Application of structural and textural features from X-ray images to predict the type of bone fracture treatment

Anam Haq
Poznan University of Technology
Poznan, Poland
anam.haq@put.poznan.pl

Szymon Wilk
Poznan University of Technology
Poznan, Poland
szymon.wilk@cs.put.poznan.pl

ABSTRACT

Analysis of medical images plays a very important role in clinical decision making. For a long time it has required extensive involvement of a human expert. However, recent progress in data mining techniques, especially in machine learning, allows for creating decision models and support systems that help to automatize this task and provide clinicians with patient-specific therapeutic and diagnostic suggestions. In this paper, we describe a study aimed at building a decision model (a classifier) that would predict the type of treatment (surgical vs. non-surgical) for patients with bone fractures based on their X-ray images. We consider two types of features extracted from images (structural and textural) and used them to construct multiple classifiers that are later evaluated in a computational experiment. Structural features are computed by applying the Hough transform, while textural information is obtained from gray-level occurrence matrix (GLCM). In research reported by other authors structural and textural features were typically considered separately. Our findings show that while structural features have better predictive capabilities, they can benefit from combining them with textural ones. Interestingly, there are no statistical differences in overall classification accuracy attained by the classifiers considered in the study (it ranges from 91.0% to 96.1%), however, the most promising one is the random forest.

KEYWORDS

clinical data, X-ray images, classification models, decision support

1 INTRODUCTION

Over the last decades medical image processing has made substantial progress and has attracted attention from researchers belonging to various fields, e.g., mathematics, computer science, engineering, physics, biology and medicine [10]. Information systems that store and process image information (e.g., PACS – picture archiving and communication systems) have become an important component of health IT infrastructure and they are regularly used throughout the patient management process. Moreover, development of various image modalities has resulted in challenges associated with their efficient processing and advanced analysis, also in combination with other available types of information. The later is often referred to as data fusion [4]. In this paper we focus on images obtained from a single modality – X-ray – that represent bone fractures.

Bone fractures constitute the most common type of injury that occurs in clinical practice. Normally, during the examination process the physician identifies the fracture and its type, and then decides how it should be treated properly. In order to examine bone fractures various medical imaging technologies are available – they include X-ray and CT (computed tomography) imaging, with the former being most commonly and widely used for bone examination. The process of manual examination of an X-ray image is very time consuming and tedious, therefore, physicians often make mistakes while inspecting such images [7]. These mistakes may result in inadequate treatment, like unnecessary surgeries. Several studies have shown that surgery is not needed in every case [6]. Moreover, surgical treatment is not only more expensive than non-invasive one, but it is also more painful.

This problem can be addressed by building computer-aided diagnostic (CAD) tools that automatically identify the presence and type (severity) of bone fracture, and then suggest the most appropriate treatment for a given patient. However, we have to keep it in mind that human skeleton consist of different types of bones (short, long, flat, irregular, and sesamoid) [9], therefore designing a decision model or a CAD system that would deal with any fracture is a significant challenge [7]. The reason behind is that every type of bone requires a different type processing workflow involving specialized image analysis algorithms. Because of the difficulty related to this problem intense research has been being conducted in the automated fracture detection and still there is room for improvement. In this paper we limit the scope of the problem by providing support only for the decision related to treatment and by focusing on long bones (arm and leg) and upper pelvic bones.

Clinical decision models that rely on images are largely dependent on segmentation and feature extraction algorithms. Moreover, building any decision model requires medical domain knowledge related to the underlying problem [4]. For example, when detecting a brain tumor in an MRI (magnetic resonance imaging) scan it is important to have information about the nature of the tumor. This domain knowledge is helpful when developing automated approaches for detection of abnormalities and their further diagnosis [12]. Such abnormalities can be defined by by their structural characteristic (e.g., area, thickness, or thinness) or by their textural features (e.g., maximum intensity value, energy, contrast). The use of the Hough transform has been proved useful in detecting fracture bones [4] as it incorporates the structural details of the bones. To perform texture analysis of the bones famous gray-level co-occurrence matrix (GLCM) is widely used, which was introduced by Haralick et al. [5]. This technique is based on the assumption that image texture consists of different regions or sub-regions defined by the characteristics like brightness, color, energy, etc., and that information about these regions may be very useful in image analysis.

Recently, application of deep learning methods (e.g., convolutional neural network) to medical imaging problems has gained a lot of attention [17]. In many problems deep learning has proven to be more efficient then tradition image processing techniques and has raised a question regarding the importance of feature extraction among researchers [17]. However, the main problem with deep learning is that it requires huge amount of data for learning (for example, a learning set considered in the recent competition considered by Kaggle on analyzing fundus images contains tens of thousands of images), and also does provides no insights into its "internals", including the discovered knowledge[17].

As already mentioned above, in this paper we deal with X-ray images of fractured bones. We apply image processing (in particular feature extraction) techniques and machine learning methods to build a decision model that would predict an appropriate type of treatment (surgical vs. non-surgical) a given patient should undergo. Currently, different image processing techniques are used to analyze X-ray images of different types of bones. While the majority of proposed methods focus on a single type of features (either structural or textural) for the identification or classification of fractures, we consider both types of features. In this way we are able to evaluate their impact on the performance of resulting decision models (classifiers) and potential benefits resulting from their synergy.

2 RELATED WORK

Work related to our research comes from the three following areas: (1) pre-processing of X-ray images (especially for noise removal), (2) segmenting bones in these images and (3) extracting features from images. Relevant research work is discussed below:

2.1 Pre-processing

Vijaykumar et al. [15] proposed an algorithm to to remove Gaussian noise present in X-ray images. Their algorithm estimates the presence of noise in image and replace the value of pixels located in the center by mean value of the neighboring pixels based on the threshold value. The proposed filtering algorithm proved to work better on X-ray images as compared to other filtering approaches like Wiener, k-means and bilateral-trilateral algorithms. Another approach for noise removal was presented by Al-Khaffaf et al. [7] where they used k-fill algorithm (calculating the number of black and white pixels in a filter window of 3x3) to eliminate salt and pepper noise. Moreover, Anu et al. [7] used Gaussian filter of size 3x3 to remove the noise when detecting bone fracture in X-ray images. Finally, Chai et al. [1] used Laplacian filters to remove noise from the X-ray while developing algorithm for fracture detection by the help of textural features (GLCM).

2.2 Bone Segmentation

In order to detect the boundaries of an object present in a noisy X-ray image Aishwariya et al. [13] proposed an approach that starts with edge detection using Canny edge detection algorithm, and then applies boundary detection techniques like active contour model or geodesic active contour model. Smith et al. [14] developed a method to detect fractures of pelvic bones. The method uses discrete wavelet transformation for automated segmentation of the bone boundary. The wavelet transformation is followed by a sequence of morphological operations – if at the end as a result a single boundary is detected, this indicates no fracture. On the other hand, if multiple boundaries are detected, then this signals one or more fractures.

2.3 Feature Extraction

Chan et al. in [2] uses three different types of transformations for feature selection i.e., curvelets, wavelets and Haar. Haar performed best as compared to wavelet and curvelet transformation. Aishwariya et al. [13] proposed to use Sobel for detection of bone boundaries, and then uses GLCM features to further detect the presence of bone fractures. This approach was tested on X-ray images and an accuracy of 85% was achieved. However, the most difficult task was the segmentation of bone boundaries. Myint et al. in [11] proposed an algorithm that used edge detection and the Hough transformation to automatically detect fracture. The authors reported that their approach works relatively better on high resolution images. We also used the Hough transform in our earlier study [4] where it was applied to extract structural features from X-ray images. These features were further fused with non-image data coming from patient record in order to develop a therapeutic model.

3 PROPOSED APPROACH

As already discussed, physicians make the decision regarding the treatment (surgical vs. non-surgical) of patients with bone fractures by manually examining their X-ray images which is a tedious process and hence error prone. Our goal is to respond to this challenge by constructing a decision model that would support physicians while making making such decisions.

An outline of our approach to feature selection and classifier construction is given in Figure 3. We considered different approaches for pre-processing of X-ray images (see their description in [7]) and selected the one that is most perceptive in detecting bone edges (visually) for the data set at hand. It starts with pre-processing of an X-ray image by applying a median filter (window size 3x3) for noise removal and contrast enhancement for amplifying bone edges. Then, two parallel branches are initiated – the first one is responsible for extracting structural features, and the second one establishes textural features. Both branches employ several specific image processing techniques and are described in details below. Once values of features have been obtained, they are merged in a single feature vector – all considered features are listed in Table 1. This vector is finally fed into the learning and classification block where a specific classifier is constructed and then applied to new objects (X-ray images also characterized by values of extracted features).

3.1 Extraction of Structural features

We use the Hough transform [11] to extract structural features. This process consists of the following steps (please refer to [4] for its illustration):

(1) The Canny operator is applied to a pre-processed image to detect edges. Moreover, disconnected components are removed from the resulting image,
(2) The Hough transform is applied to detect the bone fracture – the process is explained in detail in [11]. Parameters of the transform are set in such a way that it produces two peaks for minor fractures and more then two peak values for major fractures.

3.2 Extraction of Textural Features

Extraction of textural features employs the GLCM transformation [1]. The required steps are as follows (see also Figure 2):

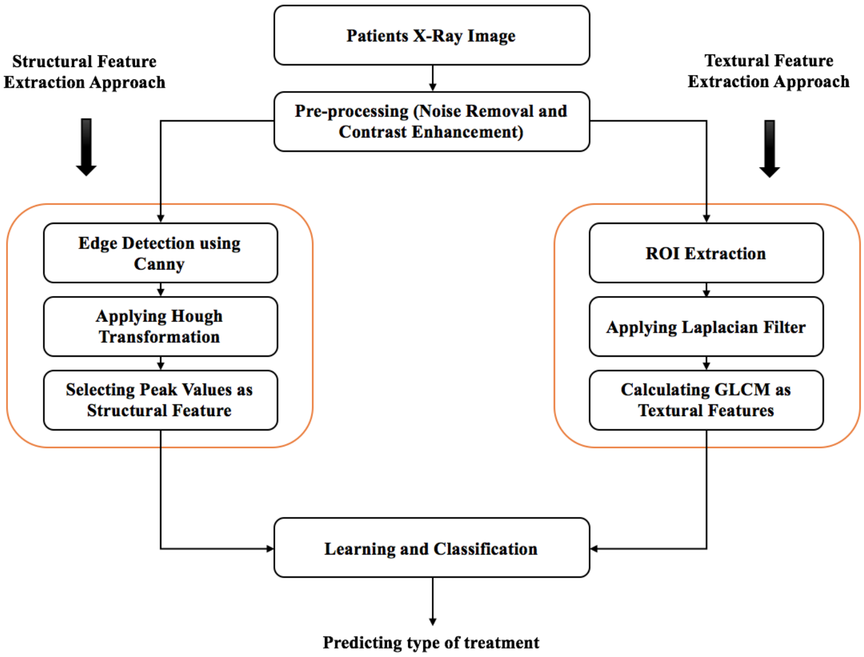

Figure 1: Outline of the proposed approach

Table 1: Description of features (S – structural feature, T – textural feature)

Feature	Type	Description
Hough peak - mean	S	Mean peak value of the Hough transform.
Hough peak - stdev	S	Standard deviation of peak values of the Hough transform.
Contrast	T	Measure of the intensity contrast between the image pixel and its neighbor over the selected ROI, the value of contrast is 0 for a constant image region.
Energy	T	Measure of sum of square of elements present in gray-level co-occurrence vector. The value of energy is 1 for a constant image. It is also known as uniformity of energy and angular second moment.
Homogeneity	T	Measure of the degree of closeness between values in the gray-level co-occurrence matrix. The value of homogeneity is 1 for diagonal GLCM vector.
Correlation	T	Measure the degree of correlation i.e., how the value of a pixel is correlated over the selected region. Value of correlation is 1 for positive image and -1 for the negative image.

(1) Region of interest (ROI) corresponding to the fracture is segmented manually from the pre-processed image,
(2) Laplacian filtering is applied to detect bone boundaries,
(3) The GLCM vector is calculated (specifically, it is obtained by identifying the number of times the pixel *i* occurred in a spatial relationship with pixel *j*), then the textural information like contrast, homogeneity, energy and correlation of an input image is obtained from this 2-Dimensional vector.

4 COMPUTATIONAL EXPERIMENT

4.1 Experimental Design

We implemented our approach using MATLAB (image processing) and WEKA (learning and classification) [16], thus combining advantages (wide choice of powerful image processing and machine learning methods) offered by both tools. This implementation was applied to a set of X-ray images coming from the data repository provided by the Wielkopolska Center of Telemedicine (https://www.telemedycyna.wlkp.pl)– a teleconsultation platform for patients with multiple injuries. The repository includes data of 2030 patients with bone fractures – 1593 (78.5%) underwent a surgery, and the remaining 437 (21.5%) were treated non-surgically. Each patient has a clinical record with non-image data (basic demographics, results of laboratory tests) and a set of 2-5 X-ray images showing fractures at different stages of treatment. From this repository we randomly selected 210 patients – 76 (36.2%) non-surgical and 134 (63.8%) surgical cases. We changed the distribution of classes to make resulting classifiers less biased towards the surgical class, and the obtained

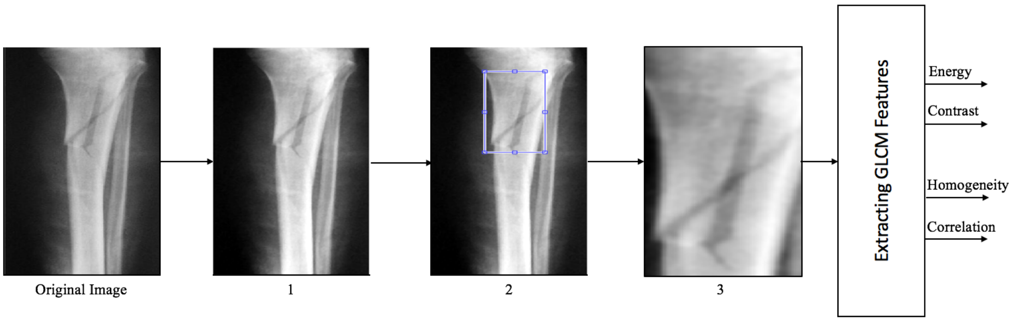

Figure 2: Extraction of textural features from an X-ray image: (1) pre-processing, (2) ROI segmentation, (3) application of Laplacian filter

distribution was established based on suggestions from [3]. Moreover, for each patient we manually selected a single X-ray image representing a fractured bone at the time when the management was initiated.

We obtained values of structural and textural features for each of the selected images and stored resulting feature vectors in an intermediary ARFF data file for further processing in WEKA. In fact, we created three versions of this data file to facilitate subsequent learning: (1) with structural features only, (2) with textural features only and (3) with all features. Extraction of features was performed in MATLAB on a MacBook Pro computer with i5 2.7GHz processor and 8GB of RAM and it took 5.23 minutes to complete.

We then constructed multiple decision models using available data. Specifically, we considered the following classifiers (in brackets with give their symbols used further in the text): a k-nearest neighbor classifier with $k = 7$ (7NN), a naive Bayes classifier (NB), a tree-based classifier induced with the C4.5 algorithm (C45), a rule-based classifier induced using the RIPPER algorithm (RIP), a random forest classifier (RF), an SVM classifier (SVM) and a multilayer perceptron classifier (MLP). Most classifiers were generated using default settings in WEKA (for a more detailed description of corresponding learning algorithms [16]) – only in NB we used supervised discretization, for SVM we used cost equal to 1e+6 and a radial basis kernel function with gamma equal to 0.01, and finally for MLP we specified 3 hidden layers. Such parameters were established during a preliminary evaluation. Here we should also note that we built three versions of each classifier – using structural, textural and all features respectively. While we are aware that building such complex decision models as RF or MLP using only two structural features may be questionable, we did it to maintain consistency of the experimental design.

Classification performance of all classifiers was evaluated in 10 runs of 10-fold cross validation (for better stability of the results) and we use classification accuracy, overall and for both decision classes (surgical and non-surgical). Computations in WEKA were run on the same MacBook Pro as the first part of the experiment (feature extraction) and it took 10.15 minutes to complete them.

4.2 Results

Classification performance of specific classifiers is given in Table 2 where we report overall accuracy along with accuracy values for both decision classes. The best results obtained for each classifier are marked with bold.

The most important observations from Table 2 are the following:

(1) Classifiers using the structural features (mean and standard deviation of peak values obtained using Hough transform) were more accurate than classifiers based on the textural features (obtained from GLCM). The overall accuracy obtained by all classifiers based on the structural features exceeded 90% for all of the considered classifiers.
(2) While the textural features alone resulted in the worse performance for each of the considered classifiers, their combination with the structural feature always improved overall classification accuracy. In fact for all considered classifiers, the best overall accuracy was achieved when using both structural and textural features. A similar observation was made for accuracies in specific classes – the only exceptions were C45 that was more accurate for the non-surgical class when using structural features, and SVM that demonstrated the same performance for structural and all combined features.
(3) The highest overall accuracy (96.1%) was achieved using RF. It also demonstrated the highest accuracy for the surgical class (99.0%). These results confirm the usefulness of ensemble classifiers, in particular RF, in the task of classifying X-ray images reported by other authors [8].

In order to get better insight into captured classification knowledge and thus to enhance explanatory capabilities of our approach we analyzed the importance of features as perceived by specific classifiers. Here we focused on classifiers that are capable of assessing the importance of specific features and considered C45, RIP and RF. In C45 more important features appear higher in the tree, for RIP such features appear more frequently and in stronger rules (i.e., rules with a larger support), and for RF the importance of features is captured by their weights.

According to RIP the most important attributes are structural features used in combination with energy from textural features (see obtained RIP rules in Figure 3). The C45 model gave more importance to structural features and used them in combination with correlation from textural features (see Figure 4). RF assign weights to features showing the most important ones at the top which are standard deviation and mean of peak values from the Hough transform followed by a sequence of textural features – contrast, energy, correlation and homogeneity (see Fig. 5).

We repeated an experiment described in [4] where we applied data fusion techniques (specifically, combination of data) to build

Table 2: Performance of classifiers based on various sets of features (standard deviation given in brackets; S – structural features, T – textural features; ⋆ indicates performance that is statistically worse than performance for all features according to two-tailed T-test)

Classifier	Feature set	Overall [%]	Non-surgical [%]	Surgical [%]
7NN	S	92.5 (6.7)	89.0 (13.0)	94.0 (7.0)
	T	73.4 (8.5)⋆	57.0 (18.0)⋆	83.0 (9.0) ⋆
	S+T	**94.7 (5.1)**	**91.0 (11.0)**	**98.0 (4.0)**
NB	S	89.2 (6.2)	84.0 (15.0)	92.0 (10.0)
	T	76.1 (8.3)⋆	73.0 (21.0)	78.0 (11.0)⋆
	S+T	**92.6 (5.8)**	**91.0 (12.0)**	**94.0 (6.0)**
C45	S	91.4 (6.4)	**92.0 (14.0)**	91.0 (8.0)
	T	77.0 (8.3)⋆	85.0 (18.0)	73.0 (10)⋆
	S+T	**94.0 (6.7)**	89.0 (11.0)	**97.0 (5.0)**
RIP	S	92.5 (6.8)	87.0 (15.0)	95.0 (7)
	T	77.0 (8.7)⋆	77.0 (18.0)	77.0 (11.0)⋆
	S+T	**94.5 (5.6)**	**88.0 (14.0)**	**98.0 (4.0)**
RF	S	91.5 (6.9)	89.0 (12.0)	93.0 (7.0)⋆
	T	77.3 (9.5)⋆	71.0 (19.0)⋆	81.0 (10.0)⋆
	S+T	**96.1 (6.5)**	**92.0 (11.0)**	**99.0 (4.0)**
SVM	S	90.6 (6.5)	**80.0 (17.0)**	**97.0 (5.0)**
	T	80.5 (7.2)⋆	69.0 (16)	81.0 (8.0)⋆
	S+T	**91.0 (6.5)**	80.0 (17.0)	**97.0 (5.0)**
MLP	S	91.5 (6.5)	88.0 (14.0)	94.0 (8.0)
	T	78.4 (8.1)⋆	72.0 (19.0)⋆	82.0 (11.0)⋆
	S+T	**94.8 (5.1)**	**94.0 (9.0)**	**96.0 (6.0)**

Figure 3: Decision rules created for the RIP classifier (the default rule for the surgical class is excluded)

```
(MeanPeakValue <= 61.5) and (StDevPeakValue >= 35.8)
and (StDevPeakValue <= 51.4) => Treatment = non-surg
(61.0/0.0)
(Energy >= 0.21) and (StDevPeakValue >= 36.2)
and (StDevPeakValue <= 57.5) => Treatment = non-surg
(11.0/0.0)
```

Figure 4: A decision tree created for C45

```
MeanPeakValue <= 61.5
|   StDevPeakValue <= 19.078784: surg (12.0)
|   StDevPeakValue > 19.078784
|   |   StDevPeakValue <= 56.568542: non-surg (65.0/1.0)
|   |   StDevPeakValue > 56.568542: surg (6.0)
MeanPeakValue > 61.5
|   StDevPeakValue <= 35.817826: surg (70.0)
|   StDevPeakValue > 35.817826
|   |   StDevPeakValue <= 57.538683
|   |   |   Correlation <= 0.974494: non-surg (10.0)
|   |   |   Correlation > 0.974494: surg (12.0/2.0)
|   |   StDevPeakValue > 57.538683: surg (35.0)
```

Figure 5: Attribute importance based on average impurity decrease in RF

```
0.52 (   377)  StDevPeakValue
0.39 (   235)  MeanPeakValue
0.34 (   122)  Contrast
0.32 (   146)  Energy
0.31 (   136)  Correlation
0.25 (    53)  Homogeneity
```

classifiers based on image and clinical data. In the additional experiment we used an expanded set of image features containing all structural and textural features introduced in this study. We observed their beneficial impact on the performance of classifiers. However, unlike previously the effect of combining image and clinical data was negligible and we hypothesize that for our data set the set of image features are so strong predictor of the type of treatment that additional clinical features become redundant. We are going to further investigate it as part of our ongoing study.

5 CONCLUSIONS

In this paper we presented the results of our study where we have considered structural and textural features extracted from X-ray images. We have used these features to build decision models aimed at predicting a proper treatment (surgical vs. non-surgical) of a patient with bone fracture and evaluated classification performance of these models. Specifically, we checked the following classifiers – k-nearest neighbor (with $k = 7$), naive Bayes, a decision tree, decision rules, a random forest, a support vector machine (with a radial basis function) and a multilayer perceptron. For each of these classifiers we observed an improvement in the overall classification accuracy when using both structural and textural features, and the largest increase occurred for the random forest and naive Bayes classifiers. At the same time, using textural features alone deteriorated the performance in comparison to structural features. Hence we can conclude that the structural features (mean and standard deviation of peak values obtained using Hough transform) have very good predictive abilities and that they may additionally benefit from combining them with the textural features (contrast, energy, homogeneity and correlation).

As future work we will compare the performance of classifiers constructed from extracted features with a convolution network network. We also plan to use more data (e.g., more than one image per patient) and to automate the process of fracture segmentation. We are planning to consider other ensembles in the context of data combining both image and clinical features as this should give ensembles greater flexibility in selecting features for component classifiers. Finally, we would like to implement our approach in form of an educational tool that is deployed on the Wielkopolska Center of Telemedicine platform and used by physicians and medical students to practice their decision making skills. This should also give us an ability to collect new data and experience from users' responses and to use this feedback to improve embedded classifiers.

REFERENCES

[1] Hum Yan Chai, Lai Khin Wee, Tan Tian Swee, and Sheikh Hussain. 2011. Gray-level co-occurrence matrix bone fracture detection. *WTOS* 10, 1 (Jan. 2011), 7–16. http://dl.acm.org/citation.cfm?id=2037119.2037121

[2] Kin-Pong Chan and Ada Wai-Chee Fu. 1999. Efficient time series matching by wavelets. In *Proceedings of 15th International Conference on Data Engineering (Cat. No.99CB36337)*. IEEE, 126–133. https://doi.org/10.1109/ICDE.1999.754915

[3] David J. Dittman, Taghi M. Khoshgoftaar, and Amri Napolitano. 2014. *Selecting the appropriate data sampling approach for imbalanced and high-dimensional bioinformatics datasets*. 304–310. https://doi.org/10.1109/BIBE.2014.61

[4] Anam Haq and Szymon Wilk. 2017. Fusion of clinical data: A case study to predict the type of treatment of bone fractures. In *New Trends in Databases and Information Systems - ADBIS 2017 Short Papers and Workshops, AMSD, BigNovelTI, DAS, SW4CH, DC, Nicosia, Cyprus, September 24-27, 2017, Proceedings*. 294–301. https://doi.org/10.1007/978-3-319-67162-8_29

[5] Robert. M. Haralick. 1979. Statistical and structural approaches to texture. *Proc. IEEE* 67, 5 (May 1979), 786–804. https://doi.org/10.1109/proc.1979.11328

[6] Mounier Hossain, V. Neelapala, and J. G. Andrew. 2008. Results of non-operative treatment following hip fracture compared to surgical intervention. *Injury* 40, 4 (April 2008), 418–421. https://doi.org/10.1016/j.injury.2008.10.001

[7] Irfan Khatik. 2017. A study of various bone fracture detection techniques. *International Journal Of Engineering And Computer Science* 6, 5 (May 2017), 21418–21423.

[8] Seong-Hoon Kim, Ji-Hyun Lee, Byoungchul Ko, and Jae-Yeal Nam. 2010. X-ray image classification using random forests with local binary patterns. In *2010 International Conference on Machine Learning and Cybernetics*, Vol. 6. 3190–3194. https://doi.org/10.1109/ICMLC.2010.5580711

[9] Kenneth J. Koval and Joseph David Zuckerman. 2006. *Handbook of Fractures*. https://books.google.pl/books?id=1x6ZQgAACAAJ

[10] Elizabeth A. Krupinski. 2010. Current perspectives in medical image perception. *Attention, Perception and Psychophysics* 72, 5 (01 Jul 2010), 1205–1217. https://doi.org/10.3758/APP.72.5.1205

[11] San Myint, Aung Soe Khaing, and Hla Myo Tun. 2016. Detecting leg bone fracture in X-ray images. *International Journal of Scientific and Technology Research* 5 (2016), 140–144.

[12] Parveen and Amritpal Singh. 2015. Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM. In *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*. 98–102. https://doi.org/10.1109/SPIN.2015.7095308

[13] R.Aishwariya, M.Kalaiselvi Geetha, and M.Archana. 2014. Computer-aided fracture detection of X-ray images. *IOSR Journal of Computer Engineering (IOSR-JCE)* 2, 1 (2014), 44–51.

[14] Rebecca Smith, Charles Cockrell, Jonathan Ha, and Kayvan Najarian. 2010. Detection of fracture and quantitative assessment of displacement measures in pelvic X-ray images. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. 682–685. https://doi.org/10.1109/ICASSP.2010.5495104

[15] V.R.Vijaykumar, P.T. Vanathi, and P. Kanagasapathy. 2007. Adaptive window based efficient algorithm for removing gaussian noise in gray scale and color images. In *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, Vol. 3. 319–323. https://doi.org/10.1109/ICCIMA.2007.367

[16] Ian H. Witten, Eibe Frank, and Mark A. Hall. 2011. *Data Mining: Practical Machine Learning Tools and Techniques* (3rd ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[17] Guoqiang Zhong, Lina Wang, and Junyu Dong. 2016. An overview on data representation learning: From traditional feature learning to recent deep learning. *The Journal of Finance and Data Science* 2 (2016), 265–278. Issue 4. https://doi.org/10.1016/j.jfds.2017.05.001