# Automated Textural Classification of Osteoarthritis Magnetic Resonance Images

Joshua D Kaggie[1,2], Rob Tovey[3], James MacKay[1,2], Fiona J Gilbert[1,2], Ferdia Gallagher[1,2], Andrew McCaskie[2,4], and Martin J Graves[1,2]

[1]Radiology, University of Cambridge, Cambridge, United Kingdom, [2]Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom, [3]Mathematics, University of Cambridge, Cambridge, United Kingdom, [4]Division of Trauma and Orthopaedic Surgery, University of Cambridge, Cambridge, United Kingdom

## Synopsis

**Osteoarthritis (OA) is the most common cause of disability in the United Kingdom and United States. Identifying the rate of OA progression remains an important clinical and research challenge for early disease monitoring. Texture analysis of tibial subchondral bone using magnetic resonance imaging (MRI) has demonstrated the ability to discriminate between different stages of OA. This work combines texture analysis with machine learning methods (Lasso, Decision Tree, and Neural Network) to predict radiographic disease progression over 3 years, trained using data from the Osteoarthritis Initiative. We achieved high sensitivity (86%), specificity (64%) and accuracy (74%) for predictions of OA progression.**

## Introduction

Osteoarthritis (OA) is a disease of the joints that results in degradation of multiple tissues including cartilage and bone [1]. OA is the greatest cause of disability in the United Kingdom and United States [2, 3]. Diagnosis is generally clear in patients with advanced disease, but remains challenging at early stages where stratification and measurement of progression is necessary for the development of effective therapies to halt or reverse disease progression [4].

Subchondral bone texture, analysed from both radiographic and MR images, is altered in severe OA [5-10]. Texture analysis promises to quantify disease state beyond arbitrary thresholds of disease presence, thereby improving discrimination of disease state[7]. Machine learning methods, such as Lasso [11], Decision Trees [12, 13] and Neural Networks [14-16], promise to improve the predictability of texture analysis despite these concerns.

## Methods

Included subjects were participants in the Osteoarthritis Initiative (OAI) bone ancillary study (BAS) [17]. The OAI provided anonymised patient records that contained age, gender, Body Mass Index (BMI) and Joint Space Width (JSW) measurements as measured from automated software. The BAS study subjects were imaged with standard OAI MRI sequences [18] and 3D fast imaging with steady state precession (FISP) MR images, acquired at 3T, with an echo time = 4.9ms, repetition time = 20ms, flip angle = 50°, matrix size = 512x512, interpolated to 1024x1024, slice thickness = 1mm, number of averages = 1, and acquisition time = 10.5 minutes.

Our study featured 61 subjects who were defined as radiographic progressors based on a reduction in minimum medial JSW of at least 0.7 mm over the 3 years following their initial BAS MRI. Progressors were age/gender matched to subjects who did not meet the radiographic progression criterion, therefore, our total dataset included 122 patient records. Training data for machine learning methods contained a ratio of 28/33 progressors vs non-progressors and the validation data contained 33/28, respectively.

A region-of-interest (ROI) was selected on a slice central to the knee, along the bone-cartilage interface and extending approximately 1cm into the subchondral bone (Figure 1). First and second order Haralick texture parameters [19], fractal textures [9], and gradient textures [20] were calculated in the ROI. The texture calculations were performed in Cython (Python Software Foundation). The texture parameters were fed into three types of machine learning algorithms: Lasso regression, the Decision Tree, and a Neural Network in the Scientific Toolkit, Scikit-Learn, Python package that were trained to predict JSW changes.

Accuracy was defined as the ratio of true positives and negatives to the total number of subjects

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}.$$

## Results

Three hundred and twenty textural parameters were calculated for 122 patients. Box plots for sample textural values of the training dataset are shown in Figure 2. Maximal correlation coefficient textures in the superior/inferior direction demonstrated the highest linear correlation with binary categorisation of JSW ($r2=0.98$), followed by directional gradients. First order values, such as mean, standard deviation, skew and kurtosis, demonstrated the lowest r2 correlations ($r2<0.40$). Skew and sum superior-inferior entropy were the primary indicators of disease according to the Decision Tree, despite the lack of linear correlation of skew with disease ($r2=0.39$).

Texture parameters that were derived from MR images and applied to a Neural Network accurately predicted disease progression in 74% of all cases (Tables 1,2). The Lasso and Decision Tree did not have the same accuracy rates (57%, 62%, respectively) as the Neural Network (74%).

## Discussion

The true positive rate (86%) of this prediction is higher than plain radiograph and histology predictions, which have a sensitivities between

11/08/2017 04:57 PM

61% and 74% [21]. Such comparisons with competitors demonstrate that texture values relate to disease progression. This work could also be applied to radiographic tomograms, which show OA-related textures.

These techniques can be applied to other diseases where any binary disease state can be determined. This platform could be used for detecting and predicting other diseases that have texture information in the images. This work could also be applied to radiographs which show features of OA to establish is similar textural information can be extracted. There are a broad range of textural patterns that could be explored to improve the performance of this work [22].

## Conclusion

We have successfully measured first and second order texture values, as well as gradient and fractal textures, which were used as training sets in Lasso, Decision Tree, and Neural Network algorithms. The Neural Network performed with the highest sensitivity (86%), specificity (64%), and accuracy (74%), which is comparable to established methods of osteoarthritis detection.

## Acknowledgements

## References

1. Forman MD, Malamet R, Kaplan D. A survey of osteoarthritis of the knee in the elderly. The Journal of Rheumatology 1983; 10: 282-287.

2. Murray CJ, Richards MA, Newton JN, Fenton KA, Anderson HR, Atkinson C, et al. UK health performance: findings of the Global Burden of Disease Study 2010. The Lancet 2013; 381: 997-1020.

3. Andersson G, Surgeons A. United States Bone and Joint Initiative: The Burden of Musculoskeletal Diseases in the United States (BMUS), 2014. Rosemont, IL. In: 2015.

4. Hunter D, Altman R, Cicuttini F, Crema M, Duryea J, Eckstein F, et al. OARSI clinical trials recommendations: knee imaging in clinical trials in osteoarthritis. Osteoarthritis and Cartilage 2015; 23: 698-715.

5. Woloszynski T, Podsiadlo P, Stachowiak G, Kurzynski M, Lohmander L, Englund M. Prediction of progression of radiographic knee osteoarthritis using tibial trabecular bone texture. Arthritis & Rheumatology 2012; 64: 688-695.

6. Kraus VB, Feng S, Wang S, White S, Ainslie M, Brett A, et al. Trabecular morphometry by fractal signature analysis is a novel marker of osteoarthritis progression. Arthritis & Rheumatology 2009; 60: 3711-3722.

7. MacKay JW, Murray PJ, Low SB, Kasmai B, Johnson G, Donell ST, et al. Quantitative analysis of tibial subchondral bone: Texture analysis outperforms conventional trabecular microarchitecture analysis. Journal of Magnetic Resonance Imaging 2015.

8. MacKay JW, Murray PJ, Kasmai B, Johnson G, Donell ST, Toms AP. MRI texture analysis of subchondral bone at the tibial plateau. European Radiology 2016; 26: 3034-3045.

9. Chung HW, Chu CC, Underweiser M, Wehrli FW. On the fractal nature of trabecular structure. Medical Physics 1994; 21: 1535-1540.

10. Fazzalari N, Parkinson I. Fractal properties of subchondral cancellous bone in severe osteoarthritis of the hip. Journal of Bone and Mineral Research 1997; 12: 632-640.

11. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) 1996: 267-288.

12. Quinlan JR. Induction of decision trees. Machine Learning 1986; 1: 81-106.

13. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press 1984.

14. Specht DF. A general regression neural network. IEEE Transactions on Neural Networks 1991; 2: 568-576.

15. Sanger TD. Optimal unsupervised learning in a single-layer linear feedforward neural network. Neural Networks 1989; 2: 459-473.

16. Nielson JL, Paquette J, Liu AW, Guandique CF, Tovar CA, Inoue T, et al. Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. Nature communications 2015; 6.

17. Nevitt M, DT F, Lester L. OAI Protocol. The Osteoarthritis Initiative: Protocol for the cohort study. In, vol. 2006. https://oai.epi-ucsf.org /datarelease/docs/StudyDesignProtocol.pdf 2006.

18. Peterfy C, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis and Cartilage 2008; 16: 1433-1441.

19. Haralick RM, Shanmugam K. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics 1973; 3: 610-621.

20. Nothdurft H. Sensitivity for structure gradient in texture discrimination tasks. Vision Research 1985; 25: 1957-1968.

21. Menashe L, Hirko K, Losina E, Kloppenburg M, Zhang W, Li L, et al. The diagnostic performance of MRI in osteoarthritis: a systematic review and meta-analysis. Osteoarthritis and Cartilage 2012; 20: 13-21.

22. Zhang J, Tan T. Brief review of invariant texture analysis methods. Pattern Recognition 2002; 35: 735-747.
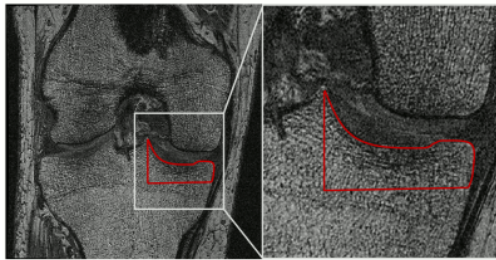
## Figures



Figure 1: An example ROI of the tibial plateau, showing an entire image of the knee (left) and a magnified image (right). The region-of-interest was selected along the bone/cartilage interface and extended approximately 1 cm into the subchondral bone.
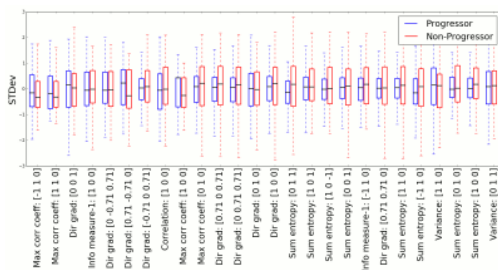


Figure 2: Box plots of several textural values calculated for progressors (red) and non-progressors (blue) for half of the dataset. The box plots represent the minimum, first quartile, median, third quartile, and maximum values for each textural value. The textural values are listed in order of decreasing r2 linear correlation, from 0.98 to 0.64.

| | JSW Measurement | |
| --- | --- | --- |
| **Method** | **Non-progressor** (negative) | **Progressor** (positive) |
| **Lasso** | | |
| Non-progressor (negative) | 42% | 25% |
| Progressor (positive) | 58% | 75% |
| **Decision Tree** | | |
| Non-progressor (negative) | 45% | 18% |
| Progressor (positive) | 55% | 82% |
| **Neural Network** | | |
| Non-progressor (negative) | 64% | 14% |
| Progressor (positive) | 36% | 86% |
| Training Samples | 28 | 33 |

Confusion matrix of measured and predicted outcomes, i.e., a table of true/false postivies/negatives of progressors and non-progressors based on joint-space-width change (by column) and based on machine learning predictions (by row). The machine learning predictions were created using a training set of textural values derived from knee images.

| Comparison | Accuracy | $\chi^2$ p-value |
|---|---|---|
| Lasso | 57% | 25% |
| Decision Tree | 62% | 5.5% |
| Neural Network | 74% | 0.02% |

Table 2 Table of accuracy and paired χ2 p-values of machine predictions for the Lasso, Decision Tree, and Neural Network methods. The p-values were calculated based on the predicted versus measured progression. The Neural Network performs better than the Lasso and Decision Tree.