

УДК 004.925

ТЕХНОЛОГИЧЕСКИЕ ПРОБЛЕМЫ ПЕРЕХОДА ОТ ОБРАБОТКИ АЭРОФОТОСНИМКОВ НИЗКОГО РАЗРЕШЕНИЯ К СНИМКАМ ВЫСОКОГО РАЗРЕШЕНИЯ

Тай Зар Лин, Абрамов Ю.А., Глебов С.А., Черепков Е.А.

Калужский филиал федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)», г. Калуга (248000, г. Калуга, ул. Баженова, д. 2.), e-mail: kominhtet94@gmail.com

Обработка информации, представленной в виде изображений, в настоящее время получила широкое распространение в научных исследованиях, промышленности, медицине, а также в некоторых других областях деятельности, где имеют дело с фотографиями. Это связано с возможностью регистрации при помощи специальной аппаратуры больших объёмов информации – терабайтов на 2018 год - с высокой точностью. Примером может служить составление карт по аэрофотоснимкам. Интенсивное развитие космических средств регистрации изображений является настоятельным фактором создания систем обработки изображений. Данная статья исследует текущие практики, проблемы и перспективы перехода от обработки аэрофотоснимков низкого разрешения к снимкам высокого разрешения. Классификация изображений является комплексным процессом, на который может воздействовать множество факторов. В статье рассматриваются следующие классификаторы: нейронные сети, повышение и случайные леса, AdaBoost, SVM – метод опорных векторов. Кроме того, обсуждаются некоторые важные проблемы, воздействующие на производительность классификации. Также исследованы различные методы обучения особенностей: использование огромного участка изображения, применение банков фильтров, неконтролируемые методы, задействование информации о высоте. Эффективное использование множественных особенностей удалённо полученных данных и выборка подходящего метода классификации особенно значимы для улучшения точности классификации. В заключение рассмотрено использование гораздо большего количества данных для обучения. Интеграция авиационных фотоснимков, географических информационных систем и экспертных систем приводит к новому авангарду исследования. Однако, нужно больше исследований для идентификации и сокращения неопределённостей в цепочке обработки изображений для улучшения качества классификации.

Ключевые слова: маркировка изображений, машинное обучение, случайные леса

TECHNOLOGICAL PROBLEMS OF TRANSITION FROM LOW RESOLUTION AERIAL IMAGES PROCESSING TO HIGH RESOLUTION IMAGES

Thai Zhar Ling, Abramov Y.A., Glebov S.A., Cherepov E.A.

Bauman Moscow State Technical University (Kaluga Branch), Kaluga (248000, Kaluga, Bazhenova st, 2.), e-mail: kominhhtet94@gmail.com

Information processing presented as images today got wide spread in scientific research, industry, medicine and also other photo business spheres. It's connected with registration possibility with help of special equipment of big information volume – terabytes in 2018 – with high accuracy. Making maps by aerial images can serve as example. Intensive development of space image registration means is an insistent factor of creating image processing systems. This article explores current practices, problems and perspectives of transition from low resolution aerial images processing to high resolution images. Image classification is a complex process which can be affected by many factors. In the article we meet the next classifiers: neural networks, boosting and random forests, AdaBoost, SVM – support vector machine. In addition, some important problems are discussed, which affect classification performance. Different features learning methods are also researched: huge image patch usage, filter banks applying, unsupervised methods, including height information. Effective multiple features usage of remote received data and suitable classification method selection are especially meaningful for classification accuracy increase. In conclusion much more amount of data for learning is considered. Integrity of aerial images, geographical informational systems and expert systems leads to a new research vanguard. However, more research for identification and uncertainty reduction in the chain of image processing for classification quality improvement is needed.

Key words: image labelling, machine learning, random forests

Спутники Ikonos и Quickbird, запущенные в 1999 и 2001 годах, начали получать панхроматические изображения поверхности Земли с разрешением примерно один квадратный метр на пиксель, что значительно увеличило число возможных применений систем интерпретации аэрофотоснимков. Поскольку в данном решении классы представляют собой искусственные объекты, такие как здания, дороги и автомобили, подход к обучению простых классификаторов по локальным спектральным и текстурным особенностям, который был успешен на изображениях с низким разрешением, больше не приводит к приемлемой точности. [7] Проблема классификации изображений с высоким разрешением намного сложнее и требует знания формы и контекста в дополнение к текстуре. [3]

Во время перехода от изображения с низким разрешением к высокому разрешению наиболее заметными были:

1. Переключение на более мощные или сложные классы, такие как AdaBoost, SVM и случайные леса. [6]

2. Использование более пространственного контекста и более богатых функций ввода.
3. Использование гораздо большего объёма данных для обучения и тестирования. [2]
4. Использование структурированных методов прогнозирования, таких как условные случайные поля.

Дискриминация между классами объектов с подобной текстурой, например, дорогами и зданиями, требует определенных знаний о форме и контексте, что, в свою очередь, приводит к гораздо более сложным границам принятия решений, чем те, которые необходимы для различения лесных и застроенных областей в изображениях с низким разрешением. Из-за необходимости изучать такие нелинейные границы решений, приложения машинного обучения к изображениям с высоким разрешением основаны на более сложных классах, таких как SVM, случайные леса и различные типы повышения. [9]

В то время как нейронные сети способны изучать нелинейные границы решений и широко используются в приложениях дистанционного зондирования, многие исследователи обнаружили, что их трудно обучать из-за наличия локальных оптимумов. SVM представляет привлекательную альтернативу нейронным сетям, поскольку, подобно нейронным сетям, они могут изучать нелинейные границы решений, но, в отличие от нейронных сетей, SVM оптимизируют выпуклую функцию потерь и не страдают от проблемы локальных оптимумов. [1] Поскольку SVM представляют собой, по существу, сложные шаблоны, они плохо работают при применении в качестве классификаторов к исходным изображениям. Они обычно используются в сочетании с функциями более высокого уровня в сообществе компьютерного зрения. Применение SVM для интерпретации аэрофотоснимков было намного более примитивным, чем в сообществе компьютерного зрения, причем большинство документов, использующих SVM, классифицировали пиксели с использованием только низкоуровневых функций.

Некоторые из наиболее успешных подходов по маркировке аэрофотоснимков высокого разрешения основывались на различных ансамблевых методах, особенно популярными были ускоряющие и случайные леса. Порвэй разработал иерархическую модель для анализа аэрофотоснимков, которая основывалась на детекторах снизу вверх для автомобилей, дорог, автостоянок и зданий, которые были обучены с использованием

различных типов повышения. Другими заметными приложениями для повышения эффективности являются работа Доллара, который разработал общую структуру для обучения обнаружению границ изображений с использованием расширенного класса пикселей и представил некоторые качественные результаты обнаружения дорожного движения, и работа Нгуена, который использовал онлайн-повышение, для обнаружения автомобилей. [4] Общей причиной использования ускорения в этих приложениях является его способность выполнять выбор из очень большого набора функций, когда ученик рассматривает одну функцию. Доллар смог использовать пул из 50 000 ответов в качестве функций. [10]

Случайный лес - это другой ансамблевый метод на основе дерева, который широко используется в приложениях для маркировки изображений. [11] Случайная лесная классификация состоит из нескольких деревьев решений, прогнозы которых обычно объединяются с использованием мажоритарного голосования. Цель учебной процедуры - уменьшить дисперсию ансамбля, стремясь произвести декоррелированные деревья. Это достигается путем изучения каждого дерева на случайном подмножестве набора данных и использовании случайного подмножества входных переменных. В ряде работ Клюкнера используются случайные леса для выполнения семантической классификации аэрофотоснимков с впечатляющими результатами.

В то время как случайные леса и усиление классов создают сборники деревьев, они делают это совершенно по-разному, что приводит к явным преимуществам и недостаткам. Такие как AdaBoost, как известно, плохо работают при наличии аномалий или несоблюдении правил обучения, поскольку они, как правило, подчеркивают сложные случаи во время обучения. Это может стать серьезным ограничением в контексте интерпретации аэрофотоснимков, поскольку зарегистрированная и актуальная информация о метках редко доступна. С другой стороны, случайные леса гораздо менее подвержены влиянию данных, не имеющих маркировки, поскольку каждое дерево построено на случайном подмножестве данных обучения с использованием случайного подмножества входных признаков, и особого внимания не уделяется трудным учебным случаям.

Кроме того, случайные леса трудно распараллеливаются, в то время как ускорение гораздо сложнее распараллеливать из-за его последовательного характера. Учитывая эти причины, случайные леса, по-видимому, немного лучше подходят для классификации аэрофотоснимков.

Подход использования значений нескольких диапазонов в одном пикселе или даже небольшого размера окна, например, 5×5 , в качестве входа в класс, является безнадёжным по сравнению с изображениями с высоким разрешением, поскольку вход просто не содержит достаточной информации для различения классов объектов. Самый простой способ решения этой проблемы - использовать в качестве входного окна большее входное окно. Mnih и Hinton показали, что увеличение размера входного участка с 24×24 на 48×48 , что уже является большим размером контекста по сравнению с другой работой, до 64×64 значительно улучшает точность и позволяет выполнить задачу обнаружения дороги.

Простое использование участка большого изображения для ввода может быть медленным даже на современных компьютерах, потому что вычислительная стоимость применения линейного фильтра для квадратного участка квадратично масштабируется с шириной участка. По этой причине недавняя работа опиралась на эффективно вычисляемые функции для масштабирования до больших размеров контекста. Наиболее широко используемый класс эффективно вычисляемых функций изображения представляет собой набор признаков, которые могут быть выражены как линейная комбинация сумм прямоугольных областей изображения. Пусть $I(x, y)$ - значение интенсивности изображения одного канала в местоположении (x, y) и определяет

$$S(x, y) = \sum_{x' \leq x} \sum_{y' \leq y} I(x', y').$$

S известен как интегральное изображение I и может быть вычислен за время, линейное относительно числа пикселей в изображении. После вычисления S сумма любого суб-прямоугольника I может быть вычислена за постоянное время. Например, сумма I по прямоугольнику $[a, b] \times [c, d]$ может быть вычислена как $S(b, d) - S(b, c) - S(a, d) + S(a, c)$. Функции, которые могут быть эффективно вычислены с использованием интегрального трюка изображения, включают в себя специальные случаи вейвлетов Хаара и гистограммы ориентированных градиентов.

Более поздняя тенденция в маркировке авиационных изображений - использование неконтролируемых методов обучения, таких как разреженное кодирование и ограниченные машины Больцмана. Было показано, что использование функций, изученных ограниченной машиной Больцмана для инициализации нейронной сети, обученной по задаче обнаружения дорожного движения, значительно улучшает как точность, так и отзыв на обучение нейронной сети по сравнению с случайной инициализацией. [5] Ригамонти использовал вариант разреженного кодирования для изучения функций для нескольких задач сегментации линейной структуры. Такие неконтролируемые подходы к изучению функций, как правило, изучают фильтры, которые напоминают ориентированные краевые детекторы и всплески Габора, и имеют такое преимущество, что они могут изучить лучшие фильтры для этой задачи, а не просто выбирать лучшие из множества функций, которые появляются только в нескольких масштабах и ориентациях.

Другим типом функции, который, как было показано, очень полезен для интерпретации аэрофотоснимков, является информация о высоте. Клюкнер показал, что использование высоты в дополнение к богатым внешним возможностям повышает точность классификации на 10-20%. Результат не должен удивлять, потому что информация о высоте может помочь различать очень путанные пары объектов, такие как дороги / здания и трава / деревья. Информация о высоте может быть получена непосредственно из данных LIDAR или из перекрывающихся пар изображений с использованием стереотехники. Будучи чрезвычайно полезной, получение информации о высоте может быть затруднено, поскольку данные LIDAR менее широко доступны и дороже, чем аэрофотоснимки, и большинство аэрофотоснимков происходит в виде неперекрывающихся ортоэкранированных фрагментов изображения, которые не содержат стереоинформации.

В ранних работах по маркировке аэрофотосъемки было распространено использование одного изображения для обучения и тестирования. Такой небольшой объем данных не может охватывать широкий диапазон изменений внешнего вида различных объектов, ведущих к классификаторам, которые вряд ли будут хорошо работать с невидимыми данными. Ограниченный размер тестового набора, а также его сходство с

данными обучения означает, что уровни точности, о которых сообщалось в более ранних работах, вряд ли можно перевести на более крупные массивы данных.

Недавние приложения машинного обучения к аэрофотоснимкам высокого разрешения использовали гораздо больше данных для обучения. Например, Rogway использовал 120 изображений размером от 640x480 до 1000x1000 для обучения своей иерархической вероятностной грамматической модели аннотаций антенных изображений. Клюкнер использовал три набора данных из примерно 100 изображений с каждым набором данных, охватывающим от 5 до 8 квадратных километров, для обучения случайной лесной классификации, в то время как Мних и Хинтон обучали большую нейронную сеть на 130 больших изображениях, которые занимают площадь около 500 квадратных километров разрешением 0,7 пикселя на квадратный метр. Наборы данных такого размера, скорее всего, будут содержать значительные изменения в появлении объектов, а при использовании для обучения эти наборы данных значительно усложняют работу мощных классов. [8]

Список литературы

1. Гришанов К.М., Рыбкин С.В. Тестирование свёрточной нейронной сети в задачах машинного зрения // Электронный журнал: наука, техника и образование. 2017. № 2 (12). С. 186-193.
2. Черепков Е.А., Рыбкин С.В. Технологии для обработки и анализа больших данных // Электронный журнал: наука, техника и образование. 2016. № 4 (9). С. 120-127.
3. Редько А.В., Рыбкин С.В. Обзор методов реконструкции трехмерной сцены по изображениям // Электронный журнал: наука, техника и образование. 2016. № 4 (9). С. 86-91.
4. Сакович И.О., Белов Ю.С. Обзор основных методов контурного анализа для выделения контуров движущихся объектов // Инженерный журнал: наука и инновации. 2014. № 12 (36). С. 11.

5. Нестеров А.Ю., Белов Ю.С. Распознавание образов по уникальным точкам на примере дорожных знаков // Электронный журнал: наука, техника и образование. 2016. № 4 (9). С. 113-119.
6. Редько А.В., Молчанов А.Н., Белов Ю.С. Использование алгоритмов определения ключевых точек изображения в задаче реконструкции трёхмерных сцен // Электронный журнал: наука, техника и образование. 2017. № СВ2 (13). С. 59-66.
7. Гуркина Е.Д., Белов Ю.С. Обнаружение частично размытых изображений и их классификация. // Научное обозрение. Технические науки. 2017. № 2. С. 38-41.
8. С. Лазебник, К. Шмид и Дж. Понсе. Beyond Bags of Features: Пространственная пирамида, подходящая для распознавания категорий естественных сцен // Компьютерное зрение и распознавание образов. 2006. № 2 С. 2169-2178.
9. Xiaomu Song, Guoliang Fan, and Mahesh Rao. Automatic CRP Mapping Using Machine Learning Approaches // IEEE Transactions on Geoscience and Remote Sensing. 2005. № 43 (4) С. 888-897.
10. Piotr Dollar, Zhuowen Tu, and Serge Belongie. Supervised learning of edges and object boundaries // 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2006. С. 1964-1971.
11. Stefan Kluckner and Horst Bischof. Semantic classification by covariance descriptors within a randomized forest // Computer Vision Workshops (ICCV), IEEE. 2009. С. 665-672.