

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра математической
кибернетики и компьютерных наук

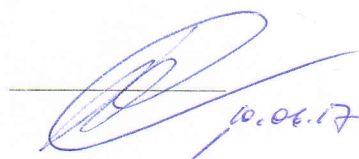

СОВРЕМЕННАЯ СИСТЕМА ИНЖЕНЕРНОГО АНАЛИЗА.
РЕГРЕССИЯ НА УРОВНЕ ПИКСЕЛЕЙ: ГЛУБИННОЕ
ОБУЧЕНИЕ ФИЛЬТРОВ.

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 4 курса 411 группы
направления 010300 — Фундаментальная информатика и информационные
технологии
факультета КНиИТ
Маляровой Марии Виталиевны

Научный руководитель
доцент, к.ф.-м.н.

Заведующий кафедрой
к.ф.-м.н.

В. М. Соловьёв

С. В. Миронов

Саратов 2017

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Основные понятия теории распознавания образов.....	6
2 Разработка структуры метода и его реализация.....	9
2.1 Результаты применения метода в эксперименте	9
2.2 Предикторы и особенности.....	10
ЗАКЛЮЧЕНИЕ	12
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	13

ВВЕДЕНИЕ

При анализе медицинских изображений было предложено множество операций фильтрации для усиления или подавления определенных структур, которые основаны на локальных производных первого или второго порядка изображения. Например, повышение емкости сосудов путем увеличений капли или фильтры для детекции пластин. Существенное различие между этими подходами и представленным в работе методом состоит в том, что подход в работе основан на обучении, то есть это обучение системы с примерами входных изображений и соответствующим желаемым результатом вместо жесткого кодирования связи между входом и выходом, чтобы достичь плавности.

Изучение связей между локальной статистикой изображений и метками в сегментации или же вероятностью того, что пиксел или воксел принадлежит определенному объекту или анатомическим структурам, широко используется в методах классификации по пикселям, [1] таких как белый материал, серое вещество и CSF-классификация по данным МРТ.

Следует подчеркнуть, что берется в рассмотрение не классификация, а регрессия. Регрессия по-своей сути более сложна, чем классификация. Одна из причин этого заключается в том, что в классификации предполагаемый результат может принимать только конечное число дискретных значений [2], тогда как в регрессии результат берется из континуума.

Тот факт, что в методе делаются попытки воспроизводить физически правдоподобные изображения, то есть изображения мягких тканей или кости, подразумевает то, что эта задача сложнее, чем, например, предсказание класса, которому принадлежат пиксели. Даже мягкое выходное изображение классификации не содержит сложные структуры, отображаемые в изображениях мягких тканей и костей. Применение регрессии к фильтрации изображений [3], таким образом, принципиально отличается от предыдущих подходов к фильтрации изображений и еще не получило большого внимания при анализе медицинских изображений. Наиболее близким к подходу является работа.

В этой работе используется массивная тренировочная искусственная нейронная сеть для прогнозирования изображения кости из стандартной рентгенограммы грудной клетки, которую можно вычесть, чтобы получить изоб-

ражение, подобное изображению мягких тканей. В работе сообщается о снижении контраста ребер в рентгенограммах грудной клетки, но не упоминается, как измеряется контраст. В книге [4] эта система используется на небольших участках, содержащих кандидатов в узлы легких, и сообщается о существенном сокращении ложных результатов, полученных с помощью компьютерной системы диагностики.

Как правило, сила таких подходов к нейронным сетям заключается в том, что они обучаются с использованием явного критерия ошибки — обычно среднего квадрата — и, таким образом, производительность сети можно сравнивать напрямую. Однако для нейронных сетей может быть трудно отследить, как происходит реальное предсказание, и поэтому изменение и улучшение такой схемы может быть сложным. Кроме того, обучение нейронной сети может занять много времени и использовать все ресурсы для большей производительности. Более того, типично веса, которые по-сути адаптированы при обучении, должны быть инициализированы случайным образом, и поэтому часто требуется несколько прогонов обучения, прежде чем можно будет принять решение об окончательной и адекватной конфигурации.

Регрессия k -ближайшего соседа, которая применена в методе, не страдает от этих недостатков. Тем не менее, последний, как правило, немного медленнее на этапе тестирования. Метод представленный в этой [1] работе использует хорошо обоснованную и классическую регрессивную методику, которая, кроме того, имеет некоторые хорошо понятные свойства сходимости и аппроксимации.

Метод «компенсации» ребер, использующий простую кусочно-линейную модель для профиля серого уровня, перпендикулярного ребру, который вычитается из входного изображения, предложен в статье [5]. Однако, вычитание реберной структуры не было основным предметом статьи, и, как представлено, метод — это просто идея, которая не была проверена каким-либо образом. Другие известные подходы к подавлению нормальной структуры от рентгенограмм грудной клетки основаны на временных [6] или контралатеральном вычитании [7].

Для применения метода должна быть доступна ранняя рентгенограмма пациента. Это изображение регистрируется на текущей рентгенограмме, а затем вычитается для удаления структур, которые не изменились. Этот

метод имеет потенциал для удаления не только костных структур, но и всех нормальных структур из изображения.

1 Основные понятия теории распознавания образов

Современные информационные системы и информационные технологии содержат в себе огромное количество процедур, которые моделируют или же поддерживают процесс интеллектуального анализа данных. Проецируя интеллектуальный анализ данных на повседневную жизнь людей их визуальное восприятие играет важную роль, помогая в обучении и общении с другими людьми. Обработка этих визуальных данных это по-сути способность к осмыслению образов [8], позволяющая обрабатывать и интерпретировать смысл информации, которую мы получаем благодаря органам зрения. За легкостью этого понятия скрывается весьма сложный процесс обработки информации. Понимание, как мы интерпретируем то, что видим, помогает нам проектировать визуальную информацию.

К простейшим процедурам такого типа, в информационных системах, относится любая классификация количественных данных по заданным пользователем критерием. Более сложные процедуры обеспечивают анализ сцен, процессов, явлений с целью выявления объектов с заданными характеристиками или свойствами.

Процедуры подобного типа присутствуют не только в задачах анализа аэрокосмических изображений, но и при обработке сигналов в технических системах, в медицинской диагностике, биологии, социологии, банковском деле и других областях человеческой деятельности. По мере расширения сферы применения геоинформационных технологий и усложнения процедур геоинформационного моделирования процедуры анализа и классификации совокупностей данных, объектов и структур занимают все более значимое место и в геоинформационных системах нового поколения.

Распознавание образов, как методология принятия решений на основе результатов наблюдений за объектами и процессами окружающего мира, возникло значительно раньше современных компьютерных систем и технологий. Первые методы распознавания разрабатывались для электронных аналоговых систем и рассматривались в рамках теории обработки сигналов. В процессе развития вычислительной техники и информационных технологий эта дисциплина, как и само понятие образа, претерпела значительные изменения и продолжает интенсивно развиваться. Сейчас трудно строго определить класс задач, который относится исключительно к распознаванию образов,

равно как и дать строгое определение самого предмета изучения.

В обыденном понимании образ включает целую совокупность наших индивидуальных ощущений, представлений и умозаключений. Образ — это описание объекта или процесса, позволяющего выделять его из окружающей среды и группировать с другими объектами или процессами для принятия необходимых решений, которые поставлены в первоначальной задаче. Распознавание образов — повседневная неотъемлемая составляющая деятельности человеческого мозга. Поэтому в спектре компьютерных дисциплин задачи распознавания относятся к проблематике искусственного интеллекта. Под искусственным интеллектом будем понимать область компьютерных наук, которая занимается автоматизацией разумного поведения машины. С большей долей вероятности это определение точно соответствует рассматриваемой проблеме искусственного интеллекта. Однако, это определение имеет существенный недостаток, так как само понятие «интеллекта» [8] не очень понятно и четко сформулировано, ведь дать конкретное для оценки предположительной разумной компьютерной программы одновременно отражающей жизнеспособность сложного человеческого разума — комплексно сложно. Итак, проблема определения искусственного интеллекта сводится к проблеме определения интеллекта вообще: является ли он чем-то единым, или же этот термин объединяет набор разрозненных способностей? В какой мере интеллект можно создать, а в какой он существует априори? Что именно происходит при таком создании? Что такое творчество? Что такое интуиция? Можно ли судить о наличии интеллекта только по наблюдаемому поведению, или же требуется свидетельство наличия некоторого скрытого механизма? Как представляются знания в нервных тканях живых существ, и как можно применить это в проектировании интеллектуальных устройств? Что такое самоанализ и как он связан с разумностью?

Вернувшись от проблематики искусственного интеллекта к составляющей образа, можно сказать, что в самом общем случае в качестве образа может рассматриваться любая информационная модель объекта или процесса абстрактного или реального мира. Отличительная особенность такой модели в задаче распознавания — это использование только того подмножества характеристик объектов исследования [9], которое обеспечивает выделение одной или нескольких групп объектов совершенно определенного типа. Целью

процедуры распознавания является ответ на вопрос: относится ли объект, описанный заданными характеристиками, к интересующим нас категориям и если относится, то к какой именно. Для системы обработки информации образ — это совокупность некоторых данных об объекте или явлении, которые включают в себя параметры и связи. Параметры в свою очередь представляют собой количественные характеристики, полученные с помощью измерительных систем и математических моделей. Связи же могут описывать внутреннюю структуру образа или же его особенности поведения, если объект имеет дело с динамикой или проходящим процессом.

Любой алгоритм распознавания можно представить как абстрактную функциональную систему R , состоящую из трех компонент: $R = \{A, S, P\}$, где $A = \{Ak\}$, $k = \{1, \dots, K\}$ — алфавит классов — множество категорий, по которым мы должны распределить наши образы, $S = \{Sj\}$, $j = \{1, \dots, n\}$ — словарь признаков — множество характеристик, из которых составляется описание образа, $P = \{Pl\}$, $l = \{1, \dots, L\}$ — множество правил принятия решения.

Функционирование этой системы сводится к следующему: на вход подается образ — некоторая конфигурация из элементов множества S , к ней применяется определенная последовательность правил из P , в результате конфигурации присваивается индекс, соответствующий одному из элементов множества A . Качество функционирования системы определяется тем, насколько часто присвоенный образу индекс совпадает с ожидаемым нами результатом. Компоненты A , S представляют собой информационную часть системы, а P — методологическую. Ясно, что смысл понятия класс для различных способов описания образов будет различным. В свою очередь, способ описания образа зависит от физической природы объектов распознавания и возможностей формализации соответствующих им понятий.

2 Разработка структуры метода и его реализация

Схема метода, которая является структурой вывода преобразования из входных изображений с использованием регрессии, состоит из пяти этапов.

- предварительная обработка;
- вычисление признаков;
- уменьшение размерности;
- регрессия и итеративная регрессия.

2.1 Результаты применения метода в эксперименте

В этом приложении два выходных изображения (изображение мягких тканей и изображение кости), которые суммируются на вход рентгенограммы грудной клетки. Это предполагает два метода оценки изображения мягких тканей: либо непосредственно напрямую, либо путем оценки изображения кости и вычитания этого изображения из входной рентгенограммы. Первый метод называется прямой оценкой мягкой ткани (STE), второй — косвенной оценкой мягкой ткани (STE).

В дополнение к выбору прямого или косвенного метода, необходимо принять решение об особенностях, которые будут использоваться в качестве предикторных переменных. Кроме того, необходимо решить, следует ли — и если да, то с какими настройками уменьшить линейную размерность, и в какой мере прогноз может улучшиться за счет повторения регрессии.

Производительность метода измерялась посредством стандартной корреляции, то есть применяя коэффициент корреляции Пирсона — r [10] между целевым изображением и заключительным изображением. Различия между методами измеряются парными t -критериями [10] на изображениях и обобщенных p -значениях.

Имея две выборки (в данном случае — целевое и заключительное изображение) $x^m = (x_1, x_2, \dots, x_m)$ и $y^m = (y_1, y_2, \dots, y_n)$, коэффициент корреляции Пирсона рассчитывается по формуле:

$$r_{xy} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2} * \sqrt{\sum_{i=1}^m (y_i - \bar{y})^2}} = \frac{\cos(x, y)}{\sqrt{s_x^2 s_y^2}} \quad (1)$$

где \bar{x} и \bar{y} — выборочные средние, x^m , y^m , $s_x^2 s_y^2$ — выборочные дисперсии.

$$r_x y \in [-1, 1].$$

2.2 Предикторы и особенности

В области компьютерного зрения N -струи [11] обычно вычисляются из представления масштабного пространства L входного изображения $f(x, y)$ и частных производных L , которые используются в качестве основы для выражения различных типов визуальных модулей. Например, алгоритмы для таких задач, как обнаружение признаков, классификация признаков, стереофоническое сопоставление, отслеживание и распознавание объектов, могут быть выражены в терминах N -струй, вычисленных в одной или нескольких масштабах в пространстве масштаба.

Небольшой эксперимент был проведен на одном изображении (при этом система обучалась на оставшихся семи изображениях), в которых изучалось несколько комбинаций Гауссовских N -струй на основе ядра [12] в нескольких масштабах. То есть, в каждой позиции пикселя, в нескольких масштабах, включены функции, полученные с использованием Гауссовых производных n -го порядка.

Чтобы избежать чрезмерного количества фильтров, выбор группы фильтров ограничивался линейными фильтрами Гаусса. В принципе это не накладывает ограничений на тип отношений ввода-вывода, которые могут быть смоделированы. Во-первых, используя Гауссовские N -струи на основе ядра, можно аппроксимировать функцию локально с произвольной точностью, и поэтому всю необходимую входную информацию можно кодировать с помощью этих струй. [12] Во-вторых, нелинейность входного/выходного отношения берется за нелинейный kNNR [2], и из-за этого никакие нелинейные преобразования фильтров не должны быть явно смоделированы.

Для прямой оценки мягкой ткани (STE) окончательный набор функций, используемых для каждой позиции пикселя, состоит из всех гауссовских функций на основе ядра вплоть до порядка 3 в масштабах 1, 2, 4, 8, 16 и 32. В дополнение к этим 60 функциям RAW значение пикселя должно быть включено, что приводит к 61-м входным переменным. Raw — это формат цифровой фотографии, содержащий в себе данные, полученные с фотоматрицы. В таких файлах содержится полная информация о хранимом сигнале, не имеющая четкой спецификации. Обработка Raw-файла позволяет менять параметры кадра (такие как экспозиция (в определённых пределах), яркость,

контрастность, баланс белого, резкость, насыщенность) непосредственно перед преобразованием, как если бы мы делали их перед съёмкой. Это позволяет получить конечное изображение, не потеряв при этом слишком тёмные или слишком светлые участки одного кадра, например, пейзажа с тёмным лесом и ярким небом или малоконтрастного объекта.

Функция же для не прямой оценки мягкой ткани (косвенный STE) — это все гауссовские функции вплоть до порядка 2 на 6 шкалах, логарифмически распределенные между 1 и 64, плюс необработанное значение серого, что приводит к 37 особым характеристикам.

Функции, извлеченные из регрессированного изображения, принимаются за те же функции, что и те, которые были взяты из исходного изображения, за исключением необработанной интенсивности. Это означает, что если не происходит уменьшения размерности, количество признаков [13], используемых для прямой оценки мягкой ткани STE после первоначальной оценки, было равно $61 + 60 = 121$, а для косвенной оценки мягкой ткани STE количество признаков равно 73.

ЗАКЛЮЧЕНИЕ

Таким образом была предложена новая структура фильтрации изображений, основанная на регрессии. Хотя возможны дополнительные улучшения, особенно применительно к конкретным применениям схемы. Представленные качественные и количественные результаты показывают, что этот каркас может выполнять крайне нетривиальное разделение костных и тканевых компонентов в рентгенограммах грудной клетки. Этот метод полностью автоматизирован и дает удовлетворительные результаты даже при данных, поступающих из другого источника.

После прохождения учебного процесса алгоритм может оценить изображение мягкой ткани и кости из любой рентгенограммы грудной клетки без необходимости в фактической двойной энергетической визуализации. Эти оценочные изображения могут использоваться для последующей обработки, например, для компьютерной диагностики. На этапе обучения алгоритма используются фактические пары рентгенограмм грудной клетки и изображения костей и мягких тканей, полученные при двойном энергетическом изображении. Можно также представить метод как моделирование изображений мягких тканей и костей только с помощью стандартной рентгенограммы в качестве входных данных.

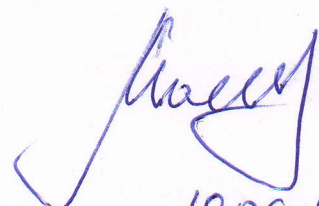
Ключевым вкладом является метод извлечения новых признаков для уменьшения размерности с учетом задач регрессии, который, как было показано, приводит к значительному улучшению производительности.

Здесь важно отметить, что рамки не ограничиваются этим приложением. Он также может использоваться для удаления шума, если схема была обучена с шумными изображениями и соответствующими изображениями с меньшим шумом. Другая возможность — предсказать, как будет выглядеть изображение, если бы оно было получено с другим способом, например, предсказание КТ-изображения из МР-исследования. Значительно сложнее будет схема оценки 3D-данных из 2D проекций (опять же, радиологи человека с их обширными знаниями об анатомии делают это регулярно).

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 *Amato, U.* Segmentation of magnetic resonance brain images through discriminant analysis / U. Amato, M. Larobina, A. Antoniadis, B. Alfano // *Journal of neuroscience methods*. — 2003. — Vol. 131, no. 1. — Pp. 65–74.
- 2 *L. Devroye L. Györfi, G. L.* A Probabilistic Theory of Pattern Recognition / G. L. L. Devroye, L. Györfi. — Springer-Verlag, New York, 1996.
- 3 *Chen, S.* Enhancement of chest radiographs obtained in the intensive care unit through bone suppression and consistent processing / S. Chen, S. Zhong, L. Yao, Y. Shang, K. Suzuki // *Physics in Medicine and Biology*. — 2016. — Vol. 61, no. 6. — P. 2283.
- 4 *Suzuki, K.* Effect of massive training artificial neural networks for rib suppression on reduction of false positives in computerized detection of nodules on chest radiographs / K. Suzuki, J. Shiraishi, F. Li, H. Abe, H. MacMahon, K. Doi // *Progress in Biomedical Optics and Imaging - Proceedings of SPIE* / Ed. by J. Fitzpatrick, J. Reinhardt. — I edition. — 2005. — Vol. 5747. — Pp. 97–103.
- 5 *Vogelsang, F.* Model-based analysis of chest radiographs / F. Vogelsang, M. Kohnen, J. Mahlke, F. Weiler, M. W. Kilbinger, B. B. Wein, R. W. Guenther // *Proc. SPIE*. — 2000. — Vol. 3979. — Pp. 1040–1052.
- 6 *Okumura, E.* Improvement of temporal and dynamic subtraction images on abdominal ct using 3d global image matching and nonlinear image warping techniques / E. Okumura, S. Sanada, M. Suzuki, A. Takemura, O. Matsui // *Physics in Medicine and Biology*. — 2007. — Vol. 52, no. 21. — P. 6461.
- 7 *Li, Q.* Contralateral subtraction: A novel technique for detection of asymmetric abnormalities on digital chest radiographs / Q. Li, S. Katsuragawa, T. Ishida, H. Yoshida, S. Tsukuda, H. MacMahon, K. Doi // *Medical Physics*. — 2000. — Vol. 27, no. 1. — Pp. 47–55.
- 8 *Хоромин,* Проблема «Искусственного интеллекта» / *Хоромин.* — URL: <http://www.structuralist.narod.ru/articles/ai.htm> (Дата обращения 13.06.2017). Загл. с экр. Яз. рус.

- 9 Чабан, Теория и алгоритмы распознавания образов / Чабан. — 2004. — URL: <http://vtiaoai.miigaik.ru/posobiya/20141021125039-9441.pdf> (Дата обращения 13.06.2017). Загл. с экр. Яз. рус.
- 10 Rice, J. Mathematical Statistics and Data Analysis / J. Rice. — Duxbury Press; 3 edition, 1995.
- 11 Koenderink, J. J. Representation of local geometry in the visual system / J. J. Koenderink. — URL: <http://www.vision.caltech.edu/savarese/papers/Presentation%20of%20local%20geometry%20-%20Koenderink%20-%2087.pdf> (Дата обращения 13.06.2017)..Загл. с экр. Яз. англ.
- 12 Florack, L. Image Structure Computational Imaging and Vision, vol. 10. / L. Florack. — Dordrecht, 1997.
- 13 Udupa, J. K. 3D Imaging in Medicine (2nd Edition) / J. K. Udupa. — CRC Press, 2000.


10.06.17