Оценка поля прозрачности облачности с использованием оптических и радарных изображений дистанционного зондирования Земли

В.Ф. Коновалов [1], В.В. Мясников [1], В.В. Сергеев [1]
[1] Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Россия, г. Самара, Московское шоссе, д. 34

Аннотация

В работе предложен метод оценки поля прозрачности облачности с использованием фундаментальной нейросетевой модели TerraMind, использующей в качестве мультимодальных данных территории съемки мультиспектральные изображения и изображения радара с синтезированной апертурой. Исходными данными для обучения модели TerraMind выступал набор SEN12MSCR, содержащий пары изображений спутников Sentinel-1 и Sentinel-2 как с облачностью, так и без нее. Для дообучения фундаментальной модели с целью решения искомой задачи мы использовали аугментацию, синтезируя облачность и соответствующие ей истинные поля прозрачности с использованием Satellite Cloud Generation Toolbox. Численные результаты проведенных экспериментальных исследования показывают высокие качественные показатели предложенного метода оценки, а качественные результаты оценки полей прозрачности на натурных снимках подтверждают работоспособность метода на реальных данных.

Ключевые слова: Sentinel 2, Sentinel 1, радар с синтезированной апертурой, мультиспектральные изображения, нейронные сети, устранение облачности, устранение дымки, синтетический набор данных, TerraMind.

Цитирование: **Коновалов, В.Ф.** Оценка поля прозрачности облачности с использованием оптических и радарных изображений дистанционного зондирования Земли / В.Ф. Коновалов, В.В. Мясников, В.В. Сергеев // Компьютерная оптика. – 2026. – Т. 50, № 5. – 2005. – doi: 10.18287/COJ2005.

Citation: Konovalov VF, Myasnikov VV, Sergeev VV. A method for cloud soft segmentation using multimodal remote sensing data. Computer Optics 2026; 50(5): 2005. doi: 10.18287/COJ2005.

Введение

Большое количество находящихся на орбите космических аппаратов позволяет получать данные дистанционного зондирования Земли (ДЗЗ) различной физической природы. К таким данным в первую очередь относятся широко используемые снимки радара с синтезированной апертурой (РСА) и мультиспектральные изображения. Периодический характер получения этих снимков позволяет оценивать изменения, происходящие на земной поверхности, детектировать и распознавать различные объекты и/или ситуации. Однако зачастую возникают ситуации, когда съемка оказывается либо невозможной в силу технических причин (положение спутника, конфликт между несколькими одновременными съемками и т.п.), либо результат съемки не может быть использован по объективным причинам.

К числу таких причин, мешающих использованию оптических (RGB, мультиспектральных, гиперспектральных) изображений, являются *облачность* и *дымка/туман* (англ.: haze). Известным фактом является то, что в любой момент времени порядка 70 % территории Земли оказывается покрыто облачностью [1], при этом указанная доля ниже на экваторе и выше к полюсам. В противоположность оптическим снимкам, РСА-снимки, в силу активного характера формирования изображений, свободны не только от проблем с облачностью или дымкой, но позволяют получать снимки даже в темное время суток. Как было ранее показано в работе авторов [2], при наличии одновременно РСА и оптических изображений целые фрагменты последних могут быть реконструированы с использованием мультимодальных нейронных сетей с высоким качеством: исследование проводилось на наборе данных SEN12MSCR [3] для областей реконструкции размером до 30 %; триплет показателей качества PSNR, MAE, SSIM для лучших полученных архитектурных решений принимал значения в диапазоне (37.12, 0.0057, 0.9786) – (39.45, 0.0033, 0.9852). Указанный факт делает вроде бы очевидным и способ удаления облачности/дымки на оптическом снимке путем детектирования/сегментации облаков и последующей реконструкции оптического изображения с использованием указанного выше способа с использованием РСА-снимков. Однако задача определения «границ» или «области» облака/дымки на снимке оказывается и формально (математически), и неформально некорректной, поскольку облако/дымка в общем случае не является непрозрачным объектом. То есть в зависимости от «плотности» облака/дымки на оптическим снимке может быть и видна, и интерпретируема (как интерактивно, так и автоматически) территория земной поверхности, которая находится «под облаком» – см. рис. 1. В результате, в зависимости от последующего процесса обработки и используемых инструментов (методов, алгоритмов и т.п.) одни и те же области оптических снимков могут как требовать, так и не требовать реконструкции в зависимости от степени «прозрачности» облачного покрова и/или

дымки, присутствующей на снимке. Более того, даже при наличии определенности с допустимой/граничной величиной «прозрачности» существует объективная сложность с оценкой этой величины, поскольку на реальных оптических снимках данные о «прозрачности» нигде не представлены, а одновременное получение оптических снимков с облачностью/дымкой и без нее является абсолютно невозможным. Представленные выше доводы делают *задачу оценки поля прозрачности облачности достаточно сложной и актуальной*.

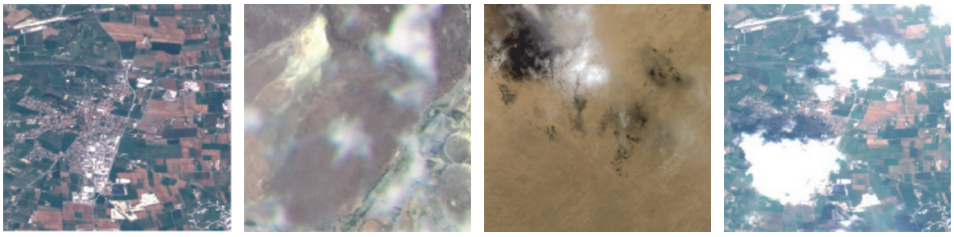


Рис. 1. Примеры изображений ДЗЗ. Слева направо: снимок без облачности, два снимка с примерами «прозрачной» облачности (дымкой), снимок с «непрозрачной» облачностью

Способы постановки и решения указанной задачи исторически менялись в зависимости как от доступных данных ДЗЗ, так и доступных математических методов и подходов. Одними из первых подходов к оценке (бинарного) поля прозрачности можно считать визуальное дешифрирование (интерпретацию) снимков экспертом и простейшие пороговые методы. Последние основывались на анализе отсчетов яркости в одном или нескольких каналах и/или отсчетов полей индексов, вычисляемых как отношение видимых (VIS) и инфракрасных (NIR/SWIR) каналов [4, 5], например, индекса снега NDSI (англ.: Normalized Difference Snow Index) [6]. Автоматические подходы впоследствии породили целое направление, обозначаемое в современной научной литературе как *обнаружение облачности* (англ.: cloud detection) [7 – 10]. Результатом методов этого класса является бинарное или мультиклассовое [11 – 13] поле прозрачности, где один из классов отвечает за видимость (или дискретную степень видимости) земной поверхности, а другой – присутствие облака. По мере улучшения разрешения оптических снимков ДЗЗ, использования низкоорбитальных (в т.ч. атмосферных) спутников и развития средств машинного обучения и искусственного интеллекта задача детектирования с бинарной маской постепенно трансформировалась в более сложные и, вообще говоря, связные задачи *устранения облачности* (англ.: cloud removal) [14] и *устранения дымки* (англ.: image dehazing) [15]. Проблема оценки поля прозрачности облачности является для этих задач внутренней и в ряде случаев явно не ставится, особенно при использовании генеративных подходов к синтезу «безоблачного» снимка по его прототипу с «облаком».

В данной работе мы предлагаем явно декомпозировать проблему устранения облачности/дымки на две последовательно решаемые: проблему оценки поля прозрачности облачности и проблему реконструкции оптического изображения в тех частях изображения, где величина прозрачности ниже определенного порога. Причины такого подхода (наличие высококачественного решения реконструкции фрагментов оптических снимков по их совместным данным с РСА-изображениями) мы изложили выше. Для решения проблемы оценки поля прозрачности облачности/дымки мы предлагаем использовать фундаментальную модель TerraMind, настроенную на специально сформированном для нашей задачи наборе данных.

Основным вкладом настоящей работы являются:

- впервые фундаментальная нейросетевая модель TerraMind использована для оценки поля прозрачности облачности по данным оптических и РСА-сенсоров;
- впервые по совместным данным оптических и РСА-сенсоров получена оценка поля прозрачности облачности;
- численно продемонстрировано преимущество по качественным показателям предложенного решения по сравнению с известными решениями;
- предложенное решение, используемое совместно с разработанными ранее авторами решениями по реконструкции фрагментов оптических снимков по данным оптических и РСА-сенсоров, позволяет достичь рекордных показателей по качеству реконструируемых изображений с устраненной облачностью.

Работа построена следующим образом. В первом параграфе дан обзор современного состояния исследований по теме работы, включающий как обзор существующих решений по оценке поля прозрачности облачности, так и более общей задачи обнаружения и устранения облачности/дымки. Во втором параграфе вначале кратко представлена общая схема устранения облачности на оптических изображениях с использованием данных РСА и оптического сенсора, использующая разработанное нами ранее решение по реконструкции оптических изображений. И более детально рассмотрена архитектура основного предлагаемого в настоящей работе решения – метода оценки поля прозрачности облачности с использованием фундаментальной нейросетевой модели TerraMind, использующей в качестве мультимодальных данных территории съемки мультиспектральные изображения и изображения РСА. Третий параграф содержит результаты экспериментальных исследований предложенного метода, содержащие как оценки показателей качества на модельных данных, так и качественные результаты на реальных данных.

1. Обзор литературы

Следуя приведенной во введении классификации задач, связанных с облачностью на снимках ДЗЗ, ниже мы представляем краткий обзор подходов, методов и алгоритмов, направленных на решение как непосредственно задачи оценки поля прозрачности облачности, так и более широкого класса задач: обнаружения облачности, устранения облачности и дымки. Дополнительно дан обзор по существующим актуальным наборам данных.

Обнаружение/распознавание облачности

Достаточно подробные обзоры литературы за последние десять-пятнадцать лет представлены в работах [7 – 10]. Типичным является деление существующих подходов на классические методы и методы с обучением, последние дополнительно подразделяются на классические методы с обучением, нейросетевые и методы с глубоким обучением. Также методы обнаружения подразделяются по используемым данным — это могут быть как монотемпоральные, так и мультитемпоральные данные, а также использующие данные одного или нескольких сенсоров, различных по природе. Использование мультитемпоральных данных неизбежно приводит к необходимости учитывать и игнорировать различия в снимках, связанные с сезонностью и другими изменениями, не связанными с облачностью [3], в то время как использование разнородных данных требует специальной обработки данных для объединения признаков из разных модальностей. В некоторых работах по распознаванию облачности также вводится и аналог поля прозрачности – облачность подразделяется на классы по плотности: дымка, полупрозрачные облака, плотные облака [16]. В ряде работ выделяются дополнительные классы как схожие с целевым классом, такие как снег [17], так и непосредственно с ним связанные, такие как тени от облаков [16]. В последнее время фокус смещается на разработку компактных и энергоэффективных решений.

Наряду с очевидными и ожидаемыми выводами о преимуществах методов глубокого обучения в задачах обнаружения/распознавания облачности (наряду с возможностью использования и полностью неспецифических методов детектирвания, таких как yolov8 [18] или SAM [19]), обзоры подчеркивают *существующие проблемы* в данной области, в частности *сложность получения корректно и точно размеченных наборов данных*, слабость кросс-модальных методов, а также необходимость переобучать сеть под различные виды сенсоров [7 – 9, 20].

Устранение облачности

Основная задача *устранения облачности* заключается в восстановлении перекрытых ею участков снимков ДЗЗ без нарушения их пространственной и спектральной целостности. Достаточно подробный разбор различных подходов к решению этой задачи представлен, например, в обзоре [14]. Так же как и в задаче обнаружения облачности, здесь также подходы подразделяются на традиционные, то есть использующие физические модели, и методы, использующие машинное обучение и нейросети. Традиционные подходы опираются на модели атмосферной коррекции и спектрального анализа для выявления и устранения облачного покрова. Наиболее простой из них является следующая модель, принимаемая во многих работах (и в настоящей работе тоже) [14]:

$$I(x) = J(x)t(x)+A(1 - t(x)). \quad (1)$$

Здесь x – пространственная координата, $t(x)$ – поле прозрачности (карту пропускания) облачности на снимке, A – глобальное атмосферное освещение, $J(x)$ и $I(x)$ – искомое «безоблачное» и наблюдаемое изображения. Несмотря на эффективность в отношении «тонкой» облачности, традиционные подходы оказываются малоэффективны при работе со сложными облачными формациями и переменчивыми атмосферными условиями, зачастую не обеспечивая качественного восстановления деталей закрытой поверхности.

Более перспективными в настоящее время выглядят методы глубокого обучения, которые извлекают закономерности данных из больших массивов снимков [9, 10, 14, 21 – 22]. Для повышения точности процесса применяются сверточные нейросети (CNN), генеративно-состязательные сети (GAN) и вариационные автокодировщики (VAE). Однако таким моделям часто не хватает гибкости при работе с разными типами облачности, и они плохо обобщают знания на задачи с ограниченным объемом обучающих данных, что снижает их надежность в практическом применении [14].

Мультимодальный подход предполагает совместное использование оптических данных и РСА-снимков, что позволяет задействовать сильные стороны разных типов съемки [3, 14, 22 – 24]. Способность радара проникать сквозь облачность повышает общую эффективность подхода, однако острой проблемой для широкого внедрения таких методов (и подготовки соответствующих наборов обучения) остается необходимость в идеальной синхронизации и геометрической привязке данных.

В последние годы генеративные модели (и, в частности, диффузионные) зарекомендовали себя как мощный инструмент восстановления изображений благодаря своей способности к итеративному уточнению [9, 10, 14, 20 – 22]. Они позволяют постепенно реконструировать скрытые или утраченные фрагменты, используя процесс контролируемого зашумления.

Следует особо отметить, что разработанный авторами высококачественный алгоритм реконструкции фрагментов оптических снимков по их совместным с РСА-изображениями данным также использует

мультимодальность и диффузионную нейросетевую модель, то есть отвечает самым современным тенденциям в обсуждаемом научном направлении.

Устранение дымки

Устранение дымки (англ.: image dehazing) [15, 25 – 32] – задача, во многом схожая с задачей устранения «тонкой» облачности, то есть допускающая, что поле прозрачности (карта пропускания) не имеет (или почти не имеет) «непрозрачных» отсчетов. В результате используемые подходы, методы, решения и тенденции во многом коррелируют с указанными в предыдущем подпараграфе «Устранение облачности»: традиционные подходы [27, 32], подходы на основе глубокого обучения [29 – 31], генеративные подходы [28], мультимодальность путем интеграции с РСА-снимками [30, 32]. Для подтверждения этого утверждения на рис. 2 представлена классификация методов устранения дымки, данная в обзорной работе [25], которая содержит очень подробный анализ данного вопроса с таксономией подходов и методов устранения дымки по нескольким критериям.



Рис. 2. Классификация методов устранения дымки (из материалов работы [15])

Среди *возможных направлений дальнейших исследований* авторы указанного обзора, опубликованного в текущем 2026 году, среди прочих указали следующие:

- *объединение информации с нескольких датчиков* (оптических, РСА и др.), что позволит улучшить эффективность удаления дымки в различных атмосферных условиях;
- *применение самостоятельного и неконтролируемого обучения*, которое позволяет использовать крупномасштабные неаннотированные наборы данных.

Настоящая работа непосредственно связана с указанными направлениями, поскольку:

- использует для построения поля прозрачности облачности наряду с оптическими и РСА-снимками, а также
- применяет предварительно (само-)обученную на неаннотированных данных фундаментальную модель TerraMind [33] в качестве остова конструируемого решения, который настраивается (англ.: finetuning) на решение искомой задачи.

Наборы данных

Основные наборы данных, используемые для обучения и сравнения методов/алгоритмов решения задач обнаружения и удаления облачности и дымки, представлены в табл. 1. Как можно видеть, существует всего пять открытых наборов, содержащих одновременно РСА и оптические снимки. Так, набор данных cloudSEN12 [16] представляет собой очень большой набор мультимодальных данных, включающий мультиспектральные и РСА-снимки, а также маски сегментации облачности различной точности – как попиксельной, так и «набросочной», небрежной, для 4 классов – плотная облачность, неплотная облачность, тень от облаков, «чистая» область, а также маски, полученные различными классическими методами сегментации облачности, такими как FMask. Этот набор является на данный момент наиболее полным и детально размеченным, а также одним из немногих мультимодальных наборов данных для этой задачи. Однако он не решает проблему с разделением и разметкой различных классов облачности. SEN12MSCR [3] представляет собой набор данных, содержащий мультиспектральные снимки с облачностью, соответствующие им снимки без облачности и радарные снимки. SEN12MSCR-TS [34] добавляет к этому набору данных дополнительные облачные снимки, образуя мультитемпоральный набор данных. Однако эти наборы данных не содержат размеченных масок сегментации облачности. Некоторые наборы данных [35] содержат только данные одной модальности, другие обладают недостаточным для современных методов количеством снимков.

Отдельно следует выделить синтетические наборы данных. Синтезирование реалистичных облаков хорошо исследовано, и они могут быть легко добавлены на изображения, так как находятся выше ландшафта. Синтетические наборы данных могут помочь решить указанные выше недостатки существующих наборов данных. Представленный в таблице набор данных RS-Haze использует данные, полученные из cirrus канала

(показывающего облачность), для синтеза облачных изображений из чистых. Для синтеза облачных изображений «с нуля» существуют специальные инструменты. Например, Satellite Cloud Generation Toolbox [38] предоставляет обширные возможности по генерации облачности, включая дополнительные параметры, такие как смещение облаков друг относительно друга в различных каналах, изменение плотности и вида генерируемых облаков, подстраиваемость под статистические параметры изображения. Синтезированные маски облачности устраняют проблему разметки, поскольку теперь задача может быть смещена из области сегментации в область попиксельной регрессии с присвоением каждому пикселю значения его прозрачности для сенсора. Synthinel1 [39] позволяет синтезировать радарные данные для задачи сегментации.

Табл. 1. Сравнение основных наборов данных (по материалам работы 2026 года [15])

Набор данных	Модальности	Пиксельное разрешение	Количество каналов	Число изображений	Синтетические данные
RS-Haze [35]	Мультиспектральный	512×512	7	52000	Только синтетические
SEN12MS [36]	Мультиспектральный + PCA	256×256	13	120000	нет
SEN12MS-CR [3]	Мультиспектральный + PCA	256×256	13	120000	нет
SEN12MS-CRTS [34]	Мультиспектральный + PCA	256×256	13	120000	нет
CloudSEN12 [17]	Мультиспектральный + PCA	509×509	13	49400	нет
SateHaze1k [28]	RGB+PCA	512×512	3	1000	Только синтетические
MRSHaze [37]	RGB+PCA	512×512	3	8000	есть

В представленном исследовании мы используем как реальные, так и синтетические данные облачности: последние используются для обучения и расчета числовых характеристик показателей качества разрабатываемого решения. А реальные данные с облачностью позволяют качественно верифицировать получаемый результат обработки. Результатом работы модели является явное представление непрерывного поля прозрачности, что отличает наш метод от существующих. Непрерывное поле прозрачности одержит больше информации об облачности по сравнению с бинарной и даже мультиклассовой маской и решает проблему интерпретации понятий дымки и плотной облачности, их разграничения.

2. Предлагаемый метод

В настоящем параграфе мы вначале кратко представим общую схему устранения облачности на оптических изображениях с использованием данных PCA и оптического сенсора, использующую разработанное нами ранее решение по реконструкции оптических изображений [2], после чего детально представим архитектуру основного предлагаемого в настоящей работе решения – метода оценки поля прозрачности облачности.

Общая схема устранения облачности на оптических изображениях

Как было указано во введении и показано в работе авторов [2], при наличии одновременно PCA и оптических изображений целые фрагменты последних могут быть реконструированы с использованием мультимодальных нейронных сетей с высоким качеством: исследование проводилось на наборе данных SEN12MSCR [3] для областей реконструкции размером до 30 %; триплет показателей качества PSNR, MAE, SSIM для лучших полученных архитектурных решений принимал значения в диапазоне (37,12, 0,0057, 0,9786) – (39,45, 0,0033, 0,9852). Также следует напомнить тот факт, отмеченный во введении настоящей работы, что в зависимости от «плотности» облака/дымки на оптическом снимке может быть и видна, и интерпретируема (как интерактивно, так и автоматически) территория земной поверхности, которая находится «под облаком». При этом интерпретируемость оказывается напрямую зависимой от последующего процесса обработки и используемых инструментов (методов, алгоритмов и т.п.). Поскольку используемые для последующей интерпретации снимка методы, алгоритмы и решения не входят в круг рассматриваемых в настоящей работе вопросов, мы полагаем, что допустимый для них уровень прозрачности облачности определяется и задается «извне». Таким образом, входными данными общей процедуры устранения облачности являются пара $I(x)$, $R(x)$ координатно-согласованных изображений оптического и PCA-сенсоров, а также пороговое значение threshold для оцененного поля прозрачности облачности. Обозначив ниже разработанную в [2] процедуру реконструкции фрагментов как Restore(…), которая принимает на вход оптическое и PCA-изображения, как $M(x)$ – «бинарную» маску облачности и как ETM(…) – разрабатываемый метод оценки поля (изображения) прозрачности $T(x)$ облачности, принимающий на вход оптическое и PCA-изображения, общая схема устранения облачности на оптическом изображении в виде псевдокода выглядит следующим образом:

T = ETM(I, R) // получение поля (изображения) прозрачности T
$M = T >$ threshold // M – маска (изображение) облачности, «1» – облако, «0» – истинное значение
J = Restore(I, R, M)

Метод оценки поля прозрачности облачности

Ключевым элементом метода оценки поля прозрачности облачности (выше обозначался как ETM) выступает фундаментальная нейросетевая модель TerraMind [33], разработанная в 2025 году рядом крупных компаний и лабораторий, включая IBM Research, European Space Agency и NASA. Это одна из самых современных и мощных фундаментальных моделей, обладающая множеством инструментов и предварительно (само-)обученная на значительном объеме данных практически со всего земного шара и поддерживающая работу с 9 различными модальностями, включая оптические и мультиспектральные изображения, цифровую карту высот, вегетационный индекс, маски сегментации/классификации, текстовые описания и координаты.

Отметим, что ключевой особенностью всех фундаментальных моделей является их способ обучения, который выполняется в два этапа. На этапе предварительного обучения модель самообучается без учителя, исключительно на неразмеченных данных/изображениях, обучаясь реконструировать патчи/токены одной модальности по данным другой(их). На втором этапе предварительно обученная модель может настраиваться на решение определенных задач путем ее (до)обучения (в ряде случаев происходит «надстройка» над исходной моделью). Такой двухэтапный подход позволяет на предварительном этапе крупным разработчикам/организациям обучать фундаментальные модели на таких объемах данных, которые для обычных не доступны по стоимости, аппаратным и временным ресурсам. Тем самым достигается максимум качества работы основной части фундаментальной модели. Процесс же дополнительного обучения и адаптации (англ.: finetuning) фундаментальной модели на втором этапе для решения конкретной задачи оказывается доступен уже для широкого круга исследователей и разработчиков.

В случае модели TerraMind на этапе предварительного обучения ее разработчиками она обучалась реконструировать оптические данные по радарным и наоборот (другие модальности этой модели здесь не рассматриваем) [34]. В результате обученный разработчиками энкодер модели TerraMind [33] выступал в нашей работе как готовый остов предлагаемого решения. А декодер модели (до)обучается нами для решения искомой задачи – оценки поля прозрачности облачности. То есть результатом на выходе декодера выступало изображение с размерами, эквивалентными входному, с вещественными значениями из интервала [0,1], интерпретируемыми как степень прозрачности. Иллюстрация нашего метода в случае его применения (не обучения) представлена на рис. 3 ниже, обучаемый нами декодер выделен жирным.

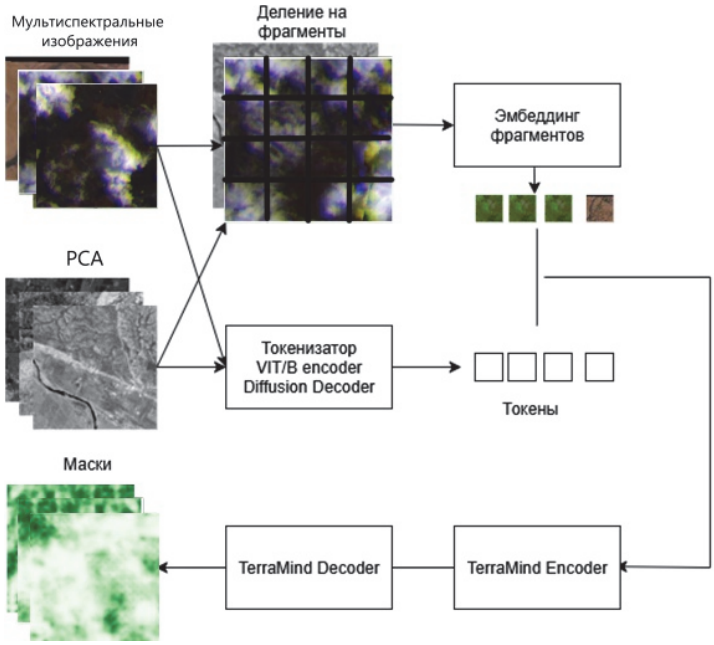


Рис. 3. Метод оценки поля прозрачности облачности с использованием фундаментальной модели TerraMind

В качестве дополнительного комментария об архитектуре используемой фундаментальной модели TerraMind следует отметить, что она построена по схеме симметричного энкодера-декодера на основе архитектуры трансформер. Это позволяет ей одновременно работать с различными типами данных, включая, например, текст и географические координаты. Для различных типов данных предусмотрены различные энкодеры, работающие с текстом, масками сегментации и изображениями по-разному. Особенностью архитектуры является

представление патчей изображений на уровне пикселей и на уровне токенов. Представление на базе токенов используется для обработки в трансформере – для извлечения более высокоуровневого концепта, а пиксельное представление позволяет работать с мелкими деталями, специфичными к данному изображению. Модель использует раннее слияние признаков в энкодере, то есть токенизированные и попиксельные признаки совмещаются на первых, имеющих наиболее высокое разрешение, слоях. Токенизатор для модальностей изображений построен на базе ViT-B энкодера [40], превращающего каждый фрагмент размером 16×16 в его скрытое векторное представление с последующей квантизацией с помощью алгоритма FSQ (англ.: finite-scalar-quantization) [41]. Декодирование этих модальностей производится с помощью диффузионной архитектуры. Именно токенизатор обеспечивает возможность межмодального взаимодействия, что особенно важно при интеграции изображений и текста.

Обучение модели для оценки поля прозрачности облачности

Для настройки модели TerraMind для решения нашей задачи мы минимизируем величину разности между требуемым и получаемым изображениями поля прозрачности облачности. А именно, если $T(x)$ – оцененное поле прозрачности, а $T^*(x)$ – истинное поле прозрачности, то критерием обучения выступает следующая величина средней абсолютной разницы (MAE – Mean Absolute Error, N – число изображений в выборке):

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\sum_{x}|T_i(x) - T_i^*(x)|. \quad (2)$$

Конструируемое преобразование, обозначенное выше ETM и реализуемое с использованием фундаментальной модели TerraMind, определяется набором параметров θ фундаментальной нейронной сети. Таким образом, цель заключается в нахождении оптимального набора параметров θ, обеспечивающего наиболее точное в смысле указанного критерия (2) совпадение оценки поля прозрачности с его истинными значениями.

Поскольку истинных полей прозрачности облачности по реальным данным съемки получить невозможно, мы для безоблачных изображений синтезируем поле прозрачности облачности, накладывая ее с учетом физической модели (1) на реальные оптические снимки. Для этого мы использовали метод синтеза искусственной облачности, описанный в работе [38]. Он позволяет синтезировать и реалистичные поля прозрачности, и соответствующие им изображения облачности, различающиеся как по плотности, так и по ее локализации на снимке.

Подготовка набора обучающих данных

Для решения задачи за основу был взят набор данных SEN12MS-CR [3]. Он содержит мультиспектральные, а следовательно, и оптические изображения с облачностью, и соответствующие им PCA-изображения, а также чистые изображения без облаков. Размеры изображений – 256×256, указанные фрагменты полностью совмещены покоординатно. Изображения обладают различным пространственным разрешением, меняющимся от 60 м/отсчет до 10 м / отсчет, оптические каналы обладают более высоким разрешением. Количество оптических каналов равно 13. Изображения были получены со спутников Sentinel 1 и Sentinel 2 и размещены в открытом доступе. Каналы мультиспектрального изображения можно поделить на оптические, инфракрасные и каналы с информацией о водяном паре (ультрафиолетовые и т.д.), каналы изображения PCA – с поляризацией VV и VH.

Из содержащихся в данном наборе чистых, безоблачных мультиспектральных (оптических) изображений путем синтеза с помощью программы Satellite Cloud Generator были получены изображения с облачностью и соответствующие поля прозрачности (см. рис. 4). Генератор был настроен так, чтобы получать различные варианты изображений с облаками (изображения с локализованной облачностью, с плотной облачностью, дымкой, частичной дымкой и т.д.), а также незначительно модифицировать цвет облачности исходя из оттенков исходного изображения.

В результате были подготовлен синтетически сгенерированный набора данных, состоящий из 120 000 триплетов «облачное изображение» – «изображение PCA» – поле прозрачности облачности. Полученный набор был поделен на тренировочную и тестовую выборку согласно разделению, предложенному в SEN12MS [36].

Метрики оценки результатов работы метода

Для оценки качества полученной оценки поля прозрачности облачности мы использовали критерии MAE, MSE (RMSE), R Squared (R2 Score). Использование таких показателей, как SSIM, LPIPS или PSNR, обычно применяемое для оценки изображений, в данном случае представляется нецелесообразным, поскольку важна величина попиксельного различия с идеальным полем прозрачности. Дадим краткую характеристику используемых показателей качества ниже.

Величина MAE, определенная выше в (2), используется и на этапе обучения как критерий качества и на этапе оценки решения как его метрика. Показатель MSE (англ.: Mean Squared Error) – среднеквадратическая ошибка, характеризующая средний квадрат отклонения истинного и оцененного поля прозрачности, задается выражением:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\sum_{x}(T_i(x) - T_i^*(x))^2.$$

Величина *R2 Score* позволяет оценить качество модели относительно «наивной» модели, которая предсказывает среднее значение пикселя в каждой точке изображения.

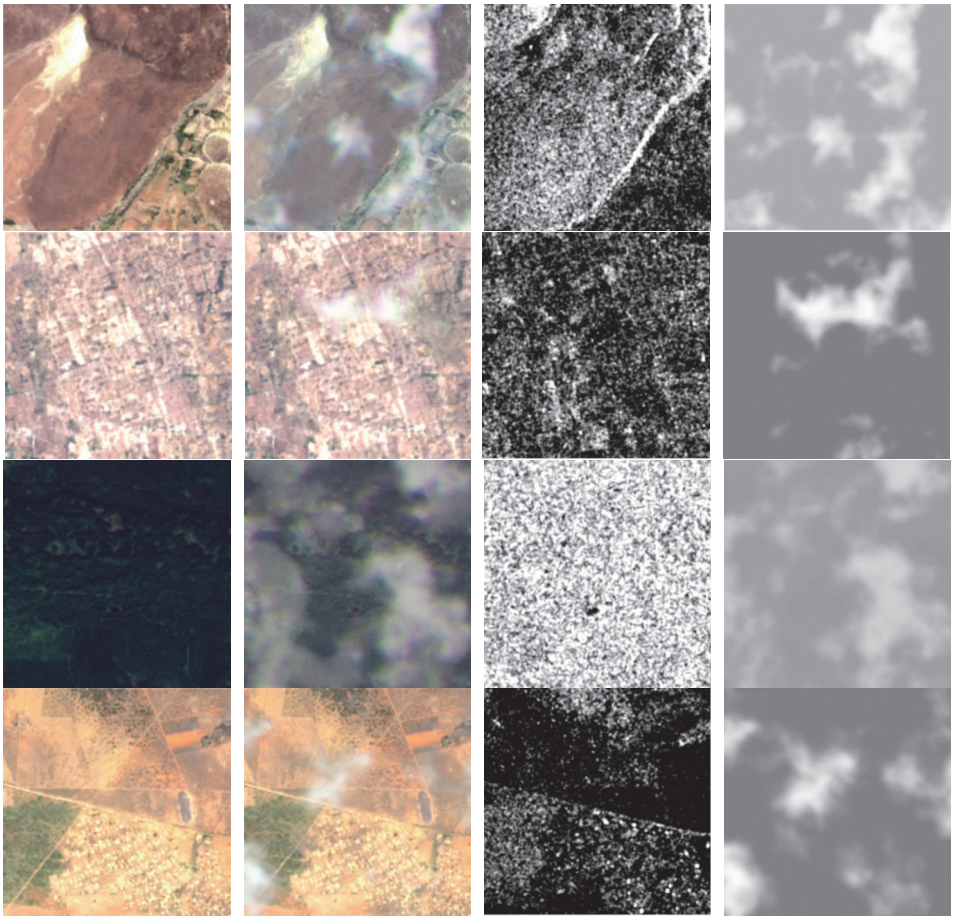


Рис. 4. Примеры обучающих данных. Слева направо: оптическое изображение, оптическое изображение с синтезированной облачностью, изображение PCA, «истинное» поле прозрачности облачности

3. Экспериментальное исследование

В процессе проведения первого эксперимента были (до)обучены две фундаментальные модели типа TerraMind. Одна из них основывается лишь на комбинации RGB-каналов Sentinel 2 и PCA-изображения, другая использует все доступные каналы мультиспектрального снимка, кроме B10 [36].

Обучение производилось в течение 50 эпох с использованием оптимизатора AdamW и скорости обучения 0,00005, в качестве функции потерь использовалась MAE (2). В качестве предобученной модели выступала terramind_base [34]. Обучение производилось с использованием RTX 4090 и процессора Ryzen 5950x. Время обучения составило около 12 часов. По окончании обучения модели вышли на «плато» функции потерь. Таким образом, модель может обучаться в приемлемые сроки на оборудовании потребительского уровня.

Результаты обучения предлагаемого метода при проверке на тестовой выборке по показателям качества представлены в табл. 2. В ней представлены средние значения показателей качества и их среднеквадратические отклонения по всему набору данных.

Табл. 2. Результаты оценки качества предлагаемого метода на тестовой выборке

Метрика	TerraMind S2 + S1	TerraMind S2 Multispectral + S1
MAE	0,0346±0,0094	0,0242±0,0070
MSE	0,0034±0,0026	0,0016±0,0015
R2 Score	0,9549±0,0363	0,9785±0,0198

Из представленных данных видно, что ошибка оценки поля прозрачности облачности мала для обоих случаев, однако для случая использования дополнительных мультиспектральных каналов качество получаемого решения выше по всем показателям. Эта модель также лучше генерализуется к реальным данным, то есть лучше показывает себя при использовании на реальных облачных изображениях.

Для модели TerraMind RGB + S1 и синтетического датасета облаков получаемые результаты и их качество могут быть оценены визуально по данным рис. 5. На этом рисунке первое изображение представляет собой входное изображение с облачностью, второе изображение – синтезированное «истинное» поле прозрачности, третье – полученная оценка поля прозрачности, четвертый столбец – изображение разности между оцененным и истинным полем прозрачности. Интенсивность зеленого цвета указывает на степень прозрачности облачности, белый цвет означает полностью прозрачные участки, темно-зеленый – непрозрачные.

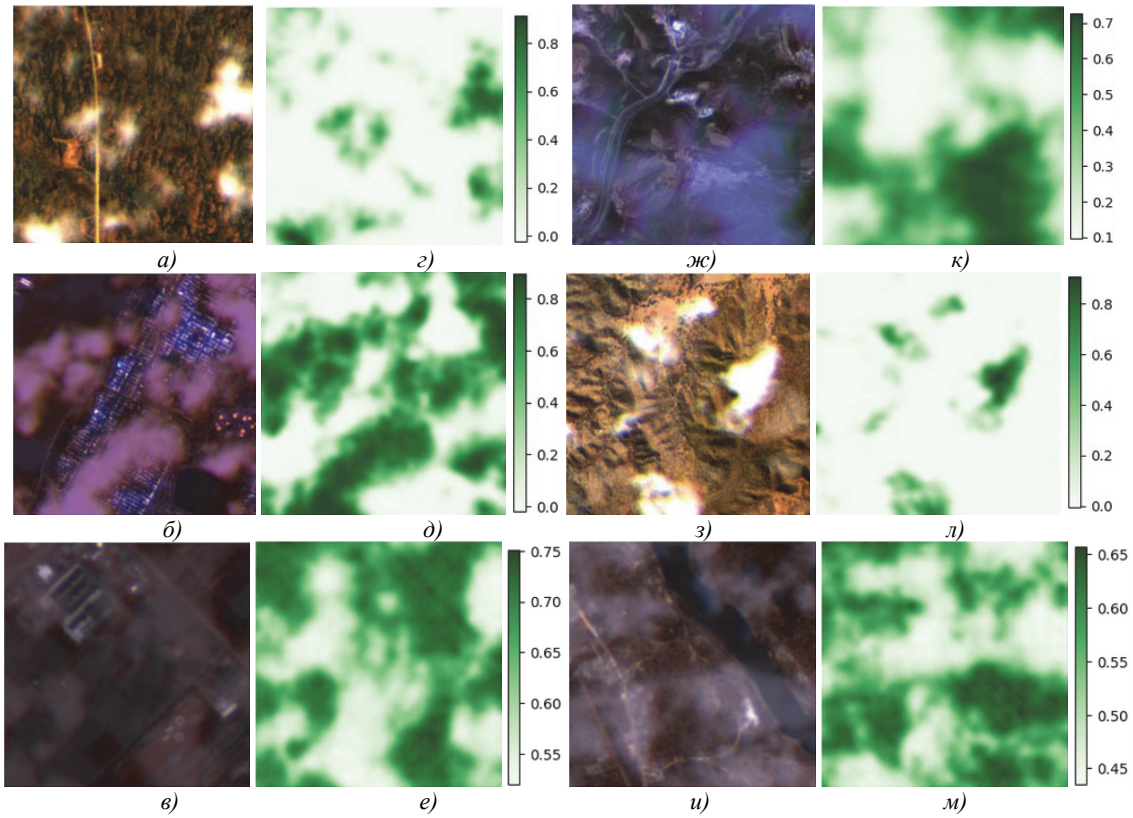


Рис. 5. Результаты распознавания облачности: (а – в, ж – и) оптическое изображение с синтезированной облачностью, (г – е, к – м) «истинные» поля прозрачности облачности

Для реальных облаков из набора данных SEN12MSCR визуально оценить качество (вторая обученная модель) оценки поля прозрачности облаков можно на рис. 6. Здесь первый и третий столбец представляют собой входные мультиспектральные изображения Sentinel 2 (показаны в псевдоцветах), второй и четвертый столбец -– полученная оценка поля прозрачности. Изображения РСА не показаны на рис. 5 и 6, поскольку облачность не меняет РСА-изображения земной поверхности.

Заключение

В работе предложен новый метод оценки поля прозрачности облачности, используемый в рамках общей технологии по устранению облачности и/или дымки на снимках ДЗЗ. Метод в качестве ключевого элемента впервые для указанной задачи применяет фундаментальную нейросетевую модель TerraMind, использующую в качестве мультимодальных данных территории съемки мультиспектральные изображения и изображения радара с синтезированной апертурой. В качестве исходных данных для обучения был использован набор SEN12MSCR, содержащий пары изображений спутников Sentinel-1 и Sentinel-2, в которых есть как изображения с облачностью, так и без нее. Предварительно обученную фундаментальную модель TerraMind мы обучали посредством аугментации данных, синтезируя облачность и соответствующие ей истинные поля прозрачности с использованием Satellite Cloud Generation Toolbox. Результаты проведенных экспериментальных исследования продемонстрировали высокие качественные показатели предложенного метода оценки, при этом качественные результаты оценки полей прозрачности на натурных снимках подтверждают его работоспособность. Высокое же качество решения общей технологии устранения облачности/дымки, в состав которой входит разработанное в настоящей работе решение, гарантируется предложенным авторами ранее методом реконструкции потерянных частей снимков, продемонстрировавшим наилучшие показатели на том же наборе данных по сравнению с существующими лучшими решениями.

Благодарности

Результаты исследования были получены при финансовой поддержке Российской Федерации в лице Минобрнауки России (Соглашение № 075-15-2024-558).

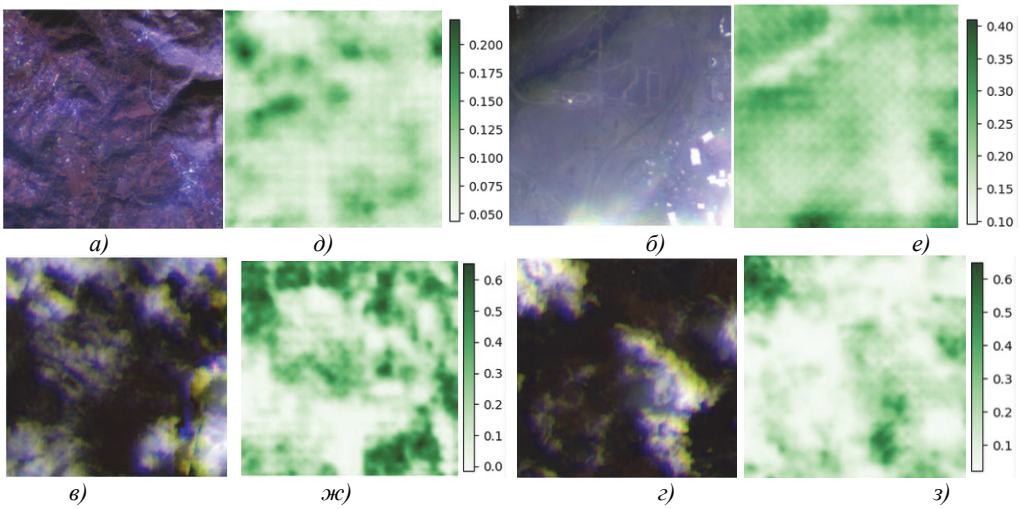


Рис. 6. Результаты оценки поля прозрачности облачности по мультиспектральным и PCA-снимкам: (а – г) входные мультиспектральные изображения (показан RGB-канал), (д – з) детектированная маска облачности

References

[1] King MD, Platnick S, Menzel WP, Ackerman SA, Hubanks PA. Spatial and temporal distribution of clouds observed by MODIS onboard the Terra and Aqua satellites. IEEE Transactions on Geoscience and Remote Sensing 2013; 51(7): 3826-3852. DOI: 10.1109/TGRS.2012.2227333.

[2] Konovalov VF, Myasnikov VV, Sergeev VV. Reconstruction of optical remote sensing images using synthetic aperture radar data and diffusion neural networks. Optoelectronics, Instrumentation and Data Processing 2025; 61(4): 449-457.

[3] Meraner A, Ebel P, Zhu XX, Schmitt M. Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion. ISPRS Journal of Photogrammetry and Remote Sensing 2020; 166: 333-346. DOI: 10.1016/j.isprsjprs.2020.05.013.

[4] Valovcin FR. Snow/Cloud Discrimination. Hanscom AFB, MA: Air Force Geophysics Laboratory, Air Force Systems Command, United States Air Force; 1976. 17 p. Air Force Surveys in Geophysics No. 349. Report AFGL-TR-76-0174.

[5] Kyle HL, Curran RJ, Barnes WL, Escoe D. A cloud physics radiometer. In: Proceedings of the Third Conference on Atmospheric Radiation; June 28-30, 1978; Davis, CA, USA. Boston, MA: American Meteorological Society; 1978: 107.

[6] Hall DK, Riggs GA, Salomonson VV. Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectroradiometer data. Remote Sensing of Environment 1995; 54(2): 127-140.

[7] Wang Z, Zhao L, Meng J, Han Y, Li X, Jiang R, Chen J, Li H. Deep learning-based cloud detection for optical remote sensing images: A survey. Remote Sensing 2024; 16(23): 4583. DOI: 10.3390/rs16234583.

[8] Li Z, Shen H, Weng Q, Zhang Y, Dou P, Zhang L. Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects. ISPRS Journal of Photogrammetry and Remote Sensing 2022; 188: 89-108. DOI: 10.1016/j.isprsjprs.2022.04.018.

[9] Huang L, Jiang B, Lv S, Liu Y, Fu Y. Deep learning-based semantic segmentation of remote sensing images: A survey. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2024; 17: 8370-8396. DOI: 10.1109/JSTARS.2024.3388508.

[10] Mahajan S, Fataniya B. Cloud detection methodologies: Variants and development—a review. Complex & Intelligent Systems 2020; 6: 251-261. DOI: 10.1007/s40747-019-00128-0.

[11] Zhu X, Helmer EH. An automatic method for screening clouds and cloud shadows in optical satellite image time series in cloudy regions. Remote Sensing of Environment 2018; 214: 135-153. DOI: 10.1016/j.rse.2018.05.024.

[12] Frantz D, Haß E, Uhl A, Stoffels J, Hill J. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. Remote Sensing of Environment 2018; 215: 471-481. DOI: 10.1016/j.rse.2018.04.046.

[13] Qiu S, Zhu Z, He B. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4-8 and Sentinel-2 imagery. Remote Sensing of Environment 2019; 231: 111205. DOI: 10.1016/j.rse.2019.05.024.

[14] Ning J, Xie L, Yin J, Liu Y. Cloud removal advances: A comprehensive review and analysis for optical remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2025; 18: 18621-18641. DOI: 10.1109/JSTARS.2025.3580718.

[15] Joglekar M, Deshpande A. Remote sensing image dehazing techniques: Systematic review, performance evaluation and future directions. Archives of Computational Methods in Engineering 2026. DOI: 10.1007/s11831-026-10500-2.

[16] Aybar C, Bautista L, Montero D, Contreras J, Ayala D, Prudencio F, Loja J, Ysuhuaylas L, Herrera F, Gonzales K, et al. CloudSEN12+: The largest dataset of expert-labeled pixels for cloud and cloud shadow detection in Sentinel-2. Data in Brief 2024; 56: 110852. DOI: 10.1016/j.dib.2024.110852.

[17] Main-Knorn M, Pflug B, Louis J, Debaecker V, Müller-Wilm U, Gascon F. Sen2Cor for Sentinel-2. In: Proceedings of the Living Planet Symposium 2016; May 9-13, 2016; Prague, Czech Republic. ESA Special Publication SP-740; 2017: 1-8.

[18] Hu J, Wei Y, Chen Z, Zhi X, Zhang W. CM-YOLO: Typical object detection method in remote sensing cloud and mist scene images. Remote Sensing 2025; 17(1): 125. DOI: 10.3390/rs17010125.

[19] Zhang J, Li Y, Yang X, Jiang R, Zhang L. RSAM-Seg: A SAM-based model with prior knowledge integration for remote sensing image semantic segmentation. Remote Sensing 2025; 17(4): 590. DOI: 10.3390/rs17040590.

[20] Yu Z, Idris MYI, Wang P. When cloud removal meets diffusion model in remote sensing. arXiv preprint arXiv:2504.14785; 2025. DOI: 10.48550/arXiv.2504.14785.
[21] Liu Y, Li W, Guan J, Zhou S, Zhang Y. Effective cloud removal for remote sensing images by an improved mean-reverting denoising model with elucidated design space. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2025: 17851-17861.
[22] Grohnfeldt C, Schmitt M, Zhu XX. A conditional generative adversarial network to fuse SAR and multispectral optical data for cloud removal from Sentinel-2 images. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. IEEE; 2018: 1726-1729. DOI: 10.1109/IGARSS.2018.8519215.
[23] Chen S, Zhang W, Li Z, Wang Y, Zhang B. Cloud removal with SAR-optical data fusion and graph-based feature aggregation network. Remote Sensing 2022; 14(14): 3374. DOI: 10.3390/rs14143374.
[24] Tu K, Yang C, Qing Y, Qi K, Chen N, Gong J. Cloud removal with optical and SAR imagery via multimodal similarity attention. ISPRS Journal of Photogrammetry and Remote Sensing 2025; 226: 116-126. DOI: 10.1016/j.isprsjprs.2025.05.004.
[25] Zhu ZH, Lu W, Chen SB, Ding C, Tang J, Luo B. Real-world remote sensing image dehazing: Benchmark and baseline. IEEE Transactions on Geoscience and Remote Sensing 2025; 63: 1-16. DOI: 10.1109/TGRS.2025.3584234.
[26] Zhao Z, Yan J, Li C, Wang X, Jiang P, Tang J. DehazeMamba: SAR-guided optical remote sensing image dehazing with adaptive state space model. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2026; 19: 778-794. DOI: 10.1109/JSTARS.2025.3634360.
[27] Makarau A, Richter R, Müller R, Reinartz P. Haze detection and removal in remotely sensed multispectral imagery. IEEE Transactions on Geoscience and Remote Sensing 2014; 52(9): 5895-5905. DOI: 10.1109/TGRS.2013.2293662.
[28] Huang B, Li Z, Yang C, Sun F, Song Y. Single satellite optical imagery dehazing using SAR image prior based on conditional generative adversarial networks. In: Proceedings of the IEEE Winter Conference on Applications of Computer Vision. IEEE; 2020: 1795-1802. DOI: 10.1109/WACV45572.2020.9093471.
[29] Shen H, Zhong T, Jia Y, Wu C. Remote sensing image dehazing using generative adversarial network with texture and color space enhancement. Scientific Reports 2024; 14: 12382. DOI: 10.1038/s41598-024-63259-6.
[30] Zhou H, Wu X, Chen H, Chen X, He X. RSDehamba: Lightweight vision Mamba for remote sensing satellite image dehazing. arXiv preprint arXiv:2405.10030; 2024. DOI: 10.48550/arXiv.2405.10030.
[31] Ma X, Wang Q, Tong X. Nighttime light remote sensing image haze removal based on a deep learning model. Remote Sensing of Environment 2025; 318: 114575. DOI: 10.1016/j.rse.2024.114575.
[32] Zhu Z, Luo Y, Wei H, Li Y, Qi G, Mazur N, Li Y, Li P. Atmospheric light estimation based remote sensing image dehazing. Remote Sensing 2021; 13(13): 2432. DOI: 10.3390/rs13132432.
[33] Jakubik J, Yang F, Blumenstiel B, et al. TerraMind: Large-scale generative multimodality for Earth observation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision; 2025.
[34] Ebel P, Xu Y, Schmitt M, Zhu XX. SEN12MS-CR-TS: A remote sensing data set for multimodal multitemporal cloud removal. IEEE Transactions on Geoscience and Remote Sensing 2022; 60: 5222414. DOI: 10.1109/TGRS.2022.3146246.
[35] Song Y, He Z, Qian H, Du X. Vision transformers for single image dehazing. arXiv preprint arXiv:2204.03883; 2022.
[36] Schmitt M, Hughes LH, Qiu C, Zhu XX. SEN12MS: A curated dataset of georeferenced multi-spectral Sentinel-1/2 imagery for deep learning and data fusion. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2019; IV-2/W7: 153-160. DOI: 10.5194/isprs-annals-IV-2-W7-153-2019.
[37] Jiang B, Chen G, Wang J, Ma H, Wang L, Wang Y, Chen X. Deep dehazing network for remote sensing image with non-uniform haze. Remote Sensing 2021; 13(21): 4443. DOI: 10.3390/rs13214443.
[38] Czerkawski M, Atkinson R, Michie C, Tachtatzis C. SatelliteCloudGenerator: Controllable cloud and shadow synthesis for multispectral optical satellite images. Remote Sensing 2023; 15(17): 4138. DOI: 10.3390/rs15174138.
[39] Kong F, Huang B, Bradbury K, Malof J. The Synthinel-1 dataset: A collection of high-resolution synthetic overhead imagery for building segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. IEEE; 2020: 1814-1823. DOI: 10.1109/WACV45572.2020.9093500.
[40] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N. An image is worth 16×16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929; 2021. DOI: 10.48550/arXiv.2010.11929.
[41] Mentzer F, Minnen D, Agustsson E, Tschannen M. Finite scalar quantization: VQ-VAE made simple. arXiv preprint arXiv:2309.15505; 2023. DOI: 10.48550/arXiv.2309.15505.

Сведения об авторах

Коновалов Виталий Федорович, 1999 года рождения, в 2022 г. окончил Самарский национальный исследовательский университет имени академика С.П. Королева (Самарский университет). В настоящее время работает инженером АО «Самара-Информспутник». Круг научных интересов включает компьютерное зрение, генеративные нейронные сети, вычисления общего назначения на графических процессорах, обработка и восстановление изображений. E-mail: *vitfvk@gmail.com*

Мясников Владислав Валерьевич, в 2008 защитил диссертацию на соискание степени доктора физико-математических наук. Работает профессором кафедры геоинформатики и информационной безопасности Самарского национального исследовательского университета имени академика С.П. Королева. Круг научных интересов: компьютерное зрение, распознавание образов, искусственный интеллект, машинное обучение и

геоинформатика. Имеет более 200 публикаций, в том числе более 100 статей и три монографии (в соавторстве). E-mail: vmyas@geosamara.ru

Сергеев Владислав Викторович, 1951 года рождения. В 1974 году окончил Куйбышевский авиационный институт (ныне Самарский национальный исследовательский университет имени академика С. П. Королёва). В 1978 г. защитил кандидатскую, а в 1993 году – докторскую диссертацию. В настоящее время является профессором кафедры геоинформатики и информационной безопасности Самарского университета. Область научных интересов: цифровая обработка сигналов, анализ изображений, распознавание образов, геоинформатика. E-mail: vserg@geosamara.ru

Поступила в редакцию 26 март 2026 г. *Окончательный вариант – 28 апреля 2026 г.*

A method for cloud soft segmentation using multimodal remote sensing data

V.F. Konovalov [1], V.V. Myasnikov [1], V.V. Sergeev [1]
[1] Samara National Research University, Moskovskoye Shosse 34, Samara, 443086, Russia

Abstract

This paper presents a method for estimating cloud transparency fields using the TerraMind foundation neural network model. The TerraMind model utilizes multispectral and synthetic aperture radar images as multimodal data for the survey area. The TerraMind model was trained on the SEN12MSCR dataset, containing pairs of Sentinel-1 and Sentinel-2 satellite images, both with and without cloud cover. To further train the fundamental model to solve the desired problem, we used augmentation, synthesizing cloud cover and the corresponding true transparency fields using the Satellite Cloud Generation Toolbox. Numerical results from the experimental study demonstrate the high quality of the proposed estimation method, and the qualitative results of estimating transparency fields on real-world images confirm the method's effectiveness on real data.

Keywords: Sentinel 2, Sentinel 1, synthetic aperture radar, multispectral imagery, neural networks, cloud detection, image dehazing, synthetic dataset, TerraMind.

Citation: Konovalov VF, Myasnikov VV, Sergeev VV. A method for cloud soft segmentation using multimodal remote sensing data. Computer Optics 2026; 50(5): 2005. doi: 10.18287/COJ2005.

Acknowledgements: The study was financially supported by the Ministry of Education and Science of the Russian Federation under contract No. 075-15-2024-558.

About authors

Vitaly Fedorovich Konovalov, (b. 1999), graduated from Samara National Research University (Samara University) in 2022. Currently works as an engineer at Samara-Informsputnik. Research interests: generative neural networks, computer vision, GPGPU computing, generative neural networks, image processing and image restoration. E-mail: *vitfvk@gmail.com*

Vladislav Valerievich Myasnikov, received his DrSc degree in Physics & Maths (2008). Currently he works as a professor at the Geoinformatics and Information Security department in Samara National Research University. The range of scientific interests: computer vision, pattern recognition and artificial intelligence, machine learning and geoinformatics. He has about 200 publications, including more than 100 articles and three monographs. E-mail: *vmyas@geosamara.ru*

Vladislav Viktorovich Sergeev, (b. 1951) graduated from Samara National Research University (Samara University) in 1974. In 1978 he received the Candidates of Technical Sciences degree and in 1993 – the Doctor of Technical Sciences degree. Currently he works as the professor of Geoinformatics and Information Security department at Samara University. Research interests are digital signals and image processing, GIS and pattern recognition. E-mail: *vserg@geosamara.ru*

Received March 26, 2026. The final version – April 28, 2026.