Fully Automatic Detection of Distal Radius Fractures from Posteroanterior and Lateral Radiographs

Raja Ebsim[1]([✉]), Jawad Naqvi[2], and Tim Cootes[1]

[1] The University of Manchester, Manchester, UK
{raja.ebsim,tim.cootes}@manchester.ac.uk
[2] Health Education North West School of Radiology, Manchester, UK
naqvi.jawad@gmail.com

Abstract. We describe a fully-automated system for analysing X-rays of the wrist to identify possible fractures. Fractures of the distal radius in the wrist are estimated to be about 18% of the fractures seen in adults and 25% of those seen in children. Unfortunately such fractures are amongst the most frequently missed by doctors in Emergency Departments (EDs). A system which can identify suspicious areas could reduce the number of misdiagnoses. We automatically locate the outline of the radius in both posteroanterior (PA) and lateral (LAT) radiographs, then use shape and texture features to classify abnormalities. We show for the first time that fractures can be better identified in the lateral view, and that combining information from both views leads to an overall improvement in performance.

Keywords: Image analysis · Image interpretation and understanding · X-ray fracture detection · Wrist fractures · Radius fractures · Machine learning

1 Introduction

Fractures of the wrist are usually identified in Emergency Departments by doctors examining lateral (LAT) and posterioanterior (PA) radiographs (Fig. 1). Unfortunately missing such fractures is one of the most common diagnostic errors in EDs, leading to delayed treatment and more suffering for the patient [7,14,18]. This is mainly because the majority of patients attending EDs are seen by junior doctors [8,13]. This problem is widely acknowledged, so in many hospitals X-rays are reviewed by an expert radiologist at a later date - however this can lead to significant delays on missed fractures which can have an impact on the eventual outcome. To address this we are developing a system which can automatically analyse radiographs of the wrist in order to identify abnormalities and thus prompt clinicians, hopefully reducing the number of errors.

In this paper we describe a fully-automated system for detecting radius fractures in PA and LAT radiographs. For each view, a global search [11] is performed

for finding the approximate position of the radius. The detailed outline of the bone is then located using a Random Forest Regression Voting Constrained Local Model (RFCLM) [10]. We use features derived from the shape and texture to train random forest classifiers on the task of detecting fractures. Features from both views are combined for better performance.

This paper is the first to show an automatic system for identifying fractures from lateral view radiographs of the wrist. We show that better performance can be achieved from this view than the PA view, and that further improvement can be obtained by combining results from both views.

2   Background

Distal radius fractures have been on increase in all age groups [15]. They alone constitute around 18% of the fractures seen in EDs in adults [4,6] and 25% of the fractures seen in children [6].

Previous work on detecting fractures in X-rays has been done on a variety of anatomical regions, including arm fractures [19], femur fractures [1,9,12,17,20], and vertebral endplates [16]. Cao *et al.* [2] used stacked random forests to fuse different feature representations to identify fractures in a range of anatomical regions. They achieved a sensitivity of 81% and precision of 25%. All reported work on automatical analysis of wrist fractures that we are aware of [5,9,12] uses only the PA view. In each case some form of shape model is used to locate the outline of the bones, then texture (and possibly shape) features are used to train a classifier to distinguish healthy from fractured bones.

[9,12] use both active shape models and active appearance models [3] to locate the approximate contour of the radius. They extract various texture features (Gabor, Markov Random Field, and gradient intensity) and classify with a Support Vector Machine SVM. They achieved encouraging performance (accuracy $\approx$ sensitivity $\approx 96\%$) but were working on a rather small dataset with only 23 fractured examples in their test set. In previous work [5] we used RFCLMs to segment both the radius and ulna in PA views and trained random forest classifiers on statistical shape parameters and eigen-mode texture features. The automatic system [5] achieved a performance of 88.6% (Area under Receiver Operating Characteristic Curve AUC) on data set of 409 radiographs including 199 fractures. The system we describe below is for the radius in both PA and LAT views achieving better performance and is tested on a data set about twice the size.

3   Method

The outline of the radius was manually annotated with 48 points in the PA view and with 64 points in the LAT View (Fig. 1). For each view, a statistical shape model and an RFCLM [10] object detection model were built from the corresponding manual annotations. The models then used to segment the bone on new radiographs automatically.

3.1   Modeling and Matching

Building Models for Shape and Texture. The outline of the radius is modeled by a linear statistical shape model [3].

Each training image is annotated with n feature points, (x_i, y_i). A $2n$-D vector $\mathbf{x} = (x_1, ..., x_n, y_1, ..., y_n)^T$ represents all the points.

PA

LAT

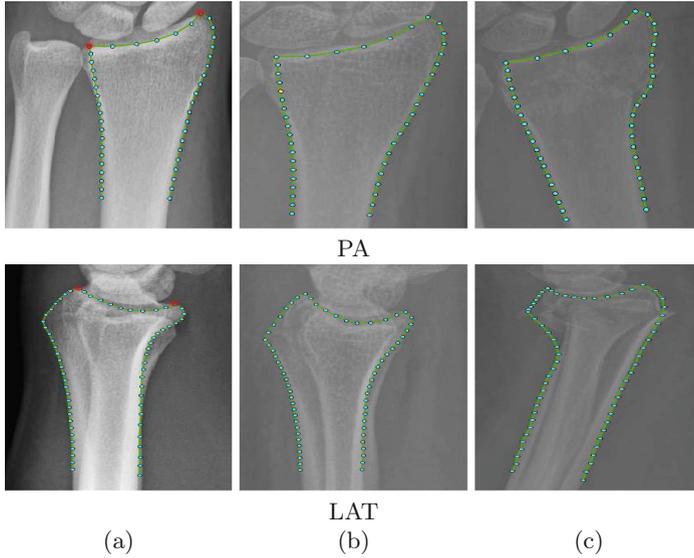(a)   (b)   (c)

Fig. 1. The annotations of both views on (a) a normal radius, (b) a radius with a subtle fracture, and (c) with an obvious fracture. The two red points, appearing on each view, are the anatomical points found by the global searcher for that view. (Color figure online)

Shapes are modelled using

$$\mathbf{x} = T(\bar{\mathbf{x}} + \mathbf{Pb} : \theta) \tag{1}$$

where $\bar{\mathbf{x}}$ is the mean shape, $\mathbf{P}$ is the set of the eigenvectors corresponding to the t highest eignvalues of the covariance matrix, $\mathbf{b}$ is the vector of shape parameters and $T(. : \theta)$ applies a similarity transformation with parameters θ. The first three modes of the built shape models are shown in Fig. 2. Similarly, statistical texture models [3] are built by applying PCA to vectors of normalised intensity ($\mathbf{g}$) sampled from the regions defined by the points of the shape model.

$$\mathbf{g} \approx \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g \tag{2}$$

After fitting the model to a new radiograph, shape parameters $\mathbf{b}$ (in Eq. 1) and the texture parameters $\mathbf{b}_g$ (in Eq. 2) are used as features on which classifiers are trained to distinguish between normal and fractured bones.

Matching Shape Models to New Radiographs. We use a technique similar to that described by Lindner *et al.* [11] to locate the outline of the targeted bone in each view separately. A single global model per view is trained to initially find approximate position of a box containing two anatomical landmarks (i.e. the red points in Fig. 1). As in [11] a random forest regressor with Hough voting is trained to find the displacement between the center of a patch and the object center. During training, patches are sampled at random displacements and scales from the object center and fed to a Random Forest to learn the relationship between an image patch and the displacement of the patch centre from the target position. By scanning a new image at different scales and orientations with the Random Forest and collecting the votes, the most likely center, scale and orientation of the object can be found. The two points estimated by the global searcher are used to initialise a local search for the outline of the radius. We used a sequence of three RFCLMs of increasing resolution to produce the final result.

3.2   Classification

After performing the full automatic search so that radius outline points are located on all radiographs, Random Forest classifiers (100 trees each) are trained on features derived from the shape (the shape parameters, **b**) and the texture (the texture model parameters, $\mathbf{b}_g$) for the task of fracture detection (i.e. normal or fractured). A series of cross validation experiments were performed with different combinations of features for each view separately, then for both views together.

4   Experiments

Data. The experiments were carried out on a dataset containing both views for 787 adult patients (378 of whom had fractures) from two local EDs gathered and anonymised by a clinician.

Automatic Annotation. In order to generate the automatic annotation for the whole dataset, we divided PA radiographs into two subsets, training models on one subset and applying them to the second. Four subsets were needed to successfully learn representative models for the LAT view. That is because of the overlap between the two bones, (i.e. radius and ulna) on lateral view can take various orientations due to different acceptable positioning in practice [6]. This is not the case for PA view as the two bones appear side by side. Figure 3 shows some examples from our LAT dataset.

The accuracy of the segmentation is calculated as the percentage of mean point-to-curve distance [11] to a reference width, and converted to mm by assuming a mean reference width of 25 mm for the radius in the PA view and 20 mm in the LAT view. The reference width of a view is the distance between the two reference points for that view (see Fig. 1(a)).
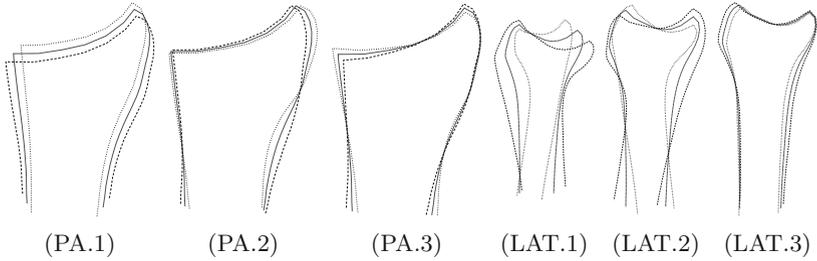
(PA.1)   (PA.2)   (PA.3)   (LAT.1)   (LAT.2)   (LAT.3)

Fig. 2. The first three modes of the shape models of the radius.

Fig. 3. Different relative radius-ulna positions appearing in lateral radiographs.

The results in Table 1 show the ability of the models to successfully segment the targeted structure even when fractured. The mean error was less than 1.4 mm for more than 95% of the radiographs in the LAT view and less than 0.6 mm for 95% of radiographs in the PA view. The PA error was less that reported in [5] (0.61 mm vs 0.78 mm for the 95%-ile, though our dataset is twice the size). The table also breaks down the results by class and shows that although in the PA view there is almost no difference in accuracy between fractured and normal cases, in the LAT view errors are roughly 50% larger. However overall the system can successfully capture a good approximation to both the normal and fractured shapes.

Classification. For each view we performed 5-fold cross validation experiments with Random Forest classifiers using 100 trees (repeated three times) on: (i) shape parameters only, (ii) texture parameters only, and (iii) the concatenation of shape and texture parameters. The results obtained from the PA view, shown in Table 2, reflects the small difference between the manual and automatic annotation. The performance of the automated system improves slightly on that described in [5] (though that was on a smaller dataset).

Table 3 shows the classification results for the LAT view. Note that classifying using the shape alone gives significantly better performance than that from the PA view. Adding the texture information makes only a small improvement (in the manual case) and slightly degrades performance of the fully automated system. The difference in performance for shape between manual and automatic

Table 1. The mean point-to-curve distance in (mm) for the fully automatic annotations.

View	Class	Mean	Median	90%	95%	99%
PA	Normal	0.18	0.10	0.40	0.50	1.08
PA	Fractured	0.18	0.11	0.47	0.63	1.04
PA	Both	0.18	0.10	0.42	0.61	1.06
LAT	Normal	0.40	0.28	0.76	0.94	1.86
LAT	Fractured	0.62	0.44	1.27	1.62	2.83
LAT	Both	0.50	0.32	1.01	1.37	2.34

Table 2. AUC for classification based on PA view.

PA view	Manual	Fully automated	Automated result from [5]
Shape	0.847 ± 0.004	0.826 ± 0.002	0.816 ± 0.007
Texture	0.896 ± 0.005	0.891 ± 0.001	**0.881 ± 0.004**
Shape & Texture	**0.898 ± 0.002**	**0.897 ± 0.002**	0.868 ± 0.002

results suggests that classification performance can be improved by improving the accuracy of the search (perhaps by increasing the size of the training set).

Table 3. AUC for classification based on lateral view.

LAT view	Manual	Fully automated
Shape	0.933 ± 0.001	**0.905 ± 0.003**
Texture	0.894 ± 0.002	0.878 ± 0.002
Shape & Texture	**0.937 ± 0.001**	0.899 ± 0.001

Combining information from both views results in the best classification performance in both manual and automatic cases. See Table 4. Figure 4 shows the ROC curves for the best results.

Table 4. AUC for classification based on features from both views.

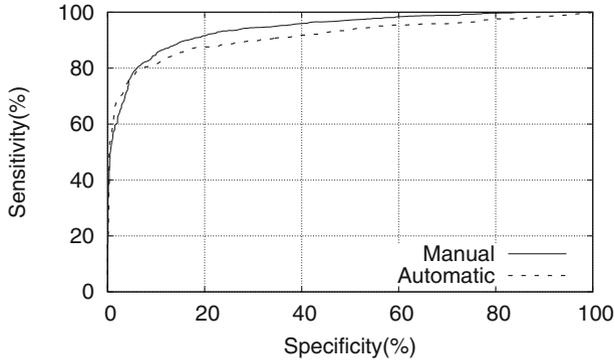| PA & LAT views	Manual	Fully automated
Shape	0.933 ± 0.001	0.913 ± 0.003
Texture	0.918 ± 0.002	0.890 ± 0.006
Shape & Texture	**0.942 ± 0.002**	**0.914 ± 0.004**

Fig. 4. The ROC curves corresponding to classification based on combining shape and texture features from both PA and LAT views for: (i) manual annotation, and (ii) automatic annotation.

5 Conclusions

This paper presents a system that automatically locates the outline of radius in both posteroanterior and lateral radiographs and extracts discriminative features for fracture detection. In future work, we will extend our current work to generate automatic description of the found fracture (i.e. fracture classification), and will explore learning alternative texture features. Our long term goal is to build a system which works well enough to help clinicians in EDs make more reliable decisions.

References

1. Bayram, F., Çakirolu, M.: DIFFRACT: diaphyseal femur fracture classifier system. Biocybern. Biomed. Eng. **36**(1), 157–171 (2016)
2. Cao, Y., Wang, H., Moradi, M., Prasanna, P., Syeda-Mahmood, T.F.: Fracture detection in x-ray images through stacked random forests feature fusion. In: International Symposium on Biomedical Imaging (ISBI 2015), pp. 801–805, April 2015
3. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 681–685 (2001)
4. Court-Brown, C.M., Caesar, B.: Epidemiology of adult fractures: a review. Injury **37**, 691–697 (2006)
5. Ebsim, R., Naqvi, J., Cootes, T.: Detection of wrist fractures in X-Ray images. In: Shekhar, R., et al. (eds.) CLIP 2016. LNCS, vol. 9958, pp. 1–8. Springer, Cham (2016). doi:10.1007/978-3-319-46472-5_1

6. Goldfarb, C.A., Yin, Y., Gilula, L.A., Fisher, A.J., Boyer, M.I.: Wrist fractures: what the clinician wants to know. Radiology **219**, 11–28 (2001)
7. Guly, H.R.: Injuries initially misdiagnosed as sprained wrist (beware the sprained wrist). Emergency Med. J. (EMJ) **19**, 41 (2002)
8. Lee, C., Bleetman, A.: Commonly missed injuries in the accident and emergency department. Trauma **6**, 41–51 (2004)
9. Lim, S.E., Xing, Y., Chen, Y., Leow, W.K., Howe, T.S., Png, M.A.: Detection of femur and radius fractures in x-ray images. In: Proceedings of 2nd International Conference on Advances in Medical Signal and Information Processing, vol. 1, pp. 249–256 (2004)
10. Lindner, C., Bromiley, P.A., Ionita, M.C., Cootes, T.: Robust and accurate shape model matching using random forest regression-voting. IEEE Trans. Pattern Anal. Mach. Intell. **37**(9), 1862–1874 (2015)
11. Lindner, C., Thiagarajah, S., Wilkinson, J.M., Consortium, T., Wallis, G.A., Cootes, T.F.: Fully automatic segmentation of the proximal femur using random forest regression voting. Med. Image Anal. **32**(8), 1462–1472 (2013)
12. Lum, V.L.F., Leow, W.K., Chen, Y., Howe, T.S., Png, M.A.: Combining classifiers for bone fracture detection in X-ray images, vol. 1, pp. I-1149–1152 (2005)
13. McLauchlan, C.A., Jones, K., Guly, H.R.: Interpretation of trauma radiographs by junior doctors in accident and emergency departments: a cause for concern? J. Accid. Emerg. Med. **14**(5), 295–298 (1997)
14. Petinaux, B., Bhat, R., Boniface, K., Aristizabal, J.: Accuracy of radiographic readings in the emergency department. Am. J. Emerg. Med. **29**, 18–25 (2011)
15. Porrino, J.A., Maloney, E., Scherer, K., Mulcahy, H., Ha, A.S., Allan, C.: Fracture of the distal radius: epidemiology and premanagement radiographic characterization. Am. J. Roentgenol. (AJR) **203**, 551–559 (2014)
16. Roberts, M.G., Oh, T., Pacheco, E.M.B., Mohankumar, R., Cootes, T.F., Adams, J.E.: Semi-automatic determination of detailed vertebral shape from lumbar radiographs using active appearance models. Osteoporos. Int. **23**(2), 655–664 (2012)
17. Tian, T.P., Chen, Y., Leow, W.K., Hsu, W., Howe, T.S., Png, M.A.: Computing neck-shaft angle of femur for X-Ray fracture detection. In: Petkov, N., Westenberg, M.A. (eds.) CAIP 2003. LNCS, vol. 2756, pp. 82–89. Springer, Heidelberg (2003). doi:10.1007/978-3-540-45179-2_11
18. Wei, C.-J., Tsai, W.-C., Tiu, C.-M., Wu, H.-T., Chiou, H.-J., Chang, C.-Y.: Systematic analysis of missed extremity fractures in emergency radiology. Acta Radiologica **47**, 710 (2006)
19. Jia, Y.: Active contour model with shape constraints for bone fracture detection. In: International Conference on Computer Graphics, Imaging and Visualisation (CGIV 2006), vol. 3, pp. 90–95 (2006)
20. Yap, D.W.H., Chen, Y., Leow, W.K., Howe, T.S., Png, M.A.: Detecting femur fractures by texture analysis of trabeculae. In: Proceedings - International Conference on Pattern Recognition vol. 3, pp. 730–733 (2004)