

УДК 004.931

DOI: 10.15587/1729-4061.2019.166887

Застосування нейронних мереж Кохонена для пошуку зон інтересу при виявленні і розпізнаванні об'єктів

В. А. Скуратов, К. О. Кузьмін, І. В. Нелін, М. К. Седанкін

Одним з найбільш ефективних способів підвищення точності і швидкодії алгоритмів розпізнавання є попереднє виділення зон інтересу на аналізованих зображеннях. Досліджено можливість застосування самоорганізуючих карт і нейронної мережі Кохонена для визначення зон інтересу на радіолокаційному або супутниковому зображенні підстильної поверхні. У знайдених зонах інтересу велика ймовірність виявлення об'єкта, що цікавить, для подальшого аналізу. Визначення зон інтересу необхідно перш за все для автоматизації та прискорення процесу пошуку і розпізнавання об'єктів, що цікавлять. Це, в силу постійно наростаючої кількості супутників, стає все більш доцільним. Представлено процес моделювання, аналіз і порівняння результатів застосування даних методів для визначення зон інтересу при розпізнаванні образів літальних апаратів на тлі підстильної поверхні. Також описано процес попередньої обробки вхідних даних. Освячено загальний підхід до побудови та навчання самоорганізуючої карти і нейронної мережі Кохонена. Застосування карт і нейронної мережі Кохонена дозволяє в 15–100 разів зменшити обсяг даних, що аналізуються. Це, відповідно, прискорює процес виявлення і розпізнавання об'єкта, що цікавить. Використання наведеного алгоритму істотно скорочує необхідну кількість навчальних образів для згортальної мережі, здійснює остаточне розпізнавання. Зменшення навчальної вибірки обумовлено тим, що розмір частин, що подаються на згортальну мережу вхідного зображення, прив'язаний до масштабу зображення і дорівнює розміру найбільшого об'єкта детектування. Нейронна мережа Кохонена показала себе більш ефективною відносно до даної задачі, так як рідше розміщує центри кластерів на підстильній поверхні в силу незалежності ваги нейронів від сусідніх центрів. Дані технічні рішення можуть застосовуватися при аналізі візуальних даних із супутників, літальних апаратів і безпілотних автомобілів, в медицині, робототехніці і т.д.

Ключові слова: розпізнавання образів, самоорганізуючі карти, нейронна мережа Кохонена, радіолокаційні та супутникові зображення, робототехніка

1. Введение

В настоящее время невысокая скорость распознавания изображений является одной из главных проблем, тормозящих развитие современных систем обработки визуальных данных. Увеличение скорости распознавания изображений приведет к росту производительности систем анализа спутниковых и радиолокационных снимков, медицинских изображений, данных с робототехнических

комплексов и беспилотных автомобилей и т.д., оказав как технологический, так и экономический эффекты.

Одним из способов повышения точности и быстродействия алгоритмов распознавания является использование нейронных сетей (НС). Самым современным типом НС, используемым при распознавании образов, является сверточная НС (convolutional neural network, CNN), которая получила свое название из-за наличия операции свертки [1]. Для качественной работы данного типа сетей необходимо, чтобы размер распознаваемого объекта был сопоставим с размерами объектов в обучающей выборке. Очевидно, что создание обучающей выборки, в которой, помимо различных типов искомым объектов, имеется еще и вариация по масштабу объекта на изображении, представляется крайне сложной задачей. Обучение такой сети будет занимать существенное время и вычислительные ресурсы, возрастет количество ошибок в распознавании, а скорость работы алгоритма будет невысока. Представленный в данной работе алгоритм позволяет уменьшить количество анализируемых сверточной НС данных при детектировании объектов на изображении в 15–20 раз и соответственно сократить время поиска объектов, увеличив точность распознавания и сэкономив вычислительные ресурсы. Это достигается за счет применения сети или карты Кохонена для определения зон интереса на входном изображении, в которых велика вероятность обнаружения искомого объекта. В данной работе рассматривается применение разработанного алгоритма для обнаружения и распознавания типа летательного аппарата (ЛА) на спутниковых снимках.

2. Анализ литературных данных и постановка проблемы

На сегодняшний день существует целый ряд работ [2–5], где исследовано применение НС, сети (Kohonen neural network) и самоорганизующейся карты Кохонена (Self-organizing map, SOM) для распознавания изображений. Одни авторы рассматривают сети и карты Кохонена для подготовки входных данных с последующим анализом иными НС, а другие – для непосредственного распознавания образов. В работе [2] приведены результаты моделирования НС для распознавания изображений, в которой сначала происходит сегментация изображения картами Кохонена, а затем анализ в гибридной НС. Карты Кохонена применялись для уменьшения обучающей выборки для гибридной НС и для ускорения анализа. В [3] представлена сеть Кохонена для распознавания цветных изображений. Существенным недостатком [2, 3] является то, что изображение для распознавания представляет собой весь образ для классификации и при изменении масштаба искомого фрагмента на входном изображении, распознавание не будет выполнено. В работе [4] показан алгоритм поиска фрагмента в серии цветных изображений с помощью сети Кохонена. В сеть для анализа последовательно подаются фрагменты изображения, полученные «окном», пиксель за пикселем проходящим по анализируемому изображению. Размер «окна» соответствует размеру искомого фрагмента. Поиск фрагмента, размером 211×169 пикселей, в изображении, размером 621×497 пикселей, был выполнен за 122 сек. За это время было проанализировано 135219 гистограмм. Временные затраты на анализ одного изображения не позволяют использовать предложен-

ный алгоритм для распознавания большого объема данных. Также для анализа одного изображения требуются серьезные вычислительные ресурсы, т.к. оно разбивается на 135219 фрагментов для анализа. В [5] произведена оценка эффективности применения для классификации изображений гибридной сети, в которой сначала происходит анализ сетью Кохонена, а потом данные подаются на перцептронную сеть. Показано, что применение конфигурации сети данной конфигурации позволяет уменьшить вычислительную сложность используемых алгоритмов. Это достигается за счет уменьшения необходимого количества нейронов более чем в 2 раза, что позволяет также уменьшить погрешность классификации в 4–5 раз.

Особый интерес представляет использование сети и карты Кохонена для выделения различных характеристик объектов на изображениях. В [6] карты Кохонена были использованы для решения сложных задач кластеризации с большим количеством объектов и выделения среди них объектов с несвойственными характеристиками. В работе [7] сеть Кохонена используется для поиска центров кластеров путем нахождения минимального эвклидова расстояния между точками, входящими в кластер.

Самым современным типом НС, используемым при распознавании образов, является сверточная НС, которая получила свое название из-за наличия операции свертки. Суть этой операции состоит в том, что каждый фрагмент изображения поэлементно умножается на ядро свертки, а результат суммируется и записывается в аналогичную позицию выходного изображения. После чего результат подается на слой субдискретизации и, в простейшем случае, полностью связную нейронную сеть. В работе [8] приведена структура сверточной НС и ее применение для распознавания образов, а в [9] для анализа видеопоследовательности.

В статьях [10–14] описано развитие алгоритма R-CNN (Region-based Convolutional Network), предназначенного для детектирования объектов на изображениях. В [10, 11] описан первоначальный вариант алгоритма R-CNN. Алгоритм выделяет на входном изображении порядка 2000 регионов, каждый из которых масштабируется, используя аффинное преобразование, и подается на вход сверточной сети, которая извлекает вектор (карту) особенностей. В статье [11] дополнительно описывается тренировка линейной регрессии для уточнения координат окна объекта. Такое уточнение локализации объекта позволяет повысить качество на 3–4 %. R-CNN имеет несколько недостатков, в основном связанных с высокими временными затратами на обучение сети, а также на непосредственную обработку изображений алгоритмом. Требуется натренировать сверточную нейронную сеть в два этапа и натренировать линейные регрессоры для каждого класса, а обработка одного изображения занимает порядка 47 секунд. Статья [12] Fast R-CNN является продолжением [10, 11] и в ней авторы предлагают подавать на вход НС полное изображение. В отличие от R-CNN, где карта особенностей генерировалась для каждого выделенного региона изображения, здесь генерируется карта особенностей для всего изображения, а затем определяются карты для каждого региона, что позволяет существенно сократить время детектирования. Следующая статья [13] Faster R-CNN предлагает

заменить процедуру генерации интересующих регионов на отдельную сверточную НС, которая извлекает признаки из изображения и определяет предполагаемые границы интересующих регионов. Карта признаков входного изображения обходится скользящим окном и для каждой позиции окна извлекается вектор признаков малой размерности. Полученные вектора подаются на вход двум слоям НС. Один из этих слоев используется для уточнения границ региона, а другой – для классификации объекта, расположенного внутри этого региона. Для каждой позиции скользящего окна одновременно могут рассматриваться несколько вариантов границ регионов, которые обладают разным размером или разным соотношением сторон. Mask R-CNN [14] развивает архитектуру Faster R-CNN добавляя возможность предсказать положение маски, покрывающей найденный объект.

В статье [15] описан алгоритм YOLO (You Only Look Once), который позволяет производить детектирование и распознавание объектов на изображениях в 10^3 раз быстрее, чем R-CNN и в 10^2 раз быстрее, чем Fast R-CNN, но с более низкой точностью. Этот алгоритм накладывает на входное изображение сетку и разделяет его на ячейки. Вокруг каждой ячейки алгоритм определяет ограничивающие рамки зоны возможного нахождения объектов с оценкой точности обнаружения и вероятностью принадлежности к классам. Затем оценка точности для каждой зоны умножается на вероятность класса и получается окончательное значение вероятности обнаружения. Алгоритм анализирует от нескольких тыс. до нескольких десятков тыс. частей изображения с ограничительными рамками разных размеров.

В работе [16] представлен алгоритм SSD: Single Shot MultiBox Detector, который по точности и быстродействию сопоставим с YOLO. Алгоритм осуществляет перекрытие всей площади входного изображения ограничивающими рамками, размер которых изменяется в установленных пределах, позволяя детектировать и распознавать объекты различных размеров. Для каждой части изображения, выделенной рамкой, выполняется оценка вероятности принадлежности к классам и производится коррекция размеров рамок. Алгоритм выполняет анализ от нескольких тыс. до нескольких десятков тыс. частей изображения.

При детектировании и распознавании объектов на радиолокационном или спутниковом изображениях необходимо применять алгоритмы, разделяющие входное изображение на совокупность изображений, подходящих по размеру для анализа в сверточной НС. Для экономии времени и ресурсов применяемый алгоритм должен уменьшать количество данных, необходимых для дальнейшего анализа. Такой алгоритм можно реализовать следующими способами:

- 1) последовательное разбиение входного изображения на кадры необходимого для сверточной НС размера со сдвигом кадра на некоторое количество пикселей относительно предыдущего;
- 2) использование алгоритмов R-CNN, Fast и Faster R-CNN, Mask R-CNN, YOLO, а также SSD.

Оба представленных способа требуют серьезных затрат времени и вычислительных ресурсов. В одном случае на анализ в последующую сверточную НС

подаются несколько тысяч кадров, на которые разбивается каждое входное изображение [4, 10, 11]. В другом – изображение анализируется целиком с перебором возможных границ регионов интереса, без предварительного отсеивания областей, не представляющих интерес [12–16].

Таким образом, существует потребность в алгоритмах, способных уменьшить количество анализируемых данных, и, тем самым, повысить быстродействие систем распознавания, основанных на применении сверточных сетей.

3. Цель и задачи исследования

Целью исследования является создание алгоритма определения зон интереса при распознавании типа ЛА сверточной нейронной сетью на радиолокационном и спутниковом изображениях подстилающей поверхности.

Для достижения цели были поставлены следующие задачи:

- выделить границы объектов, присутствующих на изображении;
- выполнить поиск центров объектов;
- проверить выбранные алгоритмы на реальных входных данных, определить с их помощью зоны интереса на изображениях и сравнить результаты применения самоорганизующейся карты и нейронной сети Кохонена при поиске центров объектов.

4. Алгоритм выделения границ объектов, расположенных на подстилающей поверхности

Перед тем, как передать входные данные для анализа в сверточную НС, необходимо провести их предварительную обработку.

Между корпусом ЛА и подстилающей поверхностью обязательно имеется граница раздела, которая образуется из-за различия яркости или уровней освещенности, т. к. отражательная способность ЛА, как правило, выше, чем у подстилающей поверхности. Также в большинстве случаев корпус ЛА отбрасывает тень на подстилающую поверхность, что также образует границу раздела. Если выделить данные границы вокруг ЛА точками на входном изображении, то центр данного кластера фактически совпадёт с центром ЛА.

Алгоритм выделения границ объектов, расположенных на подстилающей поверхности, выглядит следующим образом:

- 1) входное цветное изображение конвертируется в оттенки серого;
- 2) применение оператора Собеля – дифференциальный оператор, вычисляющий приближенное значение градиента яркости изображения;
- 3) конвертация в бинарное изображение, используя отсечение по порогу яркости. Полученные границы раздела яркостей принимают значения 1, все остальные – 0;
- 4) удаление объектов малого размера.

В табл. 1 представлено сравнение быстродействия алгоритма выделения границ объектов в зависимости от разрешения входного изображения и количества возможных объектов минимального размера. Вычисления выполнены в программе собственной разработки без применения GPU (графический процес-

сор), алгоритмов ускорения и оптимизации на персональном компьютере со следующими параметрами: процессор Intel Core i7 – 6700К, 4 ГГц, ОЗУ 32 Гб.

Таблица 1

Сравнение времени работы алгоритма выделения границ объектов

Размер входного изображения, пикс.	112×218	202×387	767×778
Количество объектов поиска, шт.	3	10	64
Предварительная обработка, сек.	0,002	0,003	0,02

Как видно из табл. 1, представленный алгоритм обладает достаточно высоким быстродействием, выделяя границы объектов входного изображения за 0,002–0,02 секунды.

5. Методика поиска центров возможных зон интереса на изображениях

Для сокращения времени анализа и необходимых вычислительных ресурсов при обнаружении на входном изображении интересующих объектов, предлагается выделить зоны их возможного нахождения с дальнейшим последовательным анализом найденных зон в сверточной НС. Это позволяет сэкономить время и ресурсы за счет уменьшения количества анализируемых кадров и, следовательно, повысить скорость работы системы в целом.

После первичной обработки входного изображения для поиска центров возможных зон интереса предлагается использовать самоорганизующиеся карты или нейронную сеть Кохонена.

Нейронные сети Кохонена – класс самоорганизующихся НС, в основе которых лежит слой Кохонена, состоящий из некоторого количества параллельно действующих адаптивных линейных элементов. После прохождения линейных элементов сигналы обрабатываются по принципу «Победитель получает всё», где наибольшему сигналу присваивается единичное значение, а остальные обнуляются.

Для решения задачи выделения центров зон интереса на входном изображении необходимо обучение НС без учителя (отсутствие обучающей выборки), т.к. обработка каждого входного изображения должна осуществляться независимо от предыдущих. Основными разновидностями сетей Кохонена, используемыми обучение без учителя, являются [17, 18]:

- сети Кохонена для векторного квантования сигналов;
- самоорганизующиеся карты Кохонена.

5. 1. Сеть Кохонена для векторного квантования сигналов

Слой Кохонена состоит из некоторого количества n параллельно действующих линейных элементов, имеющих одинаковое число входов и получают на свои входы один и тот же вектор входных сигналов $\mathbf{x}=[x_1, x_2, \dots, x_i]$. На выходе j -го линейного элемента получаем сигнал

$$y_j = \omega_{j0} + \sum_{i=1}^m w_{ji} x_i, \quad (1)$$

где w_{ji} – синаптический весовой коэффициент i -го входа j -го нейрона, i – номер входа, j – номер нейрона, ω_{j0} – пороговый коэффициент.

После прохождения слоя линейных элементов сигналы обрабатываются по правилу “Победитель получает всё”. Среди выходных сигналов выполняется поиск максимального y_j : его номер $i_{\max} = \text{argmax}(y_j)$. На выходе сигнал с номером i_{\max} равен единице, остальные – нулю. В случае, если максимум достигается одновременно на выходах нескольких сумматоров, то единичное значение присваивается одному из них по установленному правилу, например, первому. Обучение сети Кохонена представляет собой процесс подбора значений весов, минимизирующих ошибки от замены близких в смысле используемой метрики входных векторов вектором весов. Такой подход называется векторным квантованием [19] и применяется, например, в задачах сжатия аудио- и видеосигналов. Сама идея векторного квантования состоит в компактном представлении многомерных входных векторов с помощью ограниченного набора опорных векторов меньшей размерности, образующих кодовую таблицу. В сети Кохонена входные векторы кодируются номерами нейронов-победителей (номерами кластеров). Таким образом, все векторы из некоторой области входного пространства заменяются одним опорным вектором, являющимся их ближайшим соседом. Мера близости между объектами называется расстоянием. В сетях Кохонена обычно применяется евклидово расстояние:

$$d_E(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} = \|x - y\|. \quad (2)$$

Евклидово расстояние между векторами x и y представляет собой евклидову норму разности векторов, или длину отрезка, соединяющего точки x и y . Для обучения сети Кохонена “без учителя” применяются механизмы конкуренции. При подаче на вход вектора x побеждает тот нейрон, вектор весов которого в меньше всего отличается от входного вектора. Перед процессом обучения производится инициализация сети – первоначальное задание векторов весов. В базовом случае задаются случайные значения весов. Процесс обучения сети Кохонена состоит из циклического повторения следующих шагов:

1. подача исходных данных на входы сети.
2. Нахождение значений выхода каждого нейрона.
3. Определение нейрона-победителя, веса которого в наименьшей степени отличаются от соответствующих компонентов входного вектора.
4. Корректировка весов нейрона-победителя по правилу Кохонена:

$$w_j^{(k+1)} = w_j^k + \eta_j^k (x - w_j^k), \quad (3)$$

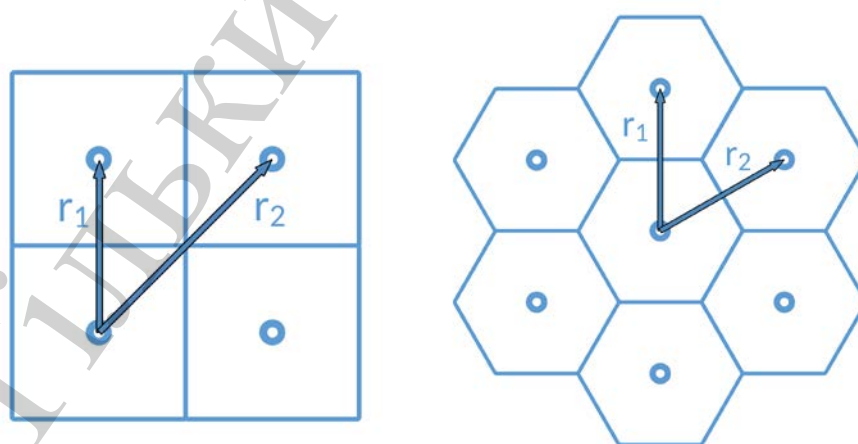
где \mathbf{x} – входной вектор, k – номер цикла обучения, η_j^k – коэффициент скорости обучения j -го нейрона в k -ом цикле обучения.

5. Переход на шаг 1, если обучение не завершено.

Следовательно, тот нейрон, чей вектор весов был ближе к входному вектору, обновляется, чтобы стать еще ближе. В итоге данный нейрон, скорее всего, выиграет конкуренцию при подаче на вход близкого вектора и проиграет при подаче существенно отличающегося. После многократной подачи на вход сети обучающих векторов будет иметься нейрон, который выдает 1, когда вектор принадлежит кластеру, и 0, когда вектор ему не принадлежит. Таким образом сеть Кохонена учится классифицировать входные векторы.

5. 2. Самоорганизующиеся карты Кохонена

Основной целью карт самоорганизации является преобразование поступающих на вход векторов, имеющих произвольную размерность, в одно- или двухмерную дискретную карту с топологически упорядоченной формой [18]. Карты, как и сети Кохонена, основаны на конкурентном обучении. Нейроны выходного слоя соревнуются за право активации, в результате чего активным оказывается только один выходной нейрон – нейрон-победитель. Одним из способов организации такого вида конкуренции между нейронами является использование отрицательных обратных связей между ними. В общем случае нейроны в SOM располагаются в узлах двумерной сетки с прямоугольными или шестиугольными ячейками (рис. 1) и взаимодействуют друг с другом. Величина этого взаимодействия определяется расстоянием r_n между нейронами на карте. Для шестиугольной сетки расстояние между отдельными нейронами больше совпадает с евклидовым расстоянием, чем для четырехугольной. Количество нейронов в сетке определяет степень детализации результата работы сети и от этого зависит точность обобщающей способности карты.



а б

Рис. 1. Начальное расположение нейронов в узлах двумерной сетки с ячейками: а – четырехугольными; б – шестиугольными.

Положения определенных в ходе конкурентного процесса нейронов (т. е. нейронов-победителей) упорядочиваются по отношению друг к другу так, что на решетке задается значимая система координат. Таким образом, SOM формируют топографические карты входных образов, в которых пространственное местоположение (координаты) нейронов решетки является индикатором статистических признаков, содержащихся в образах. Пространственное положение выходных нейронов в топографической карте соответствует конкретной области признаков данных, выделенных из входного пространства. Отсюда происходит и само название "самоорганизующиеся карты".

Карта признаков обладает следующими свойствами:

1. Аппроксимация входного пространства. Карта признаков, представленная множеством векторов синаптических весов w_j , в выходном пространстве выполняет аппроксимацию входного пространства. Главной целью алгоритма SOM является описание большого объема векторов входного пространства с помощью определения небольшого набора прототипов в выходном пространстве, которые представляют собой адекватную аппроксимацию исходного входного пространства.

2. Топологический порядок. Карта признаков, полученная после применения алгоритма SOM, является топологически упорядоченной, так как пространственное положение нейронов в решетке соответствует конкретной области или признаку входного образа.

3. Соответствие плотности. Карта признаков отражает изменения в статистиках распределения входного сигнала. Области входного образа, из которых берутся векторы с большей вероятностью, отображаются в гораздо большие области выходного пространства и с большим разрешением, чем области, из которых берутся векторы с меньшей вероятностью.

4. Выбор признаков. Для данных из входного пространства с нелинейным распределением SOM способна извлечь набор наилучших признаков для аппроксимации исследуемого распределения.

Каждый из входных образов обычно состоит из локализованной области активности, размещаемой в относительно спокойной области. Расположение и природа такой области варьируется в зависимости от вида входного примера. Работа алгоритма SOM обычно начинается с инициализации синаптических весов сети. После корректной инициализации сети для формирования карты запускаются три основных процесса: конкуренция, кооперация и синаптическая адаптация.

1. Конкуренция. Для каждого входного образа нейроны сети вычисляют относительные значения дискриминантной функции, которая является базовой для процесса конкуренции среди нейронов.

Для того чтобы подобрать наилучший вектор w_j , соответствующий входному вектору x , нужно сравнить скалярные произведения $w_j^T x$ для каждого нейрона и выбрать наибольшее значение. При этом у каждого нейрона имеется некоторое значение насыщения, равное порогу, взятому с обратным знаком. Координаты центра топологической окрестности возбужденного нейрона определяются выбором нейрона с наибольшим скалярным произведением.

Наилучший критерий соответствия, основанный на максимизации скалярного произведения $w_j^T \mathbf{x}$, математически эквивалентен минимизации Евклидова расстояния между векторами \mathbf{x} и w_j . Если использовать индекс $p(\mathbf{x})$ для идентификации того нейрона, который лучше всего соответствует входному сигналу \mathbf{x} , то эту величину можно определить с помощью следующего соотношения:

$$p(\mathbf{x}) = \arg \min_j \|\mathbf{x} - w_j\|, \quad (4)$$

где $\|\cdot\|$ – евклидова норма.

Конкретный нейрон, удовлетворяющий данному условию, называется победившим для входного вектора \mathbf{x} .

2. Кооперация. Победивший нейрон определяет расположение топологической окрестности нейронов сети, чем обеспечивает кооперацию между этими нейронами. Нейрон-победитель находится в центре топологической окрестности сотрудничающих друг с другом нейронов. Данное условие приводит к определению топологической окрестности победившего нейрона i , которая плавно уменьшается с увеличением расстояния. Обозначим символом h_{ji} топологическую окрестность с центром в победившем нейроне, состоящую из множества возбуждаемых (кооперирующихся) нейронов. Примером h_{ji} , удовлетворяющим этим условиям, является функция Гаусса:

$$h_{ji(x)} = \exp\left(-\frac{d_{ji}^2}{2\sigma^2}\right), \quad (5)$$

где $d_{ji}^2 = \|r_j - r_i\|^2$ – расстояние на плоскости между j -м нейроном и текущим нейроном-победителем, σ – эффективная ширина: числовой параметр, задающий размер окрестности вокруг нейрона-победителя, в которой производится коррекция весов. Чем меньше σ , тем меньше нейронов из окрестности победителя, на которые влияет входной вектор \mathbf{x} .

Еще одним свойством алгоритма SOM является то, что размер топологической окрестности с течением времени уменьшается. Данное требование удовлетворяется за счет постепенного уменьшения эффективной ширины σ функции топологической окрестности h_{ji} . Популярным вариантом зависимости σ от дискретных значений времени n является экспоненциальное убывание, описываемое следующей формулой:

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_1}\right), \quad (6)$$

где σ_0 – начальное значение величины σ в алгоритме SOM, τ_1 – параметр, отвечающий за скорость убывания функции $\sigma(n)$.

3. Синаптическая адаптация. Данный механизм позволяет возбужденным нейронам увеличивать собственные значения дискриминантных функций по отношению к входным векторам посредством корректировок синаптических весов. Соответствующие корректировки производятся таким образом, чтобы отклик нейрона-победителя на последующее поступление аналогичных примеров усиливался. Для дискретного времени вектора синаптических весов $w_j(n)$ в момент времени n , обновленный вектор $w_j(n+1)$ в момент времени $n+1$ можно определить в следующем образом:

$$w_j(n+1) = w_j(n) + \eta(n) h_{ji(x)}(n) (x - w_j(n)), \quad (7)$$

где $\eta(n)$ – параметр скорости обучения; $h_{ji(x)}(n)$ – функция окрестности с центром в победившем нейроне $i(x)$. Оба этих параметра динамически изменяются во время обучения с целью получения наилучшего результата. Данное выражение применяется ко всем нейронам сети, которые лежат в топологической окрестности победившего нейрона. Выражение (7) имеет эффект перемещения вектора синаптических весов победившего нейрона в сторону входного вектора. Следовательно, благодаря коррекции в окрестности победившего нейрона, векторы синаптических весов будут стремиться следовать распределению входных векторов.

Исходя из п. 5. 1 и п. 5. 2 можно сделать вывод, что нейронная сеть Кохонена и SOM могут быть использованы для поиска центров областей интереса, так как способны определять центры кластеров входных данных. При сравнении результатов поиска центров кластеров описанным методом с алгоритмами нечетких центров и вычитающей кластеризации, нейронные сети оказались существенно более быстродействующими.

6. Описание алгоритма поиска зон интереса

На вход алгоритма для анализа подается цветное изображение. Далее производится его предварительная обработка, которая осуществляется в несколько этапов:

1. Выделение границ объектов, присутствующих на изображении. Для этого использован алгоритм, основанный на применении оператора Собеля, описанный в п. 4.

2. Поиск центров объектов. Для этого используются нейронная сеть или самоорганизующаяся карта Кохонена.

3. Определение зон интереса. Вокруг найденных центров объектов формируется расширенная зона интереса, в пределах которой производится выделение частей исходного изображения определенным «окном». Выделенные части входного изображения должны содержать в себе точку центра кластера и перекрывать окрестность зоны интереса. Размеры «окна» выбираются исходя из заданных габаритов самого большого детектируемого объекта и изменяются при изменении масштаба изображения.

4. Анализ в сверточной сети. Полученные из зон интереса части входного изображения подаются на анализ в сверточную сеть, где происходит определение наличия объекта и его тип.

Схема работы алгоритма поиска зон интереса представлена на рис. 2.

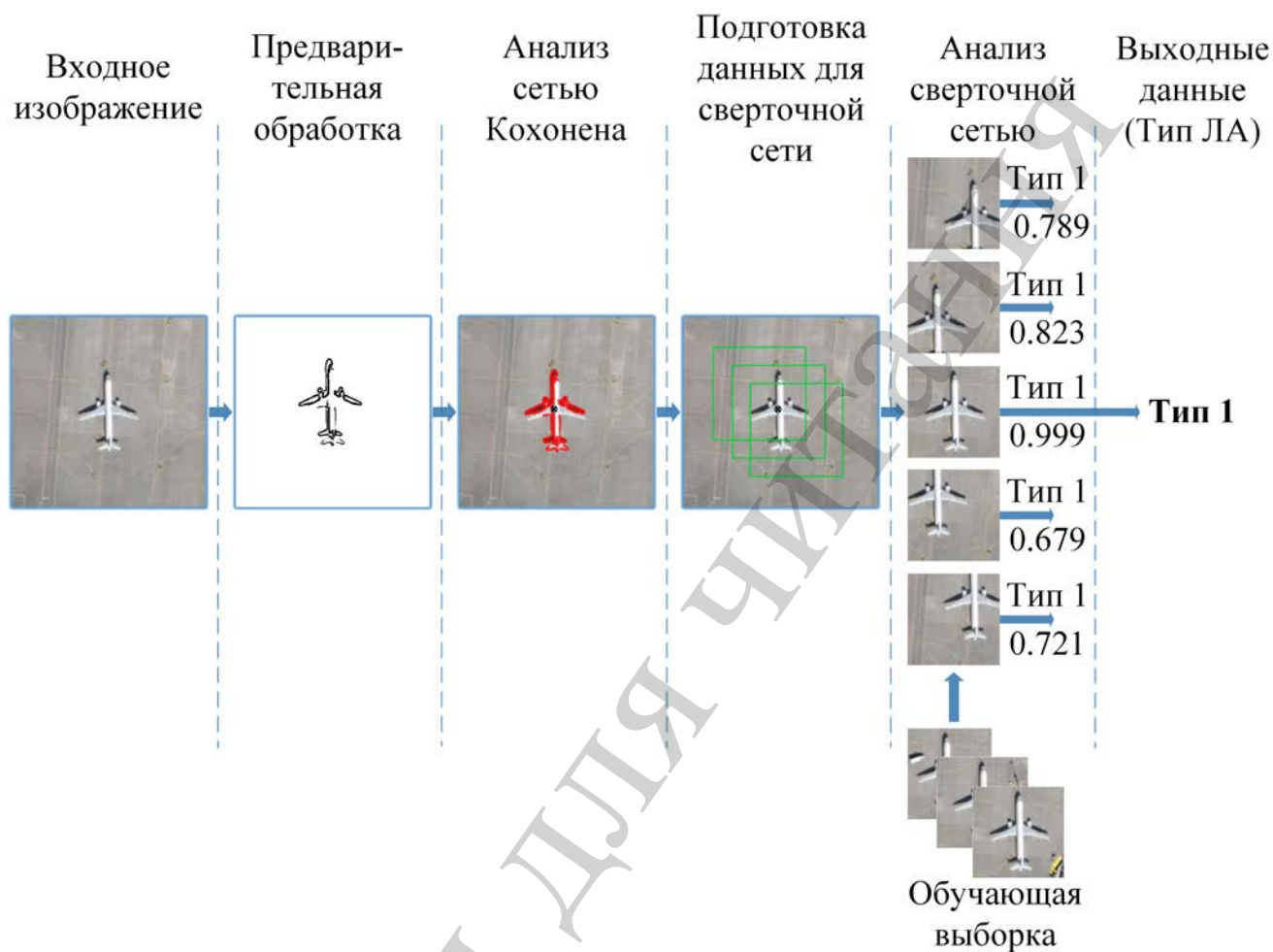


Рис. 2. Схема работы алгоритма поиска зон интереса и определения типа ЛА

Данный алгоритм может применяться для поиска и распознавания на входном изображении не только различных типов ЛА, но и других объектов, представляющих интерес, например при анализе визуальных данных беспилотных автомобилей или при испытаниях робототехнических комплексов. Для этого необходимо задать максимальные и минимальные возможные размеры объектов для работы алгоритма поиска зон интереса, составить обучающую выборку этих объектов и обучить сверточную НС.

7. Демонстрация работы алгоритма на реальных входных данных и сравнение результатов применения самоорганизующейся карты и нейронной сети Кохонена при поиске центров объектов.

Для детектирования на входном изображении нескольких ЛА необходимо учитывать реальный масштаб изображения и относительные габариты ЛА

наименьшего размера. Во избежание пропусков ЛА, необходимо, чтобы определяемое количество центров кластеров на изображении совпадало с количеством ЛА наименьшего размера, способных поместиться на входном изображении с известным масштабом. На рис. 3 представлен пример работы алгоритма определения зон интереса на реальном спутниковом снимке размером 767×778 пикселей, на котором присутствуют несколько типов ЛА, различные подстилающие поверхности и сооружения. После предварительной обработки входное изображение для сравнения было подано для анализа в однослойные сеть и самоорганизующуюся карту Кохонена с начальным расположением нейронов в узлах двумерной сетки с шестиугольными ячейками. Для SOM необходимо около 150 эпох обучения, а для сети Кохонена достаточно 1. Исходя из масштаба входного изображения количество ЛА наименьшего размера, способных поместиться на изображении, составляет 8×8 , следовательно, на выходе нейронной сети должны быть определены 64 центра кластеров. Если центры кластеров будут расположены относительно близко друг к другу, то возможно их объединение в один, расположенный в центре между ними. То же касается и “окон” вокруг соседних центров кластеров, если они перекрываются более чем на 90 % площади. Также можно исключить из анализа центры, находящиеся в непосредственной близости от края изображения, так как они не могут быть центром ЛА, которые могла бы распознать сверточная сеть при дальнейшей обработке.

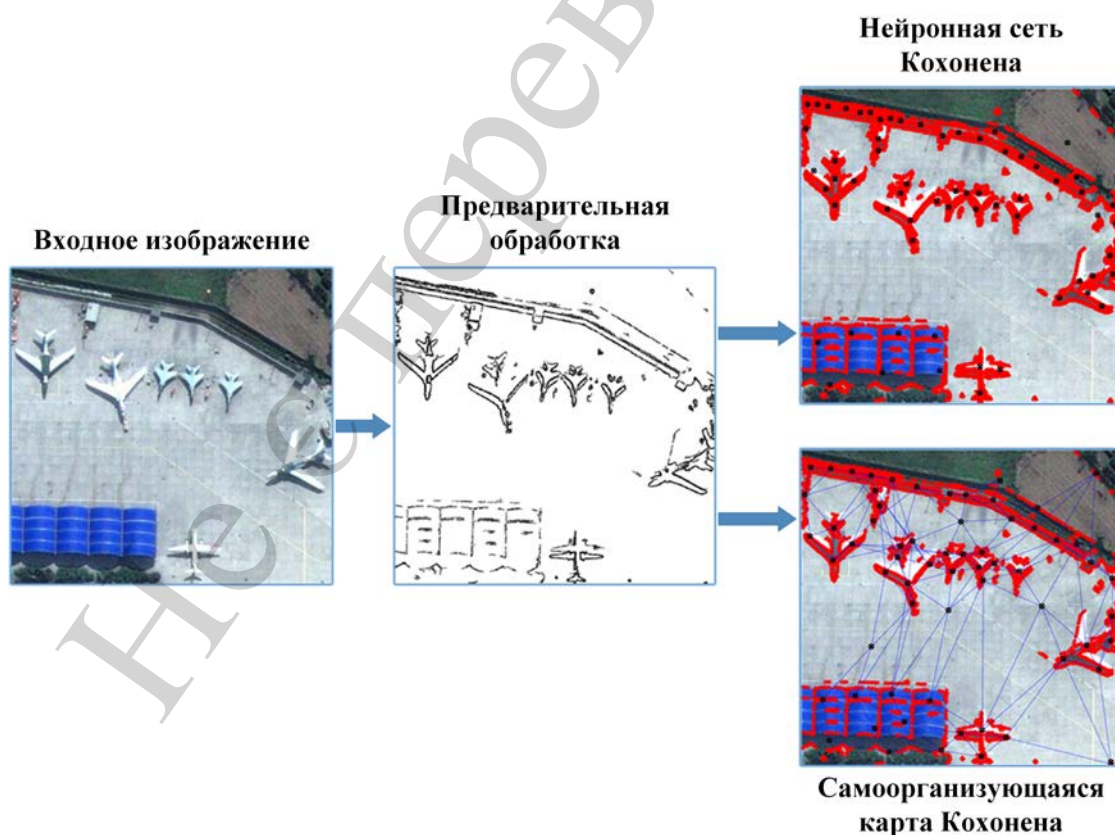


Рис. 3. Первичная обработка входного изображения с дальнейшим анализом сетью и самоорганизующейся картой Кохонена

Как видно из рис. 3, сеть и карты Кохонена определили центры кластеров таким образом, что все ЛА будут выделены “окном” и пойдут на дальнейший анализ в сверточную сеть. Очевидным минусом карты является зависимость положения центров кластеров от соседних центров, что приводит к наличию центров в областях с отсутствием перепадов яркости. Данного недостатка лишена сеть Кохонена, которая практически всегда располагает центры внутри скопления точек с перепадом яркости. Благодаря этому из дальнейшего анализа исключаются пустые области и обеспечивается выигрыш по времени анализа и затратам вычислительных ресурсов. В данном случае, после выделения “окном” зон интереса вокруг центров кластеров, сеть Кохонена отправляет на дальнейший анализ порядка 80 вырезанных кадров размером с максимальный ЛА, самоорганизующаяся карта Кохонена – около 120. Это в 1000 раз меньше, чем при использовании последовательного разбиения входного изображения на кадры или в 15–100 раз меньше, чем при использовании алгоритмов, аналогичных RCNN, YOLO и SSD. Также применение данного алгоритма существенно сокращает необходимое количество обучающих образов ЛА для сверточной сети, т.к. размер “окна” привязан к масштабу изображения и равен размеру наибольшего детектируемого ЛА. Это означает, что отпадает необходимость вариации размеров ЛА в обучающих образах. Данное решение позволяет сократить размер обучающей выборки в несколько раз, ускорить обучение сети и увеличить точность распознавания.

В табл. 2 представлено сравнение быстродействия сети и карты Кохонена в зависимости от разрешения входного изображения и количества возможных объектов минимального размера. Вычисления выполнены в программе собственной разработки без применения GPU (графический процессор), алгоритмов ускорения и оптимизации на персональном компьютере с параметрами, указанными в п. 4.

Таблица 2

Сравнение быстродействия сети и карты Кохонена

Размер входного изображения, пикс.	112×218	202×387	767×778
Количество объектов поиска, шт.	3	10	64
Обработка сетью Кохонена, сек.	0,6	1,1	5,3
Обработка картами Кохонена, сек.	0,1	0,3	8,5

Как видно из табл. 2, скорость обработки изображений сетью и картой Кохонена имеет различную зависимость от размеров изображения и количества объектов поиска. При малых и средних размерах изображения и количестве объектов поиска порядка 1–30 шт., самоорганизующаяся карта показывает в 3–6 раз более высокую скорость обработки, чем сеть Кохонена. При дальнейшем увеличении размеров изображения и количества объектов поиска, карта Кохонена постепенно отстает от сети по скорости обработки.

8. Обсуждение результатов работы созданного алгоритма поиска зон интереса на подстилающей поверхности

Существующие алгоритмы детектирования и распознавания объектов выделяют на входном изображении несколько тысяч или даже сотен тысяч частей и подают их на анализ в сверточную сеть, чем существенно увеличивают время анализа. Предлагаемый в данной статье алгоритм позволяет сократить количество анализируемых частей входного изображения и уменьшить объем обучающей выборки для сверточной сети, что позволяет существенно уменьшить время поиска и распознавания необходимых объектов. Это достигается за счет предварительного поиска зон интереса на изображении с дальнейшим анализом только найденных областей. Определение размеров “окна”, сканирующего определенные зоны интереса, исходя из масштаба изображения и заданных габаритов самого большого детектируемого объекта, исключает необходимость вариации размеров ЛА в обучающих образах. Выделение границ объектов и применение сети и самоорганизующейся карты Кохонена позволяют определить центры кластеров на изображении и создать вокруг них зоны интереса. Данные решения позволяют исключить вариации по масштабу и центрируют искомый объект, за счет чего в несколько раз сокращается размер обучающей выборки, ускоряется обучение сети и увеличивается точность распознавания.

Недостатком созданного алгоритма является тот факт, что существенный прирост скорости распознавания проявляется только при анализе изображений, на которых границы объектов не занимают большую часть пространства. Например, на спутниковых и радиолокационных снимках, когда искомые объекты расположены на однородной подстилающей поверхности большой протяженности. При анализе изображений, занятых небольшими объектами, алгоритм определит как зону интереса все изображение целиком, при этом, все равно сократив количество анализируемых частей входного изображения. Анализируемых частей будет в 300–500 раз меньше, чем при использовании последовательного разбиения или в 3–10 раз меньше, чем при использовании алгоритмов, аналогичных RCNN, YOLO и SSD. В дальнейшем планируется выявлять уникальные черты формы границ искомого объекта, что поможет выделить эти объекты на фоне границ подстилающей поверхности и границ не интересующих объектов.

Алгоритм предварительной обработки показал высокую скорость выделения границ объектов входного изображения. Поиск центров кластеров с помощью самоорганизующейся карты и сети Кохонена позволяет обрабатывать с высокой скоростью изображения размером до 500x500 пикс. и количеством объектов поиска не более 40 шт. В дальнейшем планируется применить для проведения вычислений GPU, алгоритмы ускорения и оптимизации сетей Кохонена, а также рассмотреть иные варианты поиска центров кластеров.

Представленная методика может использоваться не только для распознавания типа ЛА, но и иных объектов, различимых на фоне подстилающей поверхности и представляющих интерес. В дальнейшей работе планируется усовершенствование представленного алгоритма за счет более качественного выделения границ объектов и поиска их центров. Также планируются создание и

оптимизация сверточной нейронной сети для распознавания интересующих объектов, что позволит создать полноценную систему поиска и распознавания объектов на радиолокационном или спутниковом изображениях подстилающей поверхности.

9. Выводы

1. Для выделения границ объектов, присутствующих на изображении, использован алгоритм, основанный на применении оператора Собеля. Алгоритм обладает высоким быстродействием, выделяя границы объектов входного изображения за 0,002–0,02 сек.

2. Для поиска центров объектов использованы нейронная сеть и самоорганизующаяся карта Кохонена, так как они являются одними из наиболее быстрых алгоритмов определения центров кластеров входных данных.

3. Разработанный алгоритм для определения зон интереса и их центров был протестирован на реальных спутниковых изображениях подстилающей поверхности. Применение алгоритма позволяет уменьшить количество анализируемых сверточной сетью частей входного изображения в 15–100 раз, что соответственно сокращает время поиска и распознавания необходимых объектов. Также использование данного алгоритма сокращает необходимое количество обучающих образов для сверточной сети, т. к. размер «окна» привязан к масштабу изображения и равен размеру наибольшего детектируемого объекта. Данный факт, а также центрирование объекта на обучающих образах, позволяют ускорить обучение сети более чем в 5 раз и увеличить точность распознавания минимум на 10 %.

Сеть Кохонена оказалась эффективнее самоорганизующейся карты применительно к данной задаче, так как в 3 раза реже размещает центры кластеров на подстилающей поверхности. Однако карта Кохонена дает выигрыш по скорости анализа изображений размером до 500×500 пикс. и количеством объектов поиска не более 40 шт. При увеличении приведенных выше параметров SOM начинает уступать в скорости обработки сети Кохонена.

Литература

1. Simard P. Y., Steinkraus D., Platt J. C. Best practices for convolutional neural networks applied to visual document analysis // Proceedings of the Seventh International Conference on Document Analysis and Recognition. 2003. URL: <http://cognitivemedium.com/assets/rmnist/Simard.pdf>

2. Новикова Н. М., Дуденков В. М. Моделирование нейронной сети для распознавания изображений на основе гибридной сети и самоорганизующихся карт Кохонена // Аспирант. 2015. № 2. С. 31–34.

3. Агаян К. Ю., Ханжин В. Г. Нейронная сеть с архитектурой Кохонена для распознавания изображений // Прочность неоднородных структур – ПРОСТ 2018. Сборник трудов IX-ой Евразийской научно-практической конференции. 2018. С. 153.

4. Герасимова Н. И., Верхотурова А. Э. Поиск фрагмента изображения с использованием нейронной сети Кохонена // Информационные технологии в

науке, управлении, социальной сфере и медицине: сборник научных трудов Международной конференции с международным участием. Ч. 1. Томск, 2014. С. 68–70.

5. Солдатова О. П., Чайка П. Д. Исследование эффективности решения задачи классификации гибридными сетями Кохонена // Известия Самарского научного центра Российской академии наук. 2015. Т. 17, № 2. С. 1147–1152.

6. Нарушев И. Р. Нейронная сеть на основе самоорганизующейся карты Кохонена как средство выявления аномального поведения // Охрана, безопасность, связь. 2018. Т. 2, № 3 (3). С. 194–197.

7. Kajan S., Sekaj I., Lajtman M. Cluster Analysis Applications in Matlab Using Kohonen Network. URL: <https://pdfs.semanticscholar.org/a8ba/6977dce4bdbeec3dd370eb614de2c6f56514.pdf>

8. LeCun Y., Bengio Y. Convolutional networks for images, speech, and time series // The handbook of brain theory and neural networks. 1998. P. 255–258.

9. LeCun Y., Kavukcuoglu K., Farabet C. Convolutional networks and applications in vision // Proceedings of 2010 IEEE International Symposium on Circuits and Systems. 2010. doi: <https://doi.org/10.1109/iscas.2010.5537907>

10. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation / Girshick R., Donahue J., Darrell T., Malik J. // 2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014. doi: <https://doi.org/10.1109/cvpr.2014.81>

11. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation / Girshick R., Donahue J., Darrell T., Malik J. // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016. Vol. 38, Issue 1. P. 142–158. doi: <https://doi.org/10.1109/tpami.2015.2437384>

12. Girshick R. Fast R-CNN // 2015 IEEE International Conference on Computer Vision (ICCV). 2015. doi: <https://doi.org/10.1109/iccv.2015.169>

13. Faster R-CNN: Towards real-time object detection with region proposal networks / Ren S. et. al. // Advances in neural information processing systems. 2015. P. 91–99.

14. Mask R-CNN / He K., Gkioxari G., Dollar P., Girshick R. // 2017 IEEE International Conference on Computer Vision (ICCV). 2017. doi: <https://doi.org/10.1109/iccv.2017.322>

15. You Only Look Once: Unified, Real-Time Object Detection / Redmon J., Divvala S., Girshick R., Farhadi A. // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. doi: <https://doi.org/10.1109/cvpr.2016.91>

16. SSD: Single Shot MultiBox Detector / Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A. C. // Computer Vision – ECCV 2016. 2016. P. 21–37. doi: https://doi.org/10.1007/978-3-319-46448-0_2

17. Хайкин С. Нейронные сети: полный курс. 2-е изд. М.: Издательский дом Вильямс, 2008.

18. Kohonen T. Self-organizing maps. Vol. 30. Springer Science & Business Media, 2001. 502 p. doi: <https://doi.org/10.1007/978-3-642-56927-2>

19. Gersho A., Gray R. M. Vector quantization and signal compression. Vol. 159. Springer Science & Business Media, 1992. 732 p. doi: <https://doi.org/10.1007/978-1-4615-3626-0>

ТІЛЬКИ ДЛЯ ЧИТАННЯ