

АЛГОРИТМ И ПРОГРАММНЫЙ КОМПЛЕКС ВЫСОКОСКОРОСТНОЙ НЕЙРОСЕТЕВОЙ СЕГМЕНТАЦИИ ОБЛАЧНОСТИ НА ПАНХРОМАТИЧЕСКИХ ИЗОБРАЖЕНИЯХ ОТ КОСМИЧЕСКИХ АППАРАТОВ «РЕСУРС-П»

Москвитин А.Э., в.н.с. НИИ «Фотон» РГРТУ, д.т.н., foton@rsreu.ru

Ушенкин В.А., в.н.с. НИИ «Фотон» РГРТУ, к.т.н., foton@rsreu.ru

Ларюков С.А., м.н.с. НИИ «Фотон» РГРТУ, foton@rsreu.ru

ALGORITHM AND SOFTWARE OF HIGH-SPEED NEURAL NETWORK CLOUD SEGMENTATION OF RESURS-P SATELLITE'S PANCHROMATIC IMAGES

Moskvitin A.E., Ushenkin V.A., Laryukov S.A.

The paper considers the problem of cloud segmentation on panchromatic images obtained by the Geoton highly detailed sensor of the Resurs-P Earth remote sensing satellite. The architecture of an artificial neural network, which is characterized by lower computational costs and provides the same segmentation quality as more complex known architectures, is proposed. The question of choosing the optimal level of preliminary radiometric processing of segmented images is considered. The results of developed software experimental studies and the achieved values of segmentation quality measures are presented.

Key words: panchromatic image, cloud segmentation, remote sensing, artificial neural network.

Ключевые слова: панхроматическое изображение, сегментация облачности, дистанционное зондирование Земли, искусственная нейронная сеть.

Введение

Наличие облачности является фактором, ограничивающим применение данных дистанционного зондирования Земли (ДЗЗ) видимого и инфракрасного диапазонов. Сплошная облачность загораживает от космического аппарата объекты земной поверхности, не давая их наблюдать, а полупрозрачная дымка искажает яркость и контрастность наблюдаемых объектов, существенно затрудняя их дешифрацию и измерение физических свойств. При стереофотограмметрической обработке [1, 2] пар космических снимков наличие облачности приводит к тому, что измеряется высота не рельефа земной поверхности, а высота облака. В связи с этим при обработке космических снимков видимого и инфракрасного диапазонов [3] необходимы выявление и отбраковка участков изображения, на которых представлены не наземные объекты, а облачность.

Ручное выделение облачности на изображениях является крайне трудоемкой задачей и существенно ограничивает оперативность получения обработанных космических снимков. Поэтому актуально развитие автоматических методов сегментации облачности.

Для обнаружения облачности на изображениях могут использоваться различные признаки: яркостной (облака обычно имеют высокую яркость), колориметрический (облака обычно имеют серый или белый цвет), текстурные (облака обычно имеют гладкую текстуру и плавные границы), стереофотограмметрический (при наличии нескольких снимков, снятых со стереобазой, облака из-за возвышения над поверхностью Земли имеют большие геометрические рассогласования на таких снимках, чем наземные объекты) [4, 5]. При этом наиболее сложной выглядит задача сегментации об-

Рассматривается задача сегментации облачности на панхроматических изображениях, полученных высокодетальной съемочной аппаратурой «Геотон» космических аппаратов дистанционного зондирования Земли «Ресурс-П». Предлагается архитектура искусственной нейронной сети, позволяющая с меньшими вычислительными затратами достичь качества сегментации, не уступающего более сложным известным архитектурам. Рассматривается вопрос выбора оптимального уровня предварительной радиометрической обработки сегментируемых изображений. Приводятся результаты экспериментальных исследований разработанного программного комплекса и достигнутые значения показателей качества сегментации.

лачности на панхроматических изображениях, для которых доступны только два вида признаков: яркостной и текстурные. Задача дополнительно усложняется, если панхроматические изображения сформированы отечественной аппаратурой «Геотон», устанавливаемой на космические аппараты серии «Ресурс-П». Данная аппаратура имеет ограниченный динамический диапазон. Перед съемкой участка земной поверхности предсказываются его энергетические характеристики и выбирается соответствующий коэффициент усиления. При этом все объекты, энергетическая яркость которых превышает некоторый порог насыщения, оказываются засвеченными. За счет того, что облака, как правило, намного ярче наземных объектов, в большинстве случаев на панхроматических снимках аппаратуры «Геотон» они оказываются засвеченными, а для их автоматической сегментации не остается практически никаких признаков. Текстура облачности теряется, а яркость искажается. Лишь по наличию тени, характерной форме, площади и более темным граничным участкам, не подвергшимся искажению, с той или иной долей вероятности человек может распознать на снимке облако и отличить его от других засвеченных ярких объектов.

Для автоматического решения такой сложной задачи детерминированные алгоритмы оказываются мало пригодны. Однако развитие за последние годы методов искусственного интеллекта и в частности искусственных нейронных сетей (ИНС) позволяет получить достаточно качественное решение даже таких сложных задач.

Имеется ряд работ по применению ИНС в задаче сегментации облачности на мультиспектральных изображениях [6-9]. В работе [10] ИНС успешно применены для сегментации облачности на панхроматических изображениях. Однако известные архитектуры ИНС характеризуются большим числом весовых коэффициентов и относительно большим объемом вычислений, что существенно затрудняет их обучение и ограничивает скорость нейросетевой обработки.

В настоящей работе ставится задача разработки архитектуры ИНС, характеризующейся меньшим количеством весовых коэффициентов и меньшим объемом вычислений, при сохранении качества сегментации облачности.

Выбор показателей качества сегментации

В настоящее время для оценки качества бинарной сегментации используется множество различных показателей. Рассмотрим основные из них.

Все показатели предполагают, что имеется результат сегментации в виде бинарного растрового изображения, где один код пикселя соответствует искомому объекту, а другой – окружающему фону. В задаче сегментации облачности искомым объектом являются сплошная облачность и полупрозрачная дымка, а окружающим фоном – наземные объекты. Также должен иметься эталонный результат, полученный, например, в результате ручной сегментации.

Первый показатель – общая точность сегментации a , определяемая как отношение числа пикселей, коды которых совпадают на эталонном и полученном результатах сегментации, к общему числу сегментируемых пикселей.

До единицы показатель a дополняют ошибки первого и второго рода. Под ошибкой первого рода понимается ложное выделение искомого объекта (на результате сегментации представлен объект, а на эталоне – фон). Под ошибкой второго рода понимается пропуск объекта (на результате сегментации представлен фон, а на эталоне – объект).

Абсолютное значение числа ошибок первого и второго рода в пикселях изображения обычно не используется. Оно заменяется относительными значениями, нормированными общей площадью окружающего фона или искомым объектом. Таким образом, вероятность возникновения ошибки первого рода α – это отношение числа пикселей изображения, ложно распознанных как объект, к числу пикселей, соответствующих окружающему фону на эталоне. Вероятность возникновения ошибки второго рода β – это отношение числа пикселей изображения, ошибочно отнесенных к окружающему фону, к числу пикселей, соответствующих искомому объекту на эталоне.

Пусть эталонная доля пикселей изображения, принадлежащих искомому объекту, составляет δ , тогда общая точность сегментации может быть выражена через ошибки первого и второго рода как

$$a = 1 - (1 - \delta)\alpha - \delta\beta. \quad (1)$$

Недостатком показателя a является тот факт, что при нулевых значениях δ он вырождается в $a = 1 - \alpha$, т.е., когда искомым объектов крайне мало, оказывается неважным, пропускаются они или нет; главное, чтобы не было ложных выделений. В обратном случае при $\delta = 1$ показатель вырождается в $a = 1 - \beta$, т.е. учитываются только пропуски искомого объекта, а ложные выделения игнорируются.

В ряде задач сегментации, особенно в области медицины, приведенное поведение показателя a является недопустимым, поэтому для таких случаев были разработаны другие показатели.

Мера Жаккара J – отношение числа пикселей, верно отнесенных к объекту, к сумме эталонного числа пикселей объекта и числа пикселей, ошибочно отнесенных к объекту. Мера Жаккара при $\delta \rightarrow 0$ ведет себя еще менее приемлемым образом. Если выразить меру Жаккара через α , β и δ , получится:

$$J = \frac{\delta(1 - \beta)}{\delta + (1 - \delta)\alpha}. \quad (2)$$

При $\delta \rightarrow 0$ мера Жаккара вырождается в $J = \begin{cases} 0, & \forall \alpha > 0, \\ \text{неопределенность}, & \alpha = 0. \end{cases}$ Таким образом, на значение показателя перестают влиять не только ошибки второго рода, но и положительные значения ошибок первого рода. В отсутствие ошибок первого рода и в отсутствие искомым объектов мера Жаккара вырождается в полную неопределенность, поскольку и числитель, и знаменатель в (2) обращаются в ноль. В обратном случае (т.е. при $\delta \rightarrow 1$) $J = 1 - \beta$, т.е. ведет себя аналогично показателю a .

В качестве альтернативы a , α и β введены показатели точности обнаружения объекта p , полноты обнаружения r и F-меры.

Точность обнаружения объекта схожа с мерой Жаккара и определяется как отношение числа пикселей, верно отнесенных к искомому объекту, к общему числу пикселей, отнесенных на результате сегментации к искомому объекту:

$$p = \frac{\delta(1 - \beta)}{\delta(1 - \beta) + (1 - \delta)\alpha}. \quad (3)$$

Поведение меры p при $\delta \rightarrow 0$ точно такое же, как у меры Жаккара. При $\delta \rightarrow 1$ мера вырождается в единицу независимо от значений α и β .

Полнота обнаружения объекта определяется как отношение числа пикселей, верно отнесенных к искомому объекту, к эталонному числу пикселей объекта, т.е. фактически дополняет β до единицы:

$$r = 1 - \beta. \quad (4)$$

F-мера является средним гармоническим мер p и r :

$$F = \frac{2}{p^{-1} + r^{-1}}. \quad (5)$$

Также известна φ -мера (или коэффициент корреляции Мэттью), определяемая как:

$$\varphi = \frac{\delta(1-\beta)(1-\delta)(1-\alpha) - (1-\delta)\alpha\delta\beta}{\sqrt{\delta(1-\delta)(\delta(1-\beta) + (1-\delta)\alpha)((1-\delta)(1-\alpha) + \delta\beta)}}. \quad (6)$$

При $\delta \rightarrow 0$ и $\delta \rightarrow 1$ φ -мера вырождается в неопределенность (числитель и знаменатель в (6) обнуляются).

Таким образом, из приведенного анализа следует, что, несмотря на свои недостатки, в экстремальных случаях, когда $\delta \rightarrow 0$ или $\delta \rightarrow 1$, наименее неадекватно ведет себя мера общей точности сегментации a .

Исправить ее недостатки на несбалансированных данных, когда $\delta \ll 0,5$ или $\delta \gg 0,5$, можно путем введения поправочных коэффициентов k_α и k_β :

$$\hat{a} = 1 - k_\alpha(1-\delta)\alpha - k_\beta\delta\beta. \quad (7)$$

В экстремальных ситуациях, когда $\delta = 0$ или $\delta = 1$, показатель \hat{a} будет также игнорировать одну из величин α или β , которая в этот момент является неопределенной, но по мере приближения к экстремальным ситуациям за счет коэффициентов k_α и k_β влияние величин α и β на \hat{a} можно сделать более сбалансированным.

Применительно к задаче сегментации облачности введение коэффициентов k_α и k_β видится нецелесообразным. Ситуация пропуска облачности так же плоха, как ситуация ложного ее обнаружения. В первом случае облачность не отбракуется и негативно отразится на последующих видах обработки изображения, а во втором случае будет ошибочно проигнорирована полезная информация. Показатель a оценивает суммарное количество ситуаций обоих видов.

Переход от показателя качества сегментации к дифференцируемой функции потерь при машинном обучении

Наиболее распространенной функцией потерь при машинном обучении в задаче бинарной сегментации является бинарная перекрестная энтропия:

$$H = -\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (e_{ij} \log_2 m_{ij} + (1-e_{ij}) \log_2 (1-m_{ij})), \quad (8)$$

где $(e_{ij})_{M \times N}$ – матричное представление эталонного результата сегментации размером $M \times N$ пикселей, где $e_{ij} = 1$ соответствует наличию искомого объекта, а $e_{ij} = 0$ – отсутствию; $(m_{ij})_{M \times N}$ – матричное представление полученного результата сегментации.

Основным достоинством такой функции потерь является ее дифференцируемость, что позволяет обучать ИНС методами на основе градиентного спуска. Однако при $m_{ij} = 0$ или $m_{ij} = 1$ из-за операции логарифмирования бинарная перекрестная энтропия устремляется в бесконечность. Поэтому растровые маски искомого

объекта m_{ij} , получаемые в результате машинного обучения с данной функцией потерь, содержат лишь промежуточные значения $0 < m_{ij} < 1$, не позволяющие принять однозначное решение, соответствует данный пиксель искомому объекту или нет. Требуется дополнительный поиск некоторого порога бинаризации маски $(m_{ij})_{M \times N}$. Кроме того, поскольку для оценки качества сегментации используются совершенно другие показатели, рассмотренные ранее, целесообразно при машинном обучении применять функцию потерь, максимально близкую к этим показателям.

Напрямую использовать показатель a в качестве функции потерь невозможно, поскольку он не является дифференцируемым. При его расчете подсчитывается количество пикселей, которое является дискретной недифференцируемой величиной. Однако можно рассчитать a как частный случай дифференцируемого выражения. Покажем это.

Пусть результат сегментации $(m_{ij})_{M \times N}$ бинаризован, т.е. m_{ij} может быть только нулем или единицей. Эталонный результат сегментации $(e_{ij})_{M \times N}$ предполагается бинаризованным изначально. В этом случае число верно выявленных пикселей объекта можно записать как $\sum_{i=1}^M \sum_{j=1}^N e_{ij} m_{ij}$, а число верно сегментированных пикселей фона как $\sum_{i=1}^M \sum_{j=1}^N (1-e_{ij})(1-m_{ij})$. Таким образом, если $(m_{ij})_{M \times N}$ бинаризована, общая точность сегментации может быть выражена как

$$\tilde{a} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (e_{ij} m_{ij} + (1-e_{ij})(1-m_{ij})). \quad (9)$$

Выражение (9) является дифференцируемым и его дополнение до единицы может использоваться в качестве функции потерь $1 - \tilde{a}$. Если $(m_{ij})_{M \times N}$ содержит только нули и единицы, \tilde{a} строго соответствует общей точности сегментации. Если же $(m_{ij})_{M \times N}$ содержит промежуточные значения, \tilde{a} является некоторой приближенной оценкой показателя a .

Проведенные экспериментальные исследования показали, что на практике в ходе машинного обучения ИНС ей оказывается выгодно сформировать маску $(m_{ij})_{M \times N}$, на которой практически не содержится промежуточных значений m_{ij} , поэтому функция \tilde{a} с высокой точностью соответствует целевому показателю качества сегментации a . При этом отпадает необходимость выбора порога для бинаризации $(m_{ij})_{M \times N}$, поскольку маска на выходе ИНС уже практически бинаризована. Для тех редких пикселей, в которых $0 \ll m_{ij} \ll 1$, может быть выбран произвольный порог бинаризации, например, равный 0,5, поскольку на итоговой точности сегментации он почти не отражается.

Такое поведение ИНС может быть достигнуто следующим образом. Функцией активации выходного слоя

ИНС, используемых для бинарной сегментации, как правило, является сигмоида. Для того чтобы сигмоида была близка к нулю, ей на вход нужно подать большое по модулю отрицательное значение. Чтобы сигмоида была близка к единице, ей на вход должно поступить большое положительное значение. Последний слой ИНС обычно выполняет некоторое линейное преобразование яркости пикселей изображения. Чтобы обеспечить такой выход сигмоиды, он должен за счет больших мультипликативных коэффициентов значительно усилить сигнал. Затем к усиленному сигналу прибавляется некоторый аддитивный коэффициент, в результате чего сигнал оказывается положительным или отрицательным. По сути этот аддитивный коэффициент последнего слоя ИНС и является порогом, по которому бинаризуется результат сегментации. Он вместе с другими весовыми коэффициентами ИНС автоматически подбирается в ходе машинного обучения.

Другим достоинством функции потерь $1 - \tilde{a}$ является ее линейность. Машинное обучение ИНС обычно выполняется в пакетном режиме, когда вся обучающая выборка делится на пакеты из небольшого числа изображений. Одна итерация метода градиентного спуска и коррекции весов ИНС путем обратного распространения ошибки выполняется в пределах пакета. Соответственно функция потерь рассчитывается только по небольшому числу изображений, представленных в пакете. Перебор всех пакетов обучающей выборки составляет одну эпоху обучения. По окончании эпохи рассчитываются потери на всей выборке путем усреднения значений, полученных для каждого пакета. Результаты усреднения используются для проверки, улучшила ли свою работу ИНС в ходе эпохи обучения. Если функция потерь является линейной, то усреднение ее значений, рассчитанных для отдельных пакетов, дает строгое значение функции потерь для всей обучающей выборки. Если функция потерь является нелинейной, то для получения истинного ее значения для всей обучающей выборки необходимы более сложные алгоритмы комбинирования значений потерь, рассчитанных по отдельным пакетам. Существующие широко используемые средства машинного обучения такие сложные алгоритмы обычно не поддерживают, а усреднение потерь, рассчитанных по отдельным пакетам, искажает нелинейные целевые функции, оптимизируемые при обучении.

Для сверточных ИНС, которые по сути эквивалентны некоторому нелинейному фильтру конечной апертуры, при расчете потерь нужно учитывать наличие краевых эффектов фильтрации. Для этого либо размеры выхода ИНС должны быть уменьшены по сравнению с входом на апертуру фильтра, либо расчет потерь должен вестись лишь по центральной части выходного изображения. Если вовлекать области краевых эффектов в расчет потерь, ИНС при обучении будет вынуждена уменьшать размер окрестности, которую она анализирует, обнуляя часть своих коэффициентов, чтобы ослабить краевые эффекты. В этом случае архитектура ИНС будет использоваться не полностью, что может негативно отразиться на качестве результата.

Выбор архитектуры ИНС

Наиболее распространенной архитектурой ИНС, применяемой в задачах бинарной сегментации, является «U-Net» [11]. Это сверточная ИНС, основанная на пирамидальной обработке. Каждый следующий уровень пирамиды формируется путем выполнения двух сверток в окне 3×3 с активационной функцией типа «линейный выпрямитель» (англ. rectified linear unit, ReLU) и увеличения шага дискретизации в 2 раза с выбором максимального значения в окне 2×2 . На каждом следующем уровне пирамиды число каналов изображения увеличивается в 2 раза. Пирамида содержит 6 уровней. После достижения верхнего уровня начинается обратная операция. Шаг дискретизации изображения уменьшается в 2 раза с выполнением интерполяции и сокращением в 2 раза числа каналов. К полученному изображению в качестве дополнительных каналов добавляется предыдущий слой пирамиды, затем выполняются две свертки в окне 3×3 с активационной функцией типа «линейный выпрямитель» и снова сокращается в 2 раза число каналов. По достижении нижнего слоя пирамиды выполняется линейное преобразование яркости изображения с сигмоидой в качестве активационной функции и сокращением числа каналов до одного. Графическое представление архитектуры ИНС (рис. 1) напоминает букву «U», что определило название архитектуры. Число каналов, формируемых на нижнем слое пирамиды, в классической архитектуре «U-Net» составляет 64.

Авторы архитектуры «U-Net» применили ее для сегментации медицинских изображений. В их задаче положительного результата удалось достичь на небольшой обучающей выборке.

В задаче сегментации облачности на панхроматических изображениях обучающую выборку приходится делать больших размеров. Это связано с большим разнообразием объектов, которые могут быть представлены на космических снимках. Обучающая выборка должна включать все основные виды объектов, чтобы все из них ИНС научилась отделять от облачности. Кроме того, сама облачность отличается крайне широким разнообразием по форме, структуре, яркости, прозрачности. Одни облака могут находиться в тени других облаков и быть хуже освещенными. Полупрозрачная дымка может располагаться над любыми другими видами объектов.

Машинное обучение сложных многослойных архитектур ИНС на больших обучающих выборках крайне затруднительно, поскольку требует огромного объема вычислений. При большом числе слоев из-за накапливающихся погрешностей алгоритма обратного распространения ошибки сходимость машинного обучения крайне низкая. Требуются десятки и сотни тысяч эпох обучения для минимизации целевой функции, а если обучающая выборка большая, то одна эпоха может длиться несколько часов даже на современных графических ускорителях вычислений.

В работе [12] был предложен механизм существенно ускорения обучения многослойных ИНС – пакетная нормализация. В рамках этого механизма в ИНС после каждого сверточного слоя (кроме выходного) добавля-

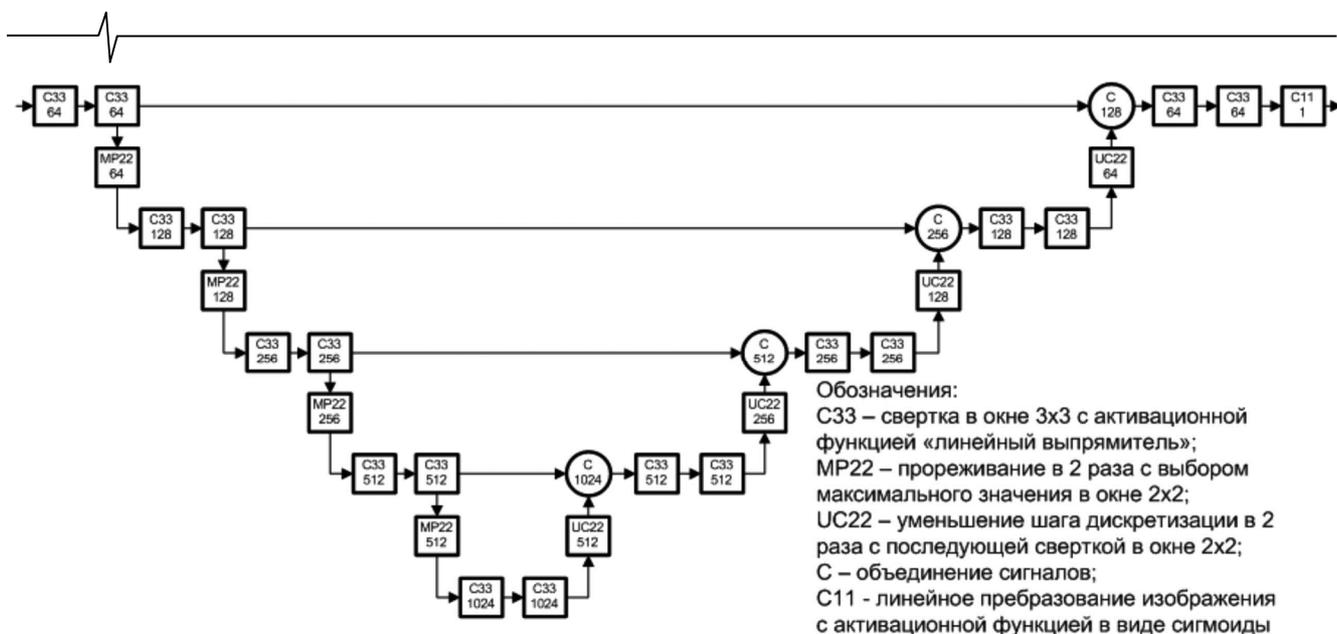


Рис. 1. Классическая архитектура ИНС «U-Net»

ются слои нормализации, которые в ходе обучения в пределах пакета рассчитывают среднее значение и дисперсию сигнала, а затем выполняют линейное преобразование сигнала, обнуляя его среднее и делая дисперсию единичной. Коэффициенты линейного преобразования оказываются уникальными для каждого пакета, однако в ходе обучения запоминаются усредненные значения этих коэффициентов, которые используются при штатной работе обученной ИНС, когда уже нет никаких пакетов, а обрабатывается то или иное отдельное изображение.

Пакетная нормализация позволяет сократить машинное обучение ИНС до нескольких десятков или сотен эпох, однако потери, рассчитываемые в ходе обучения и получаемые в ходе последующего штатного функционирования, разнятся из-за различной работы слоев пакетной нормализации в этих режимах. Поэтому для корректного обучения необходимо в пределах каждой эпохи рассчитывать функцию потерь дважды: один раз в ходе итераций градиентного спуска в режиме обучения и второй раз в конце эпохи в режиме штатной работы для оценки истинного значения потерь.

Другой проблемой машинного обучения ИНС является так называемое переобучение. С точки зрения математики минимизация целевой функции в пределах обучающей выборки не гарантирует, что целевая функция будет принимать малые значения за пределами этой выборки. Поэтому иногда случается, что обученная ИНС хорошо работает в пределах обучающей выборки, но выдает совершенно неприемлемые результаты на других изображениях. Особенно часто это происходит, когда выборка непредставительна (т.е. не охватывает все типовые варианты, которые в дальнейшем могут быть поданы на вход ИНС) или мала по объему, когда ИНС в своих коэффициентах может просто запомнить таблицу преобразования небольшого количества комбинаций на входе в требуемый выход.

Чтобы избежать переобучения, обычно выделяют две выборки: обучающую и валидационную. Обучающая выборка делится на пакеты, по которым выполняются итерации градиентного спуска, а по окончании

эпохи выполняется расчет потерь по валидационной выборке. Т.е. для коррекции весов ИНС используются значения потерь, полученные по обучающей выборке, в т.ч. с применением пакетной нормализации, а минимизируется целевая функция, рассчитанная по валидационной выборке в режиме штатной работы ИНС.

Чтобы пакетная нормализация работала корректно, необходимо выбрать достаточно большой размер пакета и более менее равномерно размещать обучающую выборку, чтобы в каждом пакете на изображениях было представлено как можно больше различных видов объектов. Тогда статистические характеристики сигналов на выходе сверточных слоев в рамках каждого пакета будут достаточно стабильными, и их усредненные значения, которые используются при штатной работе ИНС, будут близкими к ним. В этом случае ИНС обучается быстро и способна достичь малых значений функции потерь как на обучающей, так и на валидационной выборках. Если же размер пакета недостаточно большой, потери на обучающей выборке снижаются достаточно быстро, но в режиме штатной работы на валидационной выборке потери остаются стабильно высокими. Однако размер пакета, который можно выбрать, ограничен объемом памяти графического ускорителя. Нормализация требует одновременного доступа к данным всего пакета изображений для расчета их статистических характеристик. Число слоев нормализации в многослойных сетях измеряется десятками. Если все данные не будут размещены в быстросействующей памяти, время обработки одного пакета катастрофически увеличивается, что лишает смысла пакетную нормализацию как способ ускорения машинного обучения.

Классическая архитектура «U-Net» была разработана без учета применения в сочетании с пакетной нормализацией. В то же время в задаче сегментации облачности без применения пакетной нормализации обучить ИНС за приемлемое время не представляется возможным. Поэтому в архитектуру «U-Net» были добавлены слои нормализации. Однако из-за большого числа слоев и большого числа каналов на выходе промежуточных слоев ИНС «U-Net» в 11 Гбайт памяти графического ускорителя

NVIDIA GeForce RTX 2080 Ti удается разместить пакет только из 3 изображений 512×512 пикселей. Это слишком малый размер пакета, чтобы в нем удалось представить все возможные виды объектов. В результате значение функции потерь $1-\tilde{a}$ на валидационной выборке не опускается ниже 0,0584.

В связи с этим было принято решение в 2 раза сократить число каналов во всех промежуточных слоях ИНС «U-Net», за счет чего становится возможным увеличение размера пакета до 6 изображений. В этом случае удалось достичь значения функции потерь $1-\tilde{a}$ на валидационной выборке, равное 0,0228.

Однако при этом времена обучения ИНС и нейросетевой сегментации остались достаточно высокими. Поэтому были проведены эксперименты по дальнейшему упрощению архитектуры «U-Net». Основным достоинством «U-Net» является пирамидальная организация обработки, позволяющая относительно быстро анализировать широкую окрестность пикселя (радиусом порядка 100 пикселей). Было решено развить идею пирамидальной обработки, увеличив на единицу высоту пирамиды и сократив в 2 раза число сверточных слоев на каждом уровне пирамиды. Размер анализируемой окрестности при этом остается таким же, но существенно сокращается число слоев ИНС. Число каналов на каждом уровне пирамиды дополнительно сокращено в 2 раза. На нижнем уровне пирамиды, таким образом, формируется 16 каналов. При этом за счет большей высоты на верхнем уровне пирамиды число каналов оказывается лишь в 2 раза меньше, чем в классической архитектуре «U-Net». Для обработки одного изображения 512×512 пикселей с помощью модифицированной архитектуры требуется порядка 10 млрд сложений и умножений, в то время как для классической архитектуры «U-Net» – порядка

218 млрд, а для архитектуры «U-Net» с в 2 раза уменьшенным числом каналов в промежуточных слоях – порядка 55 млрд.

Модифицированная архитектура ИНС представлена на рис. 2. После модификации она напоминает более высокую и более узкую букву «U», поэтому назовем модифицированную архитектуру «Lanky U-Net». Размер пакета для «Lanky U-Net» удалось увеличить до 13 изображений. По результатам машинного обучения «Lanky U-Net» было достигнуто значение функции потерь $1-\tilde{a}$ на валидационной выборке, равное 0,0219. Таким образом, «Lanky U-Net» в задаче сегментации облачности на панхроматических изображениях аппаратуры «Геотон» космических аппаратов «Ресурс-П» обеспечивает чуть более высокое качество сегментации, чем архитектура «U-Net» с уменьшенным в 2 раза числом каналов, при в 5,5 раз меньшем объеме вычислений. Классическая архитектура «U-Net» потенциально способна обучиться до сопоставимых результатов, но лишь при развитии графических ускорителей вычислений, когда объем их памяти позволит задать приемлемый размер пакета при обучении «U-Net».

Дополнительным недостатком пакетной нормализации является небольшое увеличение объема вычислений при штатной работе ИНС из-за затрат на линейное преобразование яркостей изображения после каждого сверточного слоя. В то же время, если разместить слой нормализации не после активационной функции сверточного слоя, а между сверткой и активацией, то этих затрат при штатной работе можно избежать путем применения линейного преобразования не к свернутому изображению, а к весовым коэффициентам свертки. Если функция активации существенно нелинейная, такая перестановка положения слоев нормализации не имеет

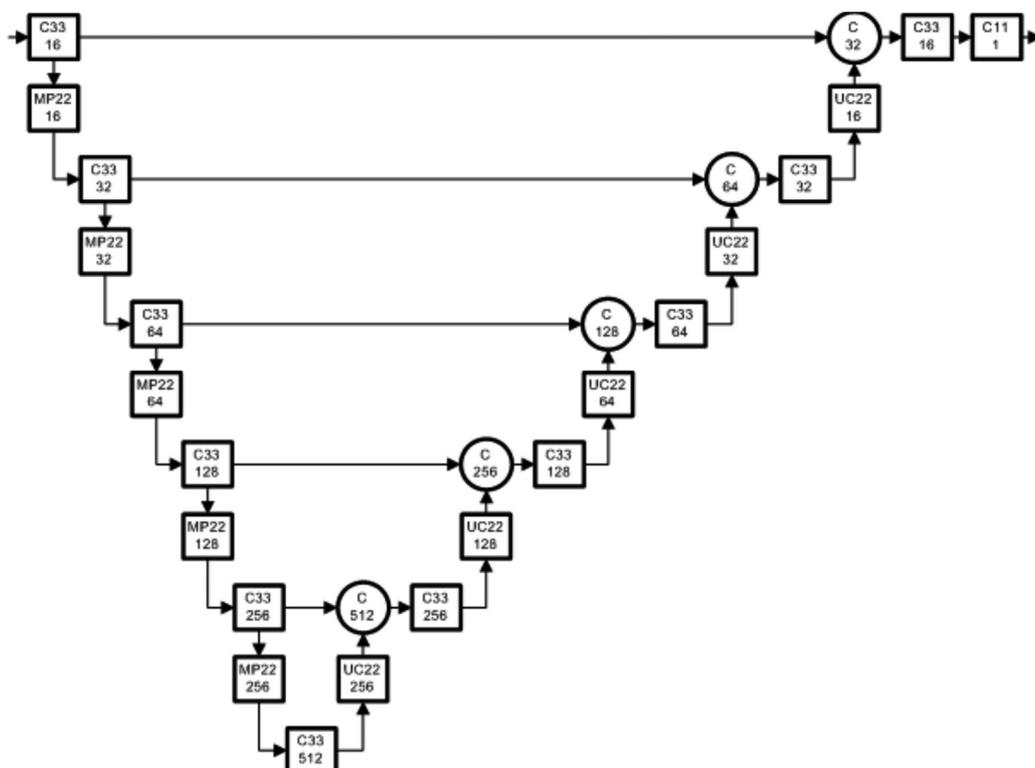
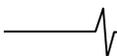


Рис. 2. Предлагаемая архитектура ИНС «Lanky U-Net»



смысла. Однако для активационной функции «линейный выпрямитель», являющейся кусочно-линейной, такая перестановка может быть допустима. Архитектура «Lanky U-Net» была обучена с измененным положением слоев пакетной нормализации. Сходимость процесса обучения от этого незначительно ухудшилась. Достигнутое значение функции потерь $1-\tilde{a}$ на валидационной выборке составило 0,0224. Таким образом, за счет перестановки слоев пакетную нормализацию можно применять при обучении ИНС без снижения скорости штатной работы, достигая при этом практически такого же качества сегментации.

Выбор уровня радиометрической обработки входных данных ИНС

В работе [10] для сегментации облачности предлагалось выполнять радиометрическую нормализацию панхроматических изображений. Космические снимки могут быть сняты с различным коэффициентом усиления и в разных условиях освещенности. Для устранения первого фактора коды яркости изображения необходимо пересчитать в единицы энергетической яркости, используя коэффициенты градуировочной характеристики съемочного устройства. Для устранения второго фактора можно предположить, что поверхность наблюдаемых объектов является горизонтально расположенной ламбертовой поверхностью, и разделить значение полученной энергетической яркости на косинус зенитного угла Солнца в момент съемки, пренебрегая поглощением и рассеянием света в атмосфере. Именно преобразованные таким образом изображения использовались для выбора и обучения архитектуры ИНС в настоящей работе.

Однако выполненная таким образом радиометрическая нормализация является приближенной.

Во-первых, для облачности, в большинстве случаев засвеченной, пересчет кодов яркости в единицы энергетической яркости является некорректным. Значения энергетической яркости получаются существенно заниженными, и облачность потенциально может быть спутана с менее яркими объектами. И если в исходных кодах яркости по абсолютному значению кода можно понять, что это область насыщения, то после преобразования получаемая энергетическая яркость может как принадлежать области насыщения при большом коэффициенте усиления, так и не принадлежать ей при малом коэффициенте.

Во-вторых, облачность не является горизонтальной ламбертовой поверхностью. Облака – это взвешенные в воздухе скопления капель воды. Рассеяние излучения этими каплями происходит не на плоской границе облака, а в некоторой его толще. Поэтому деление на косинус зенитного угла Солнца для облачности не применимо.

Таким образом, радиометрическая нормализация позволяет существенно снизить вариации яркости наземных объектов, но при этом существенно искажает яркость искомой облачности. Поэтому заранее предсказать, повышает или понижает качество сегментации облачности такая радиометрическая обработка исходного изображения, не представляется возможным. Сле-

довательно, необходимо провести соответствующие экспериментальные исследования, в которых ИНС обучаются сегментации облачности на полностью нормализованных изображениях, изображениях, пересчитанных в единицы энергетической яркости, и исходных ненормализованных изображениях.

Экспериментальные исследования показали, что ИНС «Lanky U-Net» со слоями нормализации, расположенными между сверткой и активационной функцией, обучается до следующих значений функции потерь на валидационной выборке:

- 0,0224 для полностью нормализованных изображений;
- 0,0226 для изображений, пересчитанных в единицы энергетической яркости;
- 0,0199 для исходных изображений.

Из полученных оценок можно сделать вывод, что любая нормализация в среднем снижает качество сегментации облачности. Визуальный анализ результатов сегментации показал, что на полностью или частично нормализованных изображениях плохо выделяются засвеченные облака, но достаточно неплохо отделяются от облачности другие яркие объекты: пески, снега, крыши зданий. При сегментации по исходным изображениям засвеченные облака почти не пропускаются, зато существенно учащаются случаи ложного распознавания в качестве облачности других ярких объектов. В связи с этим возникла идея подать на вход ИНС многоканальное изображение, в каналах которого содержится поразному нормализованный исходный снимок, и поручить ИНС самой определить, по какому из этих каналов ей лучше выделить облачность.

Были рассмотрены следующие варианты входных данных ИНС:

- исходное и полностью нормализованное изображение (2 канала);
- исходное и частично нормализованное (до энергетических яркостей) изображение (2 канала);
- исходное, частично нормализованное и полностью нормализованное изображения (3 канала).

В результате обучения ИНС были достигнуты следующие значения функции потерь на валидационной выборке: 0,0189 для первого варианта, 0,0181 для второго варианта и 0,0170 для третьего варианта. Видно, что добавление только полностью нормализованных изображений относительно слабо помогает лучше выделить облачность, а более существенный положительный эффект наблюдается от добавления частично нормализованных изображений. Максимальный же эффект дает добавление полностью нормализованных изображений в сочетании с частично нормализованными изображениями.

Таким образом, на вход нейронной сети наиболее целесообразно подавать в трех каналах:

- исходное изображение в кодах яркости, по которому ИНС может обнаружить засвеченные объекты;
- частично нормализованное изображение в единицах энергетической яркости, по которому ИНС наилучшим образом может обнаружить незасвеченную облачность;

– полностью нормализованное изображение, на котором минимальны возможные вариации яркости наземных объектов, что позволяет лучше отделить их от облачности.

Экспериментальные исследования

Обучающая и валидационная выборки составлены из фрагментов размером 512×512 пикселей 32 обзорных панхроматических снимков аппаратуры «Геотон» космических аппаратов «Ресурс-П» № 1-3. Снимки получены с различными коэффициентами усиления при различных зенитных углах Солнца и включают все основные типы наблюдаемых объектов: городскую застройку, водоемы, поля, леса, горы со снежными шапками, пески, засвеченную и незасвеченную сплошную облачность, а также полупрозрачную дымку. Шаг дискретизации обзорного изображения увеличен в 8 раз по сравнению с детальным и составляет около 6,5 м на земной поверхности. Это позволяет охватывать нейронной сетью в 64 раза большую площадь земной поверхности при малом объеме вычислений. За счет большого размера облаков и относительно гладкой их текстуры вся основная информация для отделения об-

лачности от других объектов на обзорном изображении сохраняется. Фрагментам, составляющим обучающую и валидационную выборки, соответствует примерно половина суммарной площади снимков, при этом объем обучающей выборки в 9 раз больше (13837 фрагментов), чем валидационной (1537 фрагментов).

Все результаты обучения, приведенные ранее, получены на описанных выборках. Сами исходные снимки составляют тестовую выборку, для которой вручную получены эталонные маски облачности. Для выполнения исследований разработан программный комплекс, реализующий рассмотренную нейросетевую сегментацию. При тестах снимки разбиваются на фрагменты с перекрытием 256 пикселей. Перекрытие выбрано так, чтобы оно было больше или равно диаметру окрестности пикселя, анализируемой ИНС, а также кратно 32. Кратность 32 обусловлена тем, что слои прореживания с выбором максимального значения инвариантны к сдвигу сигнала только на целое число своих апертур. Таким образом гарантируется отсутствие блочных артефактов. На краях тестовые изображения зеркально дополняются на 128 пикселей.

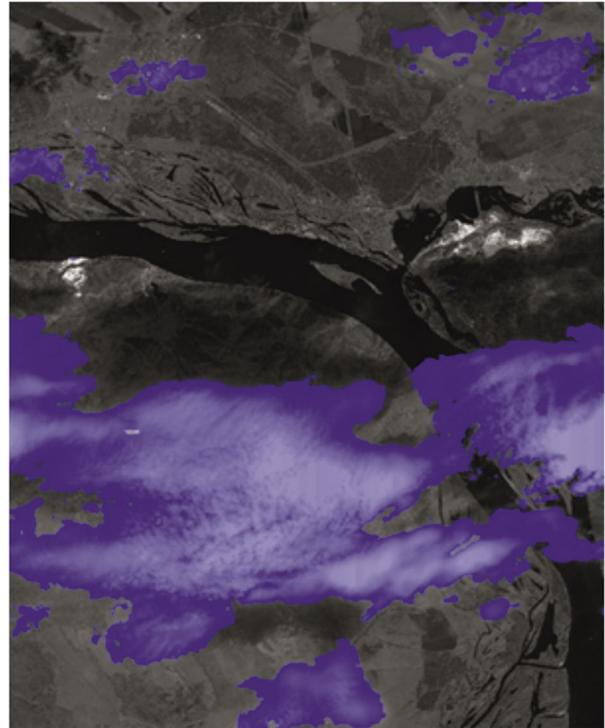
В таблице приведены значения показателей δ , α и

Показатели качества сегментации облачности на тестовых изображениях

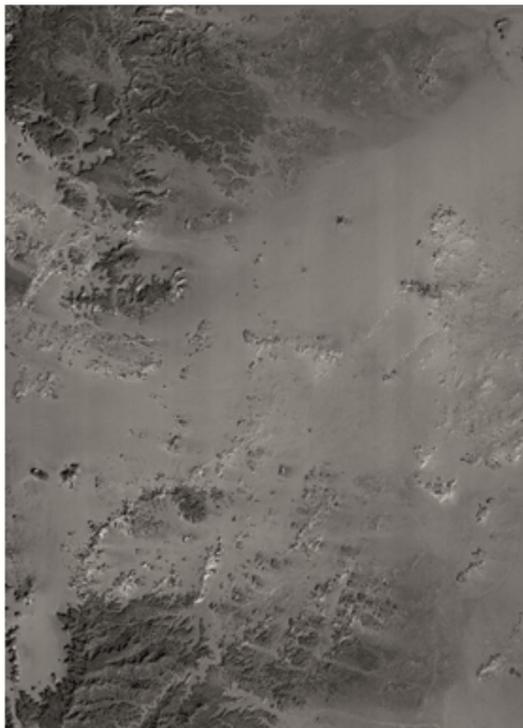
№ п/п	Ширина изображения, пикс.	Высота изображения, пикс.	δ	α	β	a
1	5625	11298	0,024	0,001	0,304	0,992
2	4500	11406	0,125	0,014	0,017	0,985
3	4500	11172	0,735	0,093	0,012	0,966
4	5625	16014	0,306	0,005	0,029	0,987
5	5625	8094	0,058	0,009	0,206	0,980
6	4500	8850	0,059	0,013	0,104	0,981
7	4500	11244	0,495	0,028	0,042	0,965
8	5625	18678	0,922	0,258	0,011	0,970
9	5625	13818	0,000	0,000	–	1,000
10	5625	13782	0,000	0,001	–	0,999
11	5625	10380	0,704	0,306	0,009	0,903
12	5625	13800	0,432	0,024	0,024	0,976
13	5625	19956	0,131	0,021	0,088	0,970
14	3375	8688	0,000	0,433	–	0,567
15	5625	20028	0,136	0,025	0,048	0,972
16	4500	10470	0,915	0,699	0,004	0,937
17	4500	8688	0,584	0,254	0,023	0,881
18	4096	8704	0,071	0,019	0,046	0,979
19	4096	9216	0,321	0,035	0,011	0,973
20	5625	6870	0,960	0,498	0,001	0,979
21	5625	7446	0,184	0,021	0,042	0,975
22	6204	4572	0,284	0,014	0,080	0,967
23	6580	10873	0,155	0,014	0,045	0,981
24	1453	4359	0,017	0,005	0,322	0,989
25	6258	34244	0,456	0,036	0,035	0,965
26	5145	17923	0,000	0,000	–	1,000
27	5469	8464	0,000	0,001	–	0,999
28	6305	30814	0,000	0,001	–	0,999
29	6289	7842	0,000	0,009	–	0,991
30	6304	10753	0,000	0,007	–	0,993
31	6338	10287	0,000	0,002	–	0,998
32	5183	8319	0,000	0,000	–	1,000
Итого			0,253	0,029	0,025	0,972



а)



б)



в)



г)



д)

Рис. 3. Примеры результатов сегментации облачности:

- а) выделение сплошной облачности и ложное обнаружение дымки на мелководном озере;
- б) корректное отделение облачности от песчаных карьеров; в) песок не выделен как облачность;
- г) наиболее яркий песок с гладкой текстурой ошибочно выделен как облачность;
- д) снежные шапки гор практически не выделяются как облачность

β для всех тестовых снимков, а также итоговые значения показателей, рассчитанные с учетом площади каждого снимка и доли облачности на нем, достигнутые нейронной сетью «Lanku U-Net» со слоями нормализации, расположенными между сверткой и активационной

функцией, на вход которой подаются в трех каналах различные виды нормализации яркостей. Итоговое значение показателя α составляет 0,972. При этом результаты сегментации оказались неплохо сбалансированы по вероятностям ошибок первого и второго рода, со-

ставляющим 0,029 и 0,025 соответственно. Скорость обработки на графическом ускорителе NVIDIA GeForce RTX 2080 Ti составила 7,4 Мпикс./с.

На рис. 3 приведены примеры сегментации облачности. Визуальный анализ показывает, что ИНС в основном удается отличить сплошную облачность от других ярких объектов: песков, крыш зданий, снегов. При этом полупрозрачная дымка выделяется плохо, и иногда за нее могут быть ошибочно приняты мелководные водоемы с просвечивающим дном и пески.

Заключение

В настоящей работе предложена архитектура ИНС «Lanky U-Net», позволяющая при в 5,5 раз меньшем объеме вычислений, чем у архитектуры «U-Net» с уменьшенным в 2 раза числом каналов, достигнуть того же качества сегментации облачности на панхроматических изображениях. Проведены исследования по выбору уровня радиометрической обработки исходных данных ИНС. Предложено подавать на вход ИНС в трех каналах: изображение, полностью нормализованное с учетом зенитного угла Солнца; изображение, частично нормализованное путем пересчета в энергетические яркости на входном зрачке съёмочного устройства; исходное изображение в кодах яркости. Достигнуты общая точность сегментации 0,972, вероятность возникновения ошибок первого рода 0,029, вероятность возникновения ошибок второго рода 0,025 и скорость обработки на графическом ускорителе NVIDIA GeForce RTX 2080 Ti 7,4 Мпикс./с.

Литература

1. Гомозов О.А., Кузнецов А.Е., Побаруев В.И., Пошехонов В.И. Программно-математическое обеспечение системы обработки космических стереоизображений. Вестник Рязанского государственного радиотехнического университета. 2009. № 27. С. 17-22.
2. Кузнецов А.Е., Пошехонов В.И. Структурно-параметрический синтез компонентов малого космического аппарата картографического назначения. Вестник Рязанского государственного радиотехнического университета. 2019. № 69. С. 185-192.

3. Кузнецов А.Е. Системы и технологии обработки аэрокосмической информации. Вестник Рязанского государственного радиотехнического университета. 2012. № 39-2. С. 7-14.

4. Ветров А.А., Кузнецов А.Е. Сегментация облачных объектов на панхроматических изображениях земной поверхности. Цифровая обработка сигналов. 2011. № 3. С. 32-36.

5. Ereemeev V.V., Kochergin A.M., Kuznetsov A.E. Automatic detection of clouds in multispectral images subjected to interchannel parallax. IEEE International Geoscience and Remote Sensing Symposium. 2015, pp. 4928-4930.

6. Li Z., Shen H., Cheng Q., Liu Y., You S., He Z. Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors. ISPRS J. of Photogram. and Rem. Sen. 2019, vol. 150, pp. 197-212.

7. Jiao L., Huo L., Hu C., Tang P. Refined UNet: UNet-Based Refinement Network for Cloud and Shadow Precise Segmentation // Remote Sensing. 2020, vol. 12(12).

8. Mohajerani S., Saeedi P. Cloud-Net: An End-to-End Cloud Detection Algorithm for Landsat 8 Imagery. IEEE International Geoscience and Remote Sensing Symposium. 2019, pp. 1029-1032.

9. Mohajerani S., Saeedi P. Cloud and Cloud Shadow Segmentation for Remote Sensing Imagery via Filtered Jacard Loss Function and Parametric Augmentation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2021, vol. 14, pp. 4254-4266.

10. Ereemeev V., Kuznetsov A., Kochergin A., Makarenkov A. Clouds segmentation on panchromatic high spatial resolution remote sensing images using convolutional neural networks. Proceedings of the SPIE. 2019, vol. 11155.

11. Ronneberger O., Fischer P. and Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention. 2015, pp. 234-241.

12. Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Proceedings of the 32nd International Conference on International Conference on Machine Learning. 2015, vol. 37, pp. 448-456.

НОВЫЕ КНИГИ

Шарамет А.В.

Информационное обеспечение систем защиты летательных аппаратов от управляющих средств поражения: Научное издание, М.: Изд-во «Горячая линия-Телеком», 2023 г. 178 с.: ил.

Рассмотрены вопросы повышения информативности бортового комплекса обороны летательного аппарата для увеличения эффективности его защиты от управляемого ракетного оружия за счет согласованного помехового, маневренного и огневого противодействия. Изложены и систематизированы основные проблемы по защите летательного аппарата от управляемых средств поражения, а также рассмотрены основные пути повышения эффективности его защиты от них. Даны оценочные расчеты отражательных свойств головок самонаведения управляемых ракет класса «воздух-воздух» и «земля-воздух», как объектов радиолокационного наблюдения. Представлено описание и проведен анализ результатов полунатурного экспериментального исследования. На основе векторно-алгебраического подхода к формализации задач определения координат объектов в многопозиционной измерительной системе проведен синтез алгоритма оценки декартовых координат атакующей ракеты по суммарно-дальномерной информации.



Для специалистов, занимающихся вопросами защиты летательных аппаратов, научных работников и инженеров. Может быть полезна аспирантам и студентам вузов.