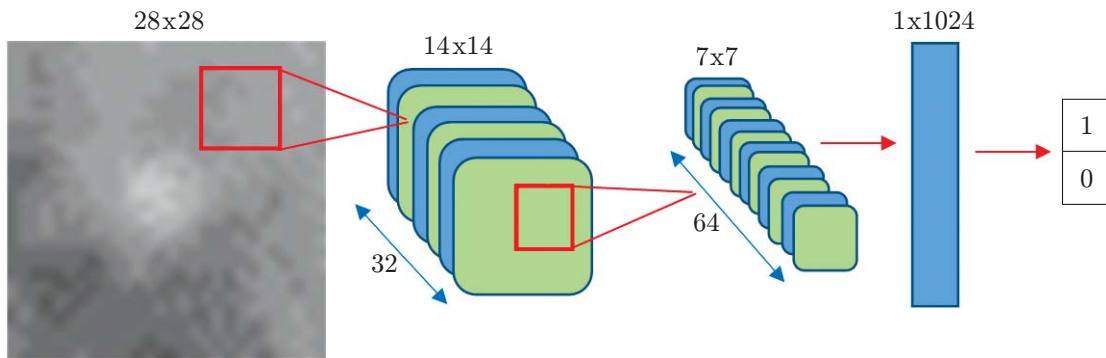


Рис. 1. Схема использованной нейронной сети

Поскольку нейронные сети используют достаточно большое количество промежуточных признаков в процессе обработки фрагмента, можно ожидать, что их применение даст возможность улучшить результаты линейной фильтрации именно на пространственно-неоднородных фонах. Кроме того, возможность обучать сеть непосредственно по наблюдаемым данным позволяет легко адаптировать этот подход к большому количеству фонов и наблюдаемых объектов.

Постановка задачи. Необходимо разработать алгоритм обнаружения объектов на неоднородных фонах, улучшающий характеристики обнаружения по сравнению с алгоритмом согласованной линейной фильтрации за счёт применения обучаемых нейронных сетей.

Обнаружение объектов с обучением классификатора по наблюдаемым фрагментам. Одним из способов применения нейронных сетей для обнаружения объектов является обучение классификатора, который характеризует каждый фрагмент наблюдаемого изображения как содержащий либо только объект, либо только фон. При этом размер обрабатываемого фрагмента выбирается равным размеру изображения объекта. Процедура обнаружения при таком подходе заключается в последовательном переборе всех фрагментов изображения и проверке их на наличие объекта с помощью обученного классификатора. Для обнаружения использовалась трёхслойная свёрточная нейронная сеть, схематически изображённая на рис. 1.

Первый слой сети выполняет свёртку с 32 различными фильтрами размерами 9×9 и уменьшение размеров полученных в результате свёртки изображений в 2 раза. Уменьшение осуществляется методом выбора наибольшего элемента из окрестности размерами 2×2 пиксела. Второй слой сети аналогичным образом выполнял свёртку с двумя фильтрами размерами 9×9 и уменьшение размеров выходных массивов в 2 раза. Третий слой переводит полученный массив данных в один вектор признаков, содержащий 1024 элемента. Полученный вектор признаков затем характеризуется как содержащий либо не содержащий объект.

Комбинация согласованной линейной фильтрации и нейронной сети. Детектирование объектов описанным выше способом трудоёмко с вычислительной точки зрения, поскольку каждый фрагмент изображения необходимо обрабатывать нейронной сетью, которая содержит каскад значительного числа фильтров. Метод согласованной линейной фильтрации даёт достаточно хорошие результаты для широкого круга реальных фонов, при этом, если ковариационная матрица фона оценена заранее, вычисление состоит в фильтрации одним линейным фильтром. Таким образом, возникает идея на первом этапе обработки использовать согласованный линейный фильтр, а затем применить обученную нейронную сеть. Зарегистрированное изображение можно представить в векторной форме

$$f_{src} = f_o + n,$$

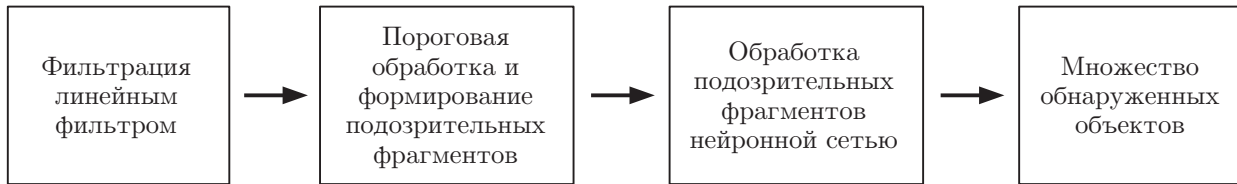


Рис. 2. Комбинация согласованной линейной фильтрации и нейронной сети

где f_0 — вектор объекта, n — вектор коррелированного шума (фона). Если K — матрица ковариации шума, то оптимальный в смысле увеличения отношения сигнал/шум линейный фильтр m имеет вид [3]

$$m = K^{-1} f_0.$$

На практике матрица K , как правило, неизвестна. В данной работе использовалась численная оценка матрицы K , полученная непосредственно по входным изображениям фона. Рассчитав таким образом линейный фильтр, дальнейшую обработку можно провести по схеме, приведённой на рис. 2.

Исходные данные и обучение сети. Для экспериментов использовались изображения Земли со спутника Электро Л-1 [20]. В работе рассмотрены точечные объекты, размеры и форма изображения которых определяются аппаратной функцией системы. Форма объекта моделировалась с использованием функции Гаусса, применялся аддитивный метод нанесения объекта. Таким образом, итоговая яркость фрагмента, содержащего объект, определялась следующим равенством:

$$D(i, j) = M(i, j) + Af(X - i, Y - j),$$

где $D(i, j)$ — итоговый фрагмент; $M(i, j)$ — исходный фрагмент; A — нормирующий множитель; $f(X - i, Y - j)$ — форма объекта с центром в точке (X, Y) , при этом функция имеет вид

$$f(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).$$

Для генерации обучающих данных центр объекта помещался в центр обрабатываемого фрагмента. Нормирующий множитель A выбирался таким образом, чтобы можно было достичь требуемого соотношения между среднеквадратичным отклонением (СКО) фона и максимальной интенсивностью объекта.

Для обучения сети распознаванию фрагментов после процедуры фильтрации данные получались следующим образом. На исходное изображение наносилось значительное количество объектов на расстоянии, в несколько раз превышающем их размер. Фрагменты изображения, содержащие объекты, сохранялись и использовались для обучения сети. Строился согласованный фильтр, и проводилась фильтрация изображения с нанесёнными объектами, а также исходного изображения без объектов. По обработанному изображению выбирался порог яркости, определяющий вероятность обнаружения объекта и ложной тревоги. На исходном изображении, не содержащем объекты, выбирались фрагменты, отклик на фильтр в которых превышал пороговое значение (ложно обнаруженные фрагменты). Эти фрагменты в дальнейшем применялись при обучении нейронной сети как примеры фона, не содержащего объекты. На рис. 3 приведены фрагменты с объектами и без них.

Экспериментальные результаты. Для сравнения эффективности порогового обнаружения при согласованной линейной фильтрации (СЛФ) и нейронной сети проводился следующий эксперимент. Одно случайно выбранное изображение фона использовалось для

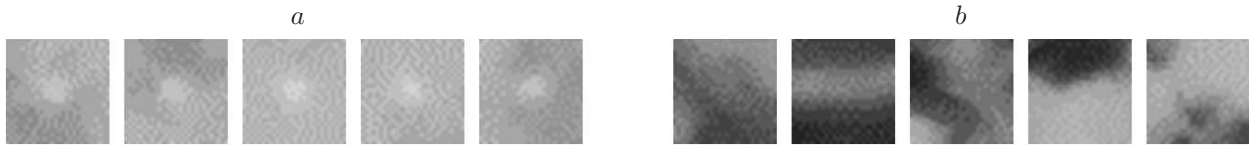


Рис. 3. Примеры исходных фрагментов: *a* — с объектами, *b* — без объектов

Таблица 1

Сравнение нейронной сети и согласованной линейной фильтрации

Текстура	Вероятность обнаружения	Ложная тревога	
		Сеть	СЛФ
Фон 1	0,9993	0,0057	0,0011
Фон 2	0,9998	0,0061	0,0007
Фон 3	0,9996	0,0085	0,0004
Фон 4	0,9997	0,0055	0,0005

обучения классификатора, дальнейшее сравнение алгоритмов осуществлялось на других фоновых изображениях. В качестве объекта взята гауссова функция с максимальной интенсивностью, равной одному СКО фона, и параметром σ , равным 3. Для получения тестового множества фрагментов на изображения фонов наносились 14000 объектов. Фрагменты с этими объектами использовались для оценки вероятности их обнаружения. Для оценки вероятности ложной тревоги из исходного изображения в произвольных местах вырезались 14000 фрагментов такого же размера. Полученные множества обрабатывались обученным классификатором. Результаты обработки нескольких фоновых изображений содержатся в табл. 1.

В столбце Сеть приведены значения ложной тревоги при использовании алгоритма обнаружения на основе обученного классификатора, а в столбце СЛФ — при использовании согласованной линейной фильтрации. Из таблицы видно, что в проведённых экспериментах нейронная сеть показала результаты хуже, чем согласованная линейная фильтрация. Скорее всего, это связано с тем, что при отсутствии чёткой методологии достаточно сложно подобрать обучающее множество для получения результата, близкого к оптимальному. Приведённые данные показывают, что оба алгоритма дают стабильные результаты при обработке различных входных изображений.

В табл. 2 приведены результаты работы алгоритма, комбинирующего согласованную линейную фильтрацию и нейронную сеть. В предыдущем эксперименте на вход нейронной сети подавались все фрагменты обрабатываемого изображения, в данном эксперименте — только фрагменты, подозрительные на наличие объекта по результатам линейной фильтрации. В данном случае использовалось изображение Фон 4 с 900 нанесёнными объектами. Эксперименты с другими фонами дали схожие результаты. На первом этапе использовался алгоритм согласованной линейной фильтрации, при этом подбирались значения порога, дающие вероятности обнаружения, указанные в таблице в столбце α_1 . Каждое значение порога определяет два множества: A_1 — множество правильно обнаруженных объектов и B_1 — множество ложно обнаруженных фрагментов, не содержащих объекты. Обозначим N_{obj} — количество всех нанесённых объектов, N_{pix} — количество всех пикселей изображения. Применяя множества A_1 и B_1 , оценим вероятности обнаружения и ложной тревоги при использовании алгоритма согласованной линейной фильтрации, приведённые в столбце СЛФ1. Вероятность обнаружения оценивалась как $\alpha_1 = |A_1|/N_{obj}$, вероятность ложной тревоги как $\beta_1 = |B_1|/N_{pix}$. Затем множества A_1 и B_1 подавались на вход нейронной сети. В результате получались два множества: A_2 — множество объектов, правильно классифицированных нейросетью, и B_2 — множество ложных фрагментов, неверно классифи-

Таблица 2

Комбинация нейронной сети и согласованной линейной фильтрации

СЛФ 1		Нейросеть		СЛФ + Сеть		СЛФ
α_1	β_1	α_N	β_N	α_C	β_C	β_0
0,999	$1,17 \cdot 10^{-3}$	0,968	0,097	0,967	$1,11 \cdot 10^{-4}$	$1,59 \cdot 10^{-4}$
0,98	$1,93 \cdot 10^{-4}$	0,969	0,404	0,95	$7,56 \cdot 10^{-5}$	$1,27 \cdot 10^{-4}$
0,96	$1,38 \cdot 10^{-4}$	0,971	0,480	0,932	$6,62 \cdot 10^{-5}$	$1,07 \cdot 10^{-4}$
0,94	$1,16 \cdot 10^{-4}$	0,976	0,5	0,918	$5,61 \cdot 10^{-5}$	$9,92 \cdot 10^{-5}$
0,92	$1,00 \cdot 10^{-4}$	0,979	0,545	0,901	$5,29 \cdot 10^{-5}$	$9,11 \cdot 10^{-5}$
0,90	$9,03 \cdot 10^{-5}$	0,980	0,564	0,882	$4,94 \cdot 10^{-5}$	$8,13 \cdot 10^{-5}$

цированных нейросетью как объекты. Эффективность работы нейросети при обработке множеств A_1 и B_1 показана в столбце Нейросеть. В нём указаны значения $\alpha_N = |A_2|/|A_1|$ и $\beta_N = |B_2|/|B_1|$. Итоговые вероятности обнаружения и ложной тревоги при рассмотренном методе обнаружения приведены в столбце СЛФ + Сеть. В нём указаны значения $\alpha_C = |A_2|/N_{obj}$ и $\beta_C = |B_2|/N_{pix}$. Для сравнения предложенного подхода были измерены вероятности ложной тревоги при применении алгоритма согласованной линейной фильтрации для вероятностей обнаружения, указанных в столбце α_C . Эти значения приведены в столбце СЛФ и обозначены как β_0 .

При сравнении значений β_C и β_0 видно, что предложенный подход позволил на 40–60 % уменьшить вероятность ложной тревоги при той же вероятности обнаружения. Характер изменений значений α_N и β_N показывает, что результаты обнаружения с помощью нейронной сети имеют корреляцию с результатами обнаружения алгоритмом согласованной линейной фильтрации. Для получения меньших значений α_1 необходимо использовать более высокий порог на этапе пороговой обработки, что даёт множества A_1 и B_1 , содержащие фрагменты с большей интенсивностью отклика на фильтр. Поскольку основным признаком наличия объекта является дополнительная зарегистрированная интенсивность, то можно предположить, что множество истинных объектов с большей интенсивностью A_1 становится проще для правильного распознавания, а множество ложных фрагментов с большей интенсивностью B_1 — сложнее. Этим можно объяснить характер изменения величин α_N и β_N . Уменьшение β_C по отношению к β_0 объясняется, скорее всего, тем, что нейросеть использует дополнительные к величине отклика на фильтр признаки, применение которых позволяет улучшить итоговые результаты.

Заключение. В проведённых экспериментах прямое применение нейронной сети для классификации фрагментов в рассмотренном диапазоне вероятностей обнаружений не позволило улучшить результаты, полученные с использованием согласованной линейной фильтрации. При этом показана возможность эффективно использовать комбинацию согласованной линейной фильтрации и нейронной сети. В результате применения предложенного подхода была повышена эффективность детектирования объектов, вероятность ложной тревоги снижена на 40–60 % при той же вероятности обнаружения объекта. Дальнейшие исследования могут быть направлены на более точную настройку параметров сети и на использование больших объёмов данных в процессе обучения.

Финансирование. Работа выполнена при поддержке Министерства образования и науки РФ (проект № АААА-А17-117052410034-6).

СПИСОК ЛИТЕРАТУРЫ

1. Cheng G., Han J. A survey on object detection in optical remote sensing images // ISPRS Journ. Photogramm. Remote Sens. 2016. **117**. P. 11–28.

2. Киричук В. С., Косых И. П., Попов С. А., Синельщиков В. В. Подавление квазистационарного фона в последовательности изображений посредством межкадровой обработки // Автометрия. 2014. **50**, № 2. С. 3–13.
3. Pratt W. K. Digital Image Processing: PIKS Scientific Inside. New Jersey: John Willey and Sons, 2007. 662 p.
4. Bae T.-W., Sohng K.-I. Small target detection using bilateral filter based on edge component // Journ. Infrar. Milli Terahz. Waves. 2010. **31**, N 6. P. 735–743.
5. Deshpande S. D., Er M. H., Ronda V., Chan Ph. Max-mean and max-median filters for detection of small-targets // Proc. SPIE. 1999. **3809**. P. 74–83.
6. Soni T., Zeidler R., Ku W. H. Performance evaluation of 2D adaptive prediction filters for detection of small object in image data // IEEE Trans. Image Process. 1993. **2**, N 3. P. 327–340.
7. Ffrench P. A., Zeidler J. R., Ku W. H. Enhanced detectability of small objects in correlated clutter using an improved 2-D adaptive lattice algorithm // IEEE Trans. Image Process. 1997. **6**, N 3. P. 383–397.
8. Hong P., Wang C., Zhang Z. Weak point target detection in the complicated infrared background // Proc. SPIE. 2011. **8200**. 820007.
9. Dong Y.-X., Li Y., Zhang H.-B. Research on infrared dim-point target detection and tracking under sea-sky-line complex background // Proc. SPIE. 2011. **8193**. 81932J.
10. Иванов В. А., Киричук В. С., Косых В. П., Синельщиков В. В. Особенности обнаружения точечных объектов в изображениях, формируемых матричным приёмником // Автометрия. 2016. **52**, № 2. С. 10–19.
11. Шакинов А. К. Алгоритмы подавления фона в задаче обнаружения точечных объектов по изображениям // Автометрия. 2014. **50**, № 4. С. 81–87.
12. Liua W., Wanga Z., Liua X. et al. A survey of deep neural network architectures and their applications // Neurocomputing. 2017. **234**. P. 11–26.
13. Zhao Sh., Song Y., Zhao Y. et al. Infrared dim small target segmentation method based on ALI-PCNN model // Proc. SPIE. 2017. **10459**. 104590A.
14. Ryu J., Kim S. Small infrared target detection by data-driven proposal and deep learning-based classification // Proc. SPIE. 2018. **10624**. 106241J.
15. Fan Z., Bi D., Xiong L. et al. Dim infrared image enhancement based on convolutional neural network // Neurocomputing. 2018. **272**. P. 396–404.
16. Zhang P., Li J. Neural-network-based single-frame detection of dim spot target in infrared images // Opt. Eng. 2007. **46**. P. 076401.
17. Козик В. И., Нежевенко Е. С., Феоктистов А. С. Адаптивное прогнозирование развития лесных пожаров на основе рекуррентных нейронных сетей // Автометрия. 2013. **49**, № 3. С. 44–55.
18. Козик В. И., Нежевенко Е. С., Феоктистов А. С. Исследование метода адаптивного прогнозирования развития лесных пожаров на основе рекуррентных нейронных сетей // Автометрия. 2014. **50**, № 4. С. 88–95.
19. Нежевенко Е. С., Феоктистов А. С., Дашевский О. Ю. Нейросетевая классификация гиперспектральных изображений на основе преобразования Гильберта — Хуанга // Автометрия. 2017. **53**, № 2. С. 79–85.
20. Электро-Л. Планета Земля за последние сутки. URL: <http://electro.ntsomz.ru> (дата обращения: 23.04.2019).

Поступила в редакцию 23.04.2019

После доработки 07.05.2019

Принята к публикации 29.05.2019