

РАСПОЗНАВАНИЕ ЗДАНИЙ НА СПУТНИКОВЫХ СНИМКАХ СВЕРХВЫСОКОГО РАЗРЕШЕНИЯ

Светлана Дьяконова

Аннотация. В работе рассматривается применение нечеткого классификатора NEFClass к задаче распознавания зданий на спутниковых снимках сверхвысокого разрешения. Проводится сравнительный анализ методов обучения нейронной сети: генетического, градиентного и метода сопряженных градиентов касательно задачи распознавания зданий с различными формами крыши.

Ключевые слова: распознавание зданий на спутниковых изображениях, нечеткий классификатор NEFClass, методы обучения нейронных сетей.

ACM Classification Keywords: 1.4.7 Computing Methodologies - IMAGE PROCESSING AND COMPUTER VISION – Feature Measurement - Size and shape. G.1.6 Mathematics of Computing – NUMERICAL ANALYSIS – Optimization- Gradient methods. I.4.8 Computing Methodologies - IMAGE PROCESSING AND COMPUTER VISION – Scene Analysis - Object recognition. I.5.1 Computing Methodologies - PATTERN RECOGNITION - Models - Neural nets.

Введение

Задача распознавания зданий на спутниковых изображениях сверхвысокого разрешения является важной частью в решение задачи автоматической интерпретации данных получаемых из систем дистанционного зондирования земли. Применение такой системы распознавания зданий возможно, например, для автоматического анализа использования земельных ресурсов, для поддержания городских геодезических данных в актуальном состоянии, в частности для контроля качества и автоматического обновления геодезических данных. Предложенные ранее работы по теме распознавания зданий разработаны с учетом того, что исследования проводятся на спутниковых снимках, взятых из одного региона, где все здания имеют определенный оттенок и форму крыши. Описываемый в данной статье подход позволяет использовать систему распознавания для различных видов населенных пунктов, это может быть город, пригород или населенный пункт в пустынном регионе. В частности в работе рассматривается три вида крыши зданий: плоская, круглая и треугольная крыша. Каждая из этих форм соответствует одному из типов построек. Здания с плоскими крышами преобладают в городских территориях, и соответствуют высокоэтажному строению, треугольные крыши наиболее часто встречаются в домах загородного типа, круглые крыши зданий преобладают в населенных пунктах находящихся в пустынных регионах. На основе знаний о геометрической форме крыши зданий были подобраны оптимальные признаки для распознавания. В качестве классификатора была выбрана архитектура системы NEFClass, которая впоследствии была программно реализована с тремя алгоритмами обучения: генетическим, градиентным и сопряженных градиентов.

Описание архитектуры нечеткого классификатора NEFClass

Модель NEFClass [Зайченко, 2008] используется для определения класса или категории полученных входных образцов (паттернов). Паттерны — это векторы признаков $x = (x_1, \dots, x_n) \in R^n$, а класс C — подмножество R^n . Принимаем, что пересечение двух разных классов является пустым. Величина признака паттерна представляется нечетким множеством, и классификация определяется множеством лингвистических правил. Для каждого входного признака x_i существует q_i нечетких множеств $\mu_1^i, \dots, \mu_{q_i}^i$. Также есть база правил, которая вмещает k нечетких лингвистических правил R_1, \dots, R_k .

База правил представляет собой аппроксимацию функции (неизвестной) $\varphi: R^n \rightarrow \{0,1\}^m$, которая представляет задачу классификации, где $\varphi(x) = (c_1, \dots, c_m)$ такая, что $c_i = 1, c_j = 0$ ($j \in \{1, \dots, m\}, j \neq i$). Значит, x принадлежит классу $C_i \rightarrow [0,1]^m$.

Обозначим наибольший компонент каждого вектора c единицей ($c = 1$), а все другие компоненты положим равными нулю.

Нечеткие множества и лингвистические правила, обеспечивающие выполнение такой аппроксимации и определяющие результирующую систему NEFClass, будут получены из множества примеров для обучения. На Рис. 1 приведена система NEFClass, которая классифицирует входные образцы с двумя признаками и двумя отдельными классами, используя пять лингвистических правил.

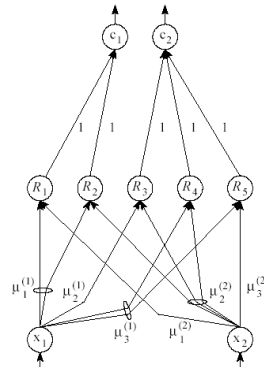


Рис.1 Пример нейросети NEFClass с 2 признаками, 5 правилами и 2 классами

NEFClass — это трехслойный нечеткий перцептрон со следующими спецификациями:

1. $U_1 = \{x_1, \dots, x_n\}$ - входной слой, $U_2 = \{R_1, \dots, R_k\}$ - слой нейронов правил, $U_3 = \{c_1, \dots, c_m\}$ - выходной слой классифицирующих нейронов.
2. Каждая связь между узлами $x_i \in U_1$ и $R_r \in U_2$ помечена лингвистическим термом $A_{jr}^{(i)}$ ($j_r \in \{1, \dots, q_i\}$).
3. $W(R, c) \in \{0, 1\}$ определяется для всех $R \in U_2, C \in U_3$.
4. Соединения, которые происходят из одного и того же входного узла x и имеют идентичные метки, одинаковый вес в любой момент времени. Такие соединения называются связанными соединениями, а их вес называется взвешенным.
5. Обозначим $L_{x,R}$ метку соединения между узлами $x \in U_1$ и $R \in U_2$, тогда для всех $R, R' \in U_2$ выполняется

$$(\forall (x \in U_1) L_{x,R} = L_{x,R'}) \Rightarrow R = R'.$$

6. Для всех узлов правил $R \in U_2$ и всех узлов $C, C' \in U_3$ имеем

$$(W(R, C) = 1) \wedge (W(R, C') = 0) \Rightarrow C = C'.$$

7. Для всех выходных узлов $C \in U_3$ выполняется

$$O_c = a_c = NET_c.$$

8. Для всех выходных узлов $C \in U_3$ NET_c рассчитывается по формуле

$$NET_c = \frac{\sum_{R \in U_2} W(R, C) O_R}{\sum_{R \in U_2} W(R, C)}.$$

Система NEFClass может быть построена по неполному знанию об образцах, а затем откалибрована путем обучения или создана с самого начала посредством обучения. Пользователь должен определить количество начальных нечетких множеств, разбивающих интервал значений, в которых лежит признак, а также наибольшее количество узлов правил, созданных в скрытом слое.

Обучение в системе NEFClass

Для обучения была использована гауссовская функция принадлежности [Зайченко, 2004], поскольку она дает более реальные данные и обеспечивает гибкость настройки нейронной сети.

$$\mu(x) = \exp\left[-\left(\frac{x-c}{\sigma}\right)^2\right]$$

Рассмотрим непосредственно алгоритм обучения. Алгоритм обучения системы NEFClass [Зайченко, 2008] состоит из двух этапов обучения. На первом этапе генерируется или задается база лингвистических правил вида «если - то». На втором этапе обучения производится настройка весов нейронной сети с помощью следующих алгоритмов обучения: градиентного, генетического и алгоритма сопряженных градиентов.

Рассмотрим систему NEFClass с n входными нейронами x_1, \dots, x_n , $k \leq k_{max}$ нейронами правил и m выходными нейронами C_1, \dots, C_m . Также задано обучающее множество образцов $L = \{(p_1, t_1), \dots, (p_s, t_s)\}$, каждый из которых состоит из входного образца $p \in R^n$ и желаемого образца $t \in \{0, 1\}^m$.

Целью первого этапа является создать k нейронов правил в системе NEFClass, данный этап состоит из следующих шагов:

1. Выбираем следующий образец (p, t) из L
2. Для каждого входного нейрона $x_i \in U_1$ находим такую функцию принадлежности μ_j , что:

$$\mu_{j_i}^{(i)}(p_i) = \max_{j \in \{1, \dots, q_i\}} \{\mu_{j_i}^{(i)}(x_i)\}$$

3. Если узлов правил осталось меньше, чем k_{max} и не существует узла правила R с $W(x_1, R) = \mu_{j_1}^{(1)}, \dots, W(x_n, R) = \mu_{j_n}^{(n)}$, то создаем такой узел и соединя его с выходным узлом C_i , если $t_i = 1$.
4. Если в L остались необработанные образцы и $k \leq k_{max}$, то идем в начало и повторяем снова все шаги, иначе останавливаемся.

Целью второго этапа обучения является настройка весов нейронной сети. В данном исследовании были использованы три алгоритма обучения нейронных сетей: градиентный, генетический и алгоритм сопряженных градиентов. Результаты исследования и сравнительная характеристика работы этих алгоритмов приводятся ниже в данной статье.

Вычисление признаков объектов

В данной части статьи рассматриваются признаки, которые станут основой для распознавания зданий на спутниковых изображениях.

В работе рассматривается 3 вида крыш: плоская, круглая и треугольная крыша. Каждый из этих видов крыши представляет различные населенные пункты, отличающиеся местоположением и ландшафтом. Так здания с плоскими крышами преобладают в городских территориях, и соответствуют высокоэтажному строению. Треугольные крыши наиболее часто встречаются в домах загородного типа, распознать на изображении большую часть таких крыш, можно говорить о том, что данный населенный пункт является небольшим и в нем преобладают невысокие строения. Круглые крыши зданий преобладают в

населенных пунктах находящихся в засушливых, пустынных регионах. Кроме того такой вид крыши часто обозначает религиозные сооружения, если вокруг преобладают плоские и/или треугольные крыши. На Рисунке 2 приведены типичные изображения населенных пунктов, которые можно охарактеризовать по форме крыш как город, пригород и населенный пункт в пустынном регионе.



Рис. 2 Изображения населенных пунктов:

(а) - город в пустынном регионе; (б) - пригород; (в), (г) - город

На основе знаний о геометрической форме крыши зданий были подобраны оптимальные признаки для распознавания. Рассмотрим подробнее используемые в работе признаки.

1. Округлость здания для каждого сегмента рассчитывается как отношение площади к периметру. Этот признак может принимать значения от 0 до 1. Вычисляется по следующей формуле:

$$\text{Округлость} = \frac{4\pi * \text{Площадь}}{\text{Периметр}^2}$$

где Площадь сегмента рассчитывается как количество принадлежащих ему пикселей, а Периметр сегмента рассчитывается как количество пикселей лежащих на границе области.

2. Среднее значение углов [Müller, Zaum, 2005]. Углы являются важным признаком, позволяющим дифференцировать здания от других объектов. Для того чтобы определить значение углов здания в выделенном сегменте необходимо построить шестиугольник по следующей схеме: 1. Проводится главная ось между двумя точками контура, которые находятся на максимальном расстоянии друг от друга. 2. Определяются две вспомогательные оси с каждой стороны главной оси. Для этого строим вертикальные линии, лежащие на главной оси и имеющие максимальное расстояние от контура до главной оси. 3. Проводим поперечную ось, это должна быть вертикальная линия на главной оси, которая соединяет две точки контура находящиеся на максимальном расстоянии друг от друга. 4. Полученные шесть точек соединяем линиями в шестиугольник. В качестве признаков возьмем среднее значение двух углов, находящиеся по обе стороны главной оси.
3. Длина здания определяется как отношение длины главной оси к длине поперечной оси полученного выше шестиугольника [Müller, Zaum, 2005]:

$$\text{Длина} = \frac{m_{20} + m_{02} + \sqrt{(m_{20} - m_{02})^2 + 4m_{11}^2}}{m_{20} + m_{02} - \sqrt{(m_{20} - m_{02})^2 + 4m_{11}^2}}$$

4. Полнота вычисляется как отношение общего количества пикселей, которые лежат внутри шестиугольника и не принадлежат к исследуемой области к количеству пикселей, которые лежат за пределами шестиугольника и принадлежат исследуемому сегменту.
5. Градиент яркости поверхности крыши. Для определения градиента яркости будет использован оператор Собеля, который вычисляет приближенное значение градиента яркости изображения. В результате применения оператора Собеля в каждой точке изображения получим вектор градиента яркости в этой точке. Таким образом, будет найдено направление наибольшего увеличения яркости и величина её изменения в этом направлении. В разработанном алгоритме изображение делится на 20 сегментов, значение оператора Собеля вычисляется в центре каждого такого сегмента. В зависимости от полученных направлений увеличения яркости можно говорить об одном из видов крыш.

Экспериментальные исследования

Для проверки эффективности работы рассматриваемой нейронной сети были проведены тесты на изображениях с различными видами крыш. Для проведения экспериментов было использовано 90 изображений, из них 60 изображений использовались для обучения нейронной сети. Изображения с круглыми крышами брались из регионов с преобладанием пустынного ландшафта, треугольные - из пригородных территорий, а изображения с плоскими крышами – из территорий крупных городов. Для тестирования было использовано 30 изображений.

В разработанной системе NEFClass входам нейронной сети соответствуют описанные выше 5 признаков: округлость, среднее двух углов, лежащих по обе стороны главной оси, построенного шестиугольника, длина здания, полнота, градиент яркости поверхности крыши. Выходам сети соответствуют 3 вида крыш: круглая, треугольная, плоская.

Эксперимент был поставлен следующим образом. Каждое изображение было предварительно обработано и сегментировано на отдельные образцы – паттерны. Каждый такой паттерн содержит изображение крыши одного из видов зданий. Для генерации базы правил каждому паттерну на изображении ставится в соответствие подходящая форма крыши. Таким образом, создается база правил. Проводим обучение нейронной сети одним из выбранных методов: генетическим, градиентным или методом сопряженных градиентов. Далее на обученную сеть подаем тестовое изображение, на выходе получаем изображение, на котором отмечены распознанные крыши различными цветами, в зависимости от формы крыши. Так круглые крыши отмечены желтым цветом, треугольные – синим, а плоские – красным.

Если сравнить реализованные алгоритмы обучения между собой, то можно увидеть, что не все три метода справляются хорошо и можно отметить следующие тенденции. Метод сопряженных градиентов в начале обучения сходится значительно быстрее, чем обычный градиентный метод и в итоге дает лучший результат. Генетический же метод сходится гораздо медленнее градиентного метода и метода сопряженных градиентов. Кроме того, генетический метод отличается этапным улучшением суммарной погрешности, то есть при общей тенденции сходимости и медленном уменьшении ошибки, есть моменты довольно резкого уменьшения ошибки. Необходимо отметить, что градиентный метод и метод сопряженных градиентов сошлись к ошибке, которая была ниже установленного порога, поэтому оказались лучше генетического алгоритма. Однако в случае большего количества итераций и, возможно, других данных генетический алгоритм мог оказаться лучше, о чем свидетельствуют результаты, полученные при выполнении работы и теоретические данные [Зайченко, 2004].

Сказанное выше проиллюстрировано на Рис. 3, который показывает зависимость суммарной погрешности от пройденных итераций по каждому из выбранных методов.

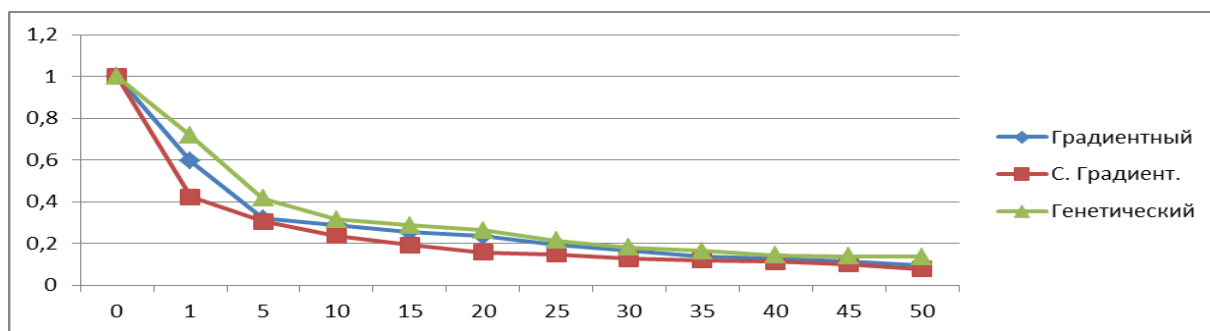


Рис.3. Графическое представление зависимости суммарной погрешности от пройденных итераций

Рассмотрим график на Рис. 4, который показывает процесс обучения нейронной сети для каждого вида крыши, на примере обучения методом сопряженных градиентов, поскольку этот метод дал наилучший результат.

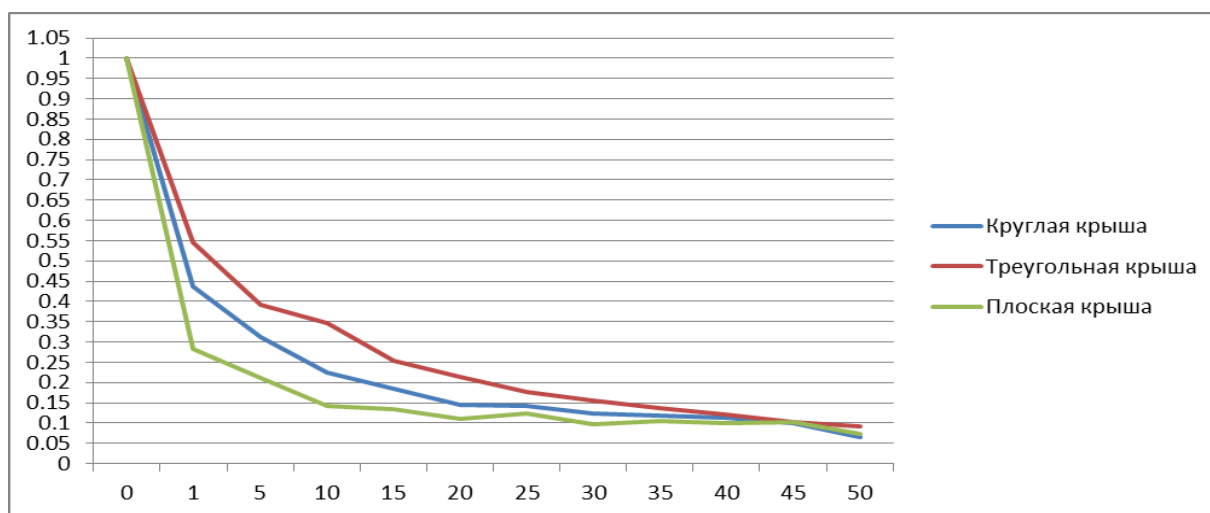


Рис.4. Графическое представление зависимости погрешности от пройденных итераций для каждого вида крыши

Можно увидеть, что наиболее легким примером для обучения нейронной сети были здания с плоскими крышами. Среди зданий этого типа наибольшее количество распознанных. В тоже время треугольная форма крыши была наиболее сложным примером для обучения и дольше других удерживала значение погрешности выше желаемого порога. Также треугольная форма крыши чаще всего встречалась в нераспознанных образцах, что можно объяснить более сложной геометрической формой.

Для эксперимента было отобрано 90 изображений – 60 для обучения и 30 для тестирования. Результаты обучения и классификации приведены в Табл. 1.

Алгоритмы обучения весовых коэффициентов	Обучение		Тестирование	
	Величина ошибки	% ошибочно классифицированных	Величина ошибки	% ошибочно классифицированных
Генетический	0,13	0	0,19	23,3
Градиентный	0,09	0	0,12	13,3
Сопряженного градиента	0,07	0	0,09	6,6

Табл. 1. Результаты обучения и тестирования

По полученным результатам эксперимента можно сделать вывод, что наилучший результат был получен при обучении с помощью метода сопряженных градиентов. Необходимо отметить, что система не всегда правильно определяет тип крыши, что говорит о необходимости коррекции используемых признаков. Кроме того, наличие вокруг дома на близком расстоянии в большом количестве деревьев увеличивает процент ошибочно классифицированных образцов. Это верно как для критерия количества неправильно классифицированных символов, так и по скорости обучения. При тестировании было ошибочно классифицировано 14,4 % из проверочного набора данных.

Выводы и перспективы дальнейших исследований

Практическая реализация модели NEFClass показала, что эта система дает достаточно высокие результаты применительно к задаче распознавания зданий на спутниковых снимках. Программное обеспечение, разработанное при выполнении работы, показывает основные возможности нечетких нейронных сетей типа NEFClass и может служить базой для развития темы распознавания зданий в дальнейшем. Кроме того, разработанное программное обеспечение может быть использовано для проведения экспериментов по совершенствованию построения и обучения нечетких нейронных сетей. Реализовав на практике градиентный, генетический алгоритмы, и алгоритм сопряженных градиентов, и проанализировав полученные данные, можно заключить, что лучшим для решения поставленной перед системой задачи оказался метод сопряженных градиентов, который быстро сходится и дает наилучшие результаты. Необходимо отметить, что требуются дальнейшие исследования, направленные на усовершенствования нейросетевой архитектуры и методов ее обучения, об этом говорит достаточно высокий процент ошибочно классифицированных зданий. Это улучшение может быть получено путём коррекции признаков, которые поступают на вход сети и увеличения обучающей выборки. Также возможно улучшение результата путем дообучения базы правил с помощью алгоритма «наилучший для каждого класса» [Зайченко, 2008]. Такой подход позволит снизить процент неверно классифицированных зданий. В заключение следует ещё раз отметить, что по результатам проведенных исследований можно говорить о достаточно высокой результативности применения нечеткого классификатора к задачам распознавания зданий.

Благодарности

Статья частично финансирована из проекта ITHEA XXI Института Информационных теорий и Приложений FOI ITHEA и консорциума FOI Bulgaria (www.ithea.org, www.foibg.com)

Литература

1. Зайченко Ю.П. Основы проектирования интеллектуальных систем. Учебное пособие. – К.: «Издательский Дом «Слово»», 2004. – С. 352
2. Зайченко Ю.П. Нечёткие модели и методы в интеллектуальных системах. Учебное пособие для студентов высших учебных заведений. – К.: «Издательский Дом «Слово»», 2008. – С. 344
3. S. M'uller and D. W. Zaum, 2005, Robust building detection from aerial images, IAPRS, Vol. XXXVI, Part 3/W24, Vienna, Austria.

Информация про автора



Дьяконова Светлана Владимировна – аспирантка Национального технического университета Украины «КПИ», адрес электронной почты: edeta@bigmir.net

Основные сферы научных исследований автора: применение нечеткого классификатора к задачам распознавания зданий на спутниковых снимках.