ПРОГРАММНАЯ СИСТЕМА ДЛЯ ОЦЕНКИ ИНФОРМАТИВНОСТИ ПРИЗНАКОВ 2D ИЗОБРАЖЕНИЙ КОСТНЫХ ОБЪЕКТОВ В СУДЕБНО-ЭКСПЕРТНЫХ ИССЛЕДОВАНИЯХ

А.А. Дудкин, В.В. Старовойтов, А.А. Воронов, В.В. Ганченко, Е.Е. Марушко, А.В. Инютин, Л.П. Поденок

Объединенный институт проблем информатики Национальной академии наук Беларуси, ул. Сурганова, 6, 220012, г. Минск, Беларусь, <u>voronov@lsi.bas-net.by</u>

Приведено описание программного обеспечения (ПО) для экспериментального исследования способа анализа разрушенных костных объектов по 2D-данным, описана архитектура ПО, а также результаты анализа разрушений костных объектов. Проведенный анализ на основе оценки информативности позволил выбрать признаки, наиболее подходящие для решения задачи классификации переломов. Результаты будут использованы для судебно-экспертного исследования сложных поверхностей переломов с применением автоматизированной системы анализа цифровых изображений.

Ключевые слова: Костные объекты; информативные признаки; текстурные характеристики; локальные бинарные шаблоны.

SOFTWARE FOR ASSESSING THE INFORMATIVITY OF FEATURES IN THE ANALYSIS OF 2D IMAGES OF BONE OBJECTS IN FORENSIC EXAMINATION

A. Doudkin, V. Starovoitov, A. Voronov, V. Ganchenko, E. Marushko, A. Inytin, L. Podenok

The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, 6 Surhanava Street, Minsk 220012, Belarus, voronov@lsi.bas-net.by

This article describes the software implementation of the system for extracting and evaluating information features from 2D images of bone fractures and bone objects for classifying fractures and identifying the alleged instrument that caused the injury. As parameters, the textural characteristics of Haralick, local binary patterns of pixels for 2D images, Gabor filters, Laws energy texture characteristics for 2D images are considered. The analysis carried out on basis of information content estimation to select the features that are most suitable for solving the problem of bone fractures classification. The results will be used for development of methods for complex forensic examination of complex polygonal surfaces of solid objects for automated system for analyzing digital images.

Keywords: Bone fracture; bone objects; texture features; local binary patterns.

Введение

Существуют разные методы исследования костных объектов и переломов [1-2]. Для исследования строения переломов применим оптический метод. Существуют различного рода автоматизированные системы и экспертные системы для классификации переломов по изображениям костных объектов, однако они обладают определенными недостатками и как правило импортные. Предлагается программная система для выделения информативных признаков по фотографиям переломов костных объектов, что полезно для классификации перелома и идентификации предполагаемого орудия, которым нанесена травма. В предлагаемой системе используется один из подходов к оценке поверхности перелома - выделение информативных признаков.

1. Структура системы

В общем виде разработанную систему можно описать как совокупность следующих подсистем:

– графического интерфейса пользователя, которая используется для отображения данных и взаимодействия с пользователем (ГИП);

 хранения данных, которая выполняет взаимодействие с базой данных (БД) и файловой системой, обеспечивая чтение и запись данных (ПХД);

– обработки и анализа, которая получает от подсистемы хранения данных указанные пользователем данные, осуществляет их обработку и анализ, после чего сохраняет результат, обращаясь к подсистеме хранения (ПОА).



Рисунок 1 – Архитектура системы

Все загруженные в систему файлы (фотографии и 3D-сканы), хранятся в виде файлов на файловой системе. Информация о загруженных в систему данных и их взаимосвязях хранится в виде записей в базе данных (БД).

2. Описание работы системы

Обрабатываемые данные связываются в один объект «Набор данных», который содержит описание добавленных данных, несколько изображений и/или 3D-сканов, результатов их обработки. В общем виде работа системы может быть представлена в виде диаграммы последовательности, приведенной на рисунке 2.



Рисунок 2 – Диаграмма последовательности работы системы

Пользователь, работая с системой, всегда работает с набором данных, каждый из которых имеет свой уникальный идентификатор (ID), описание набора данных (название, время создания, время последнего обновления, текстовое описание и т.п.), а также связанные 2D и 3D данные. Модули системы используют этот ID для конкретизации запроса. Пересылка больших объемов данных (изображение, 3D-скан) осуществляется только между модулем, требующим этой операции, как правило, модули из подсистемы обработки и анализа данных, и подсистемой хранения данных.

3. Анализ 2D изображений костных объектов

В подсистеме обработки и анализа данных содержится реализация алгоритма анализа 2D изображений костных объектов. Для этого реализованы все необходимые функции.

С исходными фотоизображениями, зарегистрированными фотокамерой с разным разрешением, осуществили предварительную подготовку и вырезали прямоугольные фрагменты, содержащие 100% области интереса без фона.

Все имеющиеся исходные изображения были разделены на две группы. Первая группа для анализа была сформирована таким образом, что включала кроме непосредственно фотографий переломов также неповрежденные фрагменты кости черепа.

Во второй группе с фотографиями повреждений черепа согласно классификации основных следообразующих частей твердых тупых предметов (орудия травмы костей свода черепа) были выделены шесть классов повреждений: плоская ограниченная поверхность (прямоугольной, продолговатой или треугольной формы), плоская ограниченная поверхность круглой формы, плоская преобладающая поверхность, цилиндрическая поверхность, сферическая поверхность, предметы с ребром или двугранным углом.

4. Исследование информативности текстурных признаков для первой группы изображений

Были детально исследованы следующие группы текстурных признаков на предмет различения повреждений костей из первой группы: фильтры Габора [3]; энергетические текстурные характеристики Лавса [4]; 13 признаков Харалика на базе матриц смежности значений яркости [5]; локальные бинарные шаблоны пикселов (LBP) [6].

Использовались следующие варианты признаков LBP: радиус = 3, количество шаблонов = 24; радиус = 5, количество шаблонов = 40; радиус = 7, количество шаблонов = 56.

Имя LBP признака содержит его параметры: lbp_r<paquyc>_h<индекс_шаблона>. После вычисления признаков в каждом пикселе формируется гистограмма их значений. Эта гистограмма и рассматривается как набор признаков. Всего вычислялось 26 значений гистограммы LBP радиуса 3; 42 значений гистограммы LBP радиуса 5; 58 значений гистограммы LBP радиуса 7. Всего исследовалось 139 признаков. Фильтры Габора применялись со следующими параметрами: размеры ядра фильтров – 15, 21, 31; углы поворота в градусах – 0, 22, 24, 67, 90, 112, 135, 157.

Имя фильтра описывает его параметры: ks_<pазмер ядра>_th_<угол поворота>. В качестве оценки используется среднее значение матрицы, полученной при фильтрации изображения.

Использовались следующие энергетические текстурные характеристики Лавса: $L_N L_N$, $E_N E_N$, $S_N S_N$, $L_N E_N$, $E_N S_N$, $L_N S_N$, где N – размеры базового вектора, которые могут принимать значения из диапазона [3, 5, 7]. В качестве оценки, также, как и для фильтров Габора, используется среднее значение матрицы, полученной при фильтрации изображения.

Всего исследовался 181 текстурный признак. Оценка информативности признаков выполнялась на основе статистических процедур ANOVA (ANalysis Of VAriance), на базе F-статистики [7], рекурсивного исключения признаков (Recursive feature elimination, RFE) [8] на основе L1 нормы (Linear Suppert Vector Machine, SVM) [9]. Кроме этого использовались классификаторы: ансамбль деревьев, который строит ряд рандомизированных деревьев решений [10], и случайный лес [11]. Выделялись лучшие оценки, найденные каждым методом в отдельности и по сумме нормализованных значений всех признаков (рисунки 3–4, таблицы 1–2).



Рисунок 3 - Суммарные нормализованные оценки информативности всех признаков

Модель ExtraTrees (Extremely Randomized Trees) реализует ансамбль деревьев, который соответствует ряд рандомизированных деревьев решений на различных подвыборках набора данных, и использует усреднение для повышения точности прогнозов и контроля переобучения. ExtraTrees на каждом уровне дерева выбирает разделение признаков случайным образом.

Метод построения случайного леса реализует набор случайно построенных деревьев решений. Случайный лес на каждом уровне дерева выбирает разделение признаков на основании критерия Gini. В двух последних методах строилось по 1000 деревьев. На исследуемых данных оба метода выделили 55 признаков, но их значимость несколько отличалась, что показано в таблице 1.

повреждении поверхности и переломов костеи									
N⁰	Признак	Суммарная оценка							
1	lbp_r3_h6	0.026883							
2	lbp_r5_h7	0.026713							
3	lbp_r3_h5	0.023050							
4	lbp_r7_h0	0.020107							
5	lbp_r5_h29	0.020078							
6	lbp_r3_h17	0.019377							
7	lbp_r7_h9	0.018229							
8	lbp_r7_h7	0.017201							
9	lbp_r5_h30	0.016604							
10	lbp_r3_h18	0.015955							
11	lbp_r5_h6	0.015771							
12	lbp_r7_h29	0.014985							
13	lbp_r7_h27	0.014879							
14	lbp_r7_h8	0.014655							
15	lbp_r3_h23	0.014529							
16	lbp_r3_h4	0.013667							
17	lbp_r7_h31	0.013111							
18	lbp_r7_h6	0.013103							
19	lbp_r3_h19	0.013058							
20	lbp_r7_h28	0.012983							
19 20	lbp_r3_h19 lbp_r7_h28	0.013058 0.012983							

Таблица 1 – Двадцать наиболее информативных признаков для идентификации

В таблице 2 и на рисунке 4 представлены результаты выбора наиболее информативных признаков для повреждений поверхности и переломов костей. На исследованном наборе из 27 изображений 13 костных объектов наиболее информативными оказались признаки типа LBP с разными радиусами.



Рисунок 4 – Суммарные нормализованные оценки информативности всех признаков переломов кости

ANOVA		RFE		LinearSVC L1		ExtraTrees		Случайный лес		Комбинированная оценка	
Признак	Оценка	Признак		Признак	Коэфф. модели	Признак	Оценка	Признак	Оценка	Признак	Оценка
lbp_r3_h17	13.64	ks_21_th_22	1	L7E7	0.0010	lbp_r3_h5	0.028	lbp_r7_h7	0.037	lbp_r5_h7	2.47
lbp_r3_h18	12.06	ks_31_th_22	I	Variance	-0.0010	lbp_r3_h6	0.028	lbp_r7_h0	0.034	lbp_r3_h5	2.37
lbp_r7_h47	11.98	Contrast	-	L7S7	-0.0023	lbp_r5_h7	0.024	lbp_r3_h5	0.032	lbp_r3_h18	2.33
lbp_r5_h29	11.97	S3S3	-	-	-	lbp_r3_h18	0.019	lbp_r7_h6	0.031	lbp_r7_h9	2.26
lbp_r3_h23	11.39	E3S3	I	-	-	lbp_r3_h17	0.018	lbp_r3_h18	0.030	lbp_r3_h6	2.23
lbp_r5_h33	10.96	L3S3	-	-	-	lbp_r7_h0	0.017	lbp_r5_h9	0.029	lbp_r7_h7	2.22
lbp_r7_h9	10.72	S5S5	-	-	-	lbp_r7_h31	0.016	lbp_r7_h9	0.028	lbp_r3_h17	2.20
lbp_r3_h6	10.29	E5S5	I	-	-	lbp_r3_h23	0.015	lbp_r5_h7	0.027	lbp_r7_h0	2.10
lbp_r3_h5	10.26	E7E7	-	-	-	lbp_r7_h47	0.015	lbp_r3_h6	0.027	lbp_r5_h9	2.04
lbp_r5_h7	10.09	E7S7	-	-	-	lbp_r5_h6	0.015	lbp_r5_h6	0.026	lbp_r7_h6	1.85

Таблица 2 – Наиболее информативные признаки, выделенные шестью методами

5. Исследование информативности текстурных признаков для второй группы изображений

Всего исследовался 181 признак с использованием инструментов, описанных выше, но для второй группы изображений. В таблице 3 и на рисунке 5 представлены результаты выбора наиболее информативных признаков для повреждений этой группы.

На исследованном наборе из 45 изображений 6 костных объектов наиболее информативными оказались признаки типа LBP с разными радиусами и текстурная характеристика Харалика Correlation.



Рисунок 5 – Суммарные нормализованные оценки информативности всех признаков поверхности и переломов кости

ANOVA		RFE		LinearSVC L1		ExtraTrees		Случайный лес		Комбинирован ная оценка	
Признак	Оценка	Признак		Приз нак	Коэф. моде ли	Признак	Оценк а	Признак	Оценк а	Признак	Оценк а
lbp_r5_h 33	24,5 7	ks_21_th_1 57	1	L7E 7	0.004 0	Correlation	0,01 3	lbp_r7_h5	0,018 5	lbp_r5_h 34	2,78
lbp_r5_h 34	23,7 3	ks_31_th_1 57	-	L5L 5	0.000 9	lbp_r5_h30	0,01 3	L7L7	0,018 3	lbp_r7_h 50	2,58
lbp_r7_h 50	23,3 8	Contrast	-	L7S 7	0.000 7	lbp_r5_h31	0,01 2	lbp_r5_h6	0,018 2	lbp_r5_h 33	2,55
lbp_r3_h 17	22,6 3	SumAverag e	-	L7L 7	0.000 6	lbp_r5_h29	0,01 2	lbp_r5_h34	0,018 1	lbp_r3_h 18	2,55
lbp_r3_h 18	22,3 3	\$5\$5	-	-	-	lbp_r7_h45	0,01 1	DifferenceVarian ce	0,018 1	lbp_r5_h 6	2,50
lbp_r7_h 51	22,1 8	L5S5	-	I	-	lbp_r5_h34	0,01 1	Correlation	0,017 9	lbp_r7_h 5	2,43
lbp_r5_h 6	21,5 9	E7E7	-	-	-	lbp_r3_h18	0,01 1	lbp_r7_h50	0,017 3	lbp_r3_h 19	2,28
lbp_r5_h 30	21,5 5	S7S7	1	-	-	SumAverage	0,01 0	lbp_r5_h33	0,016 6	lbp_r7_h 51	2,28
lbp_r5_h 35	21,0 8	E7S7	-	-	-	SumVariance	0,01 0	Variance	0,015 6	lbp_r3_h 17	2,25
lbp_r7_h 48	20,9 1	L7S7	-	-	-	MeasOfCorrelatio n1	0,01 0	MeasOfCorrelatio n2	0,015 5	Correlatio n	2,25

Таблица 3 – Наиболее информативные признаки, выделенные шестью методами для второй группы

Заключение

Разработано программное обеспечение для экспериментального исследования и оценки информативности параметров поверхностей переломов костей по 2D-данным для классификации переломов. В качестве параметров рассмотрены текстурные характеристики Харалика, локальные бинарные образцы. Для анализа повреждений костной ткани по фотографиям реализован выбор наиболее информативных признаков. По результатам экспериментов наиболее информативными оказались признаки типа LBP с разными радиусами и текстурная характеристика Харалика Correlation. Отобранные признаки позволяют выполнять классификацию количественную И оценку повреждений костных объектов по их фотоизображениям.

Библиографические ссылки

- 1. Волегов П.С., Грибов Д.С., Трусов П.В. Поврежденность и разрушение: обзор экспериментальных работ // Физическая мезомеханика. 2015. Т. 18. № 3. С. 11–24.
- 2. Разрушение. Т. 1: Микроскопические и макроскопические основы механики разрушения / Под ред. Г. Либовица. М.: Мир, 1973. 620 с.

- 3. Feichtinger, Hans G. Gabor analysis and algorithms : theory and applications./ Feichtinger, Hans G. and Thomas Strohmer. Boston: Birkhäuser, 1998. 496 p.
- 4. Laws, K. Rapid texture identification / Kenneth I. Laws // Proc. SPIE 0238. Image Processing for Missile Guidance. 1980. San Diego, USA. Vol. 238. P. 376–381.
- 5. Haralick R.M., Shanmugam K., Dinstein I.H. Textural features for image classification // IEEE Trans. on systems, man, and cybernetics, 1973. № 6. P. 610–621.
- 6. Ojala T., Pietikainen M., Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns // IEEE Transactions on pattern analysis and machine intelligence, 2002. Vol. 24. № 7. P. 971–987.
- 7. Наследов Д. Математические методы психологического исследования. СПб.: Речь, 2008. 172 с.
- 8. Kuhn M. Applied Predictive Modeling / Kuhn Max and Kjell Johnson. New York, NY: Springer. 2018. 600 P.
- 9. Ben-Hur, Asa; Horn, David; Siegelmann, Hava; Vapnik, Vladimir N. "Support vector clustering" // Journal of Machine Learning Research. 2001. № 2. P. 125–137.
- 10. Geurts P., Ernst D. and Wehenkel L. Extremely randomized trees // Machine Learning. 2006. Vol.63. P.3–42.
- 11. Breiman L. Random Forests // Machine Learning. 2001. Vol. 45. P. 5–32.