

Introduction

Osteoporotic fractures are associated with significant morbidity, mortality and public health costs, and will increase with an ageing population. Many osteoporotic vertebral fractures (VF) present on images do not come to clinical attention or lead to fracture prevention treatment. Furthermore, DXA vertebral fracture assessments (VFA) are often reported subjectively. VFA computer-aided systems offer potential advantages. Methods based on statistical shape models (e.g. active appearance models, AAMs) have been used to segment vertebrae in radiographs and DXA VFA. However, results achieved using AAMs exhibit significant numbers of large errors due to model fitting failure, particularly on more severely fractured vertebrae. We evaluate an alternative, Random Forest Regression Voting Constrained Local Models (RFRV-CLMs), which have proved more robust and generalizable than AAMs in landmark annotation on clinical images. We investigate whether this will reduce the number of fitting failures in vertebral segmentation.

Method

Statistical shape model training begins with manual annotation of a set of landmark points on a large number of training images (Fig. 1), outlining the structures of interest. Images are aligned into a reference coordinate system using a similarity transformation, to remove non-shape variation such as translations and rotations. The coordinates of the landmark points in a training image, in this reference frame, are then concatenated into a single vector \mathbf{x}_i , which specifies a point in a high-dimensional shape space. The complete set of training images produces a cloud of points in shape space. The mean and covariance matrix of this distribution constitute a statistical model of the shape of the annotated structures.

The modes of variation, i.e. the axes of the covariance matrix of the distribution in shape space, are identified using Principle Component Analysis (PCA). A linear model can then be constructed, such that any point within the distribution can be represented by a combination of these modes, multiplied by a vector of parameters,

$$\mathbf{x}_i = T_\theta (\boldsymbol{\mu} + \mathbf{P}\mathbf{b}_i + \mathbf{r}_i)$$

where $\boldsymbol{\mu}$ is the mean shape (i.e. the mean of the distribution in shape space), \mathbf{P} is a matrix of the modes of variation, \mathbf{b}_i is a vector of shape parameters, and \mathbf{r}_i are residuals, allowing small variations from the model. This model can then be fitted to a query image to locate shapes similar to those in the training images; T_θ is a similarity transformation with parameters θ that locates the model within the query image.

Algorithms based on statistical shape models vary in how they use intensity data from the training images. The Active Appearance Model (AAM) applies the same technique of linear model building to the intensities in a patch of image data covering the landmarks. This generates a single model of both shape and appearance covering all landmarks. However, such holistic models have been shown [3] to generalise poorly which, in the context of vertebral fracture, results in significant numbers of model fitting failures on the most severely fractured vertebrae. More recent algorithms, such as the Random Forest Regression Voting Constrained Local Model (RFRV-CLM; [1]), build intensity models for each individual landmark, and use the shape model as a constraint during fitting. In RFRV-CLM, the intensity model is a Random Forest i.e. a binary decision tree trained on image features sampled from a region around the landmark. Such models have been shown to be more robust to shape variation than AAMs in the location of landmarks on facial images [3]. We investigate the hypothesis that this robustness will reduce the number of fitting failures during vertebral segmentation on DXA VFA images.

320 DXA VFA images obtained on various Hologic (Bedford, MA) scanners had manual annotations of 405 landmark points of vertebrae T7 to L4, with fracture classifications from an expert radiologist. RFRV-CLMs were applied to these data in a leave-1/4-out fashion. Errors were calculated as the mean, across each vertebra, of the minimum distances between the automatic annotations and a Bezier spline passing through the manual annotations. The results were compared to those presented in [2], which applied AAMs to the same task and data set (Fig. 2, Table 1), to compare the performance of RFRV-CLM to that of AAM in terms of random errors on the results, and the proportion of fitting failures.

Conclusion

Errors on automatic vertebral segmentations come from two distributions; random errors on successful fits, and systematic errors from fit failures, where the model fails to locate the vertebrae. The results indicate that the AAM achieves a slightly smaller random error (mean error of 0.60mm across all vertebrae, vs. 0.65mm for RFRV-CLM). However, the difference is small compared to the height reductions caused by osteoporotic vertebral fractures. In contrast, the RFRV-CLM achieves 66% fewer large errors due to fitting failures across all classifications. Each fitting failure represents an image for which manual inspection will be required to correct the segmentation prior to measurement of vertebral height ratios and classification of fracture grade. We conclude that the RFRV-CLM is more reliable for automatic analysis of DXA VFA.

This publication presents independent research supported by the Health Innovation Challenge Fund (grant no. HICF-R7-414/WT100936), a parallel funding partnership between the Department of Health and the Wellcome Trust. The views expressed in this publication are not necessarily those of the Department of Health or Wellcome Trust.

References

1. P.A. Bromiley, J.E. Adams and T.F. Cootes. "Localisation of Vertebrae on DXA Images using Constrained Local Models with Random Forest Regression Voting". Proc. MICCAI 2014.
2. M.G. Roberts, T.F. Cootes and J.E. Adams. "Automatic Location of Vertebrae on DXA Images using Random Forest Regression Voting", Proc. MICCAI 2012 pp. 361-368.
3. P. Sauer, T.F. Cootes and C.J. Taylor. "Accurate Regression Procedures for Active Appearance Models". Proc. BMVC 2011 pp. 30.1-30.11.

Example Model Fits

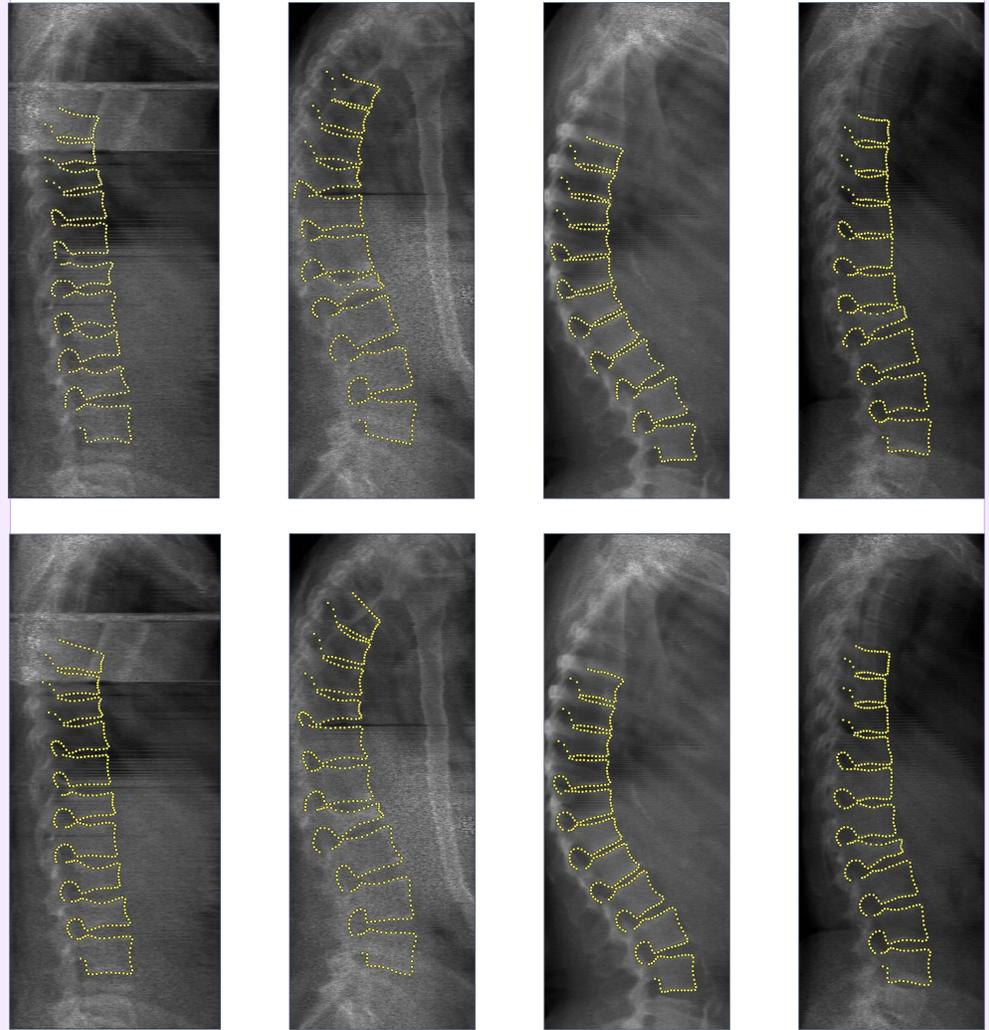


Figure 1. DXA VFA images with (top row) 405-point manual annotations and (bottom row) RFRV-CLM automatic annotation.

Results

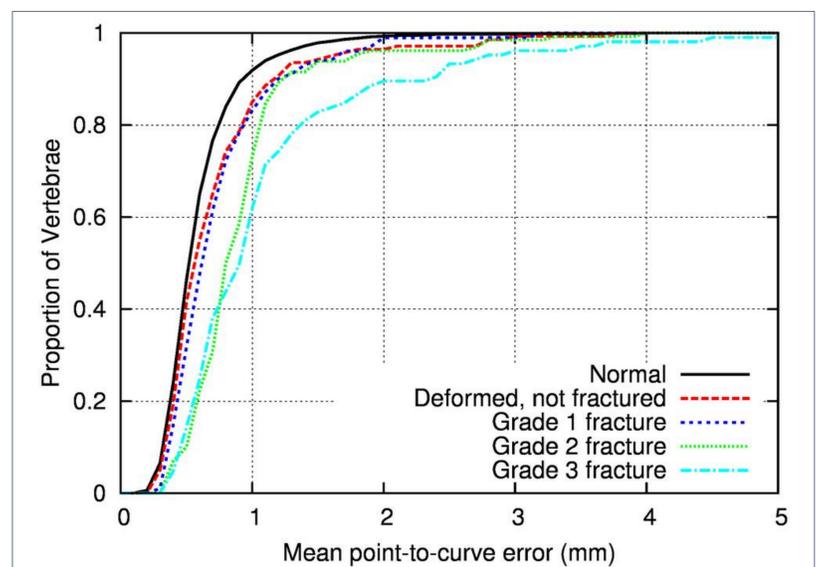


Figure 2. Cumulative distribution function (CDF) of errors in ten vertebral levels in all 320 images, for each vertebral classification. Mean errors of <4mm were achieved for 95% of grade 3 fractures, and 100% of other classifications.

Vertebra Status	%age of sample	AAM			RFRV-CLM		
		Median (mm)	Mean (mm)	%>2mm	Median (mm)	Mean (mm)	%>2mm
Normal	83.25%	0.40	0.55	2.5	0.52	0.59	0.68
Deformed	4.38%	-	-	-	0.57	0.72	3.57
Grade 1	3.16%	0.49	0.70	4.8	0.60	0.73	0.99
Grade 2	4.06%	0.61	0.92	10.2	0.80	0.91	3.84
Grade 3	3.28%	0.72	1.19	16.5	0.90	1.11	10.48

Table 1. Error statistics derived from Fig. 2, compared to those presented in [2]. Bold figures are the best result for each status/statistic.