

# Automatic Modic Changes Classification in Spinal MRI

Amir Jamaludin<sup>1</sup>, Timor Kadir<sup>2</sup>, and Andrew Zisserman<sup>1</sup>

<sup>1</sup> University of Oxford, UK

<sup>2</sup> Mirada Medical, UK

{amirj, az}@robots.ox.ac.uk

timor.kadir@mirada-medical.com

<http://www.robots.ox.ac.uk/~vgg/>

**Abstract.** This paper describes a novel automatic system for Modic changes classification of vertebral endplates. Modic changes are classes of vertebral degenerations visible as intensity variations in Magnetic Resonance Images (MRI). The system operates on T1 and T2 MRI. We introduce three main novelties: (1) a vertebrae alignment scheme via precise bounding boxes obtained through corner localisation, (2) vertebral endplate classification in 3D, and (3) Modic changes classification. The system was trained and validated using a large dataset of 785 patients, containing MRIs sourced from a wide range of acquisition protocols. The proposed system achieved 87.8% classification accuracy on our dataset.

**Keywords:** Modic Changes, Marrow Changes, Vertebrae Analysis

## 1 Introduction

The objective of this work is the automated classification of Modic changes in MRI sagittal lumbar scans. Modic changes are classes of vertebral degenerations visible as intensity variations in MRI. There are three types of Modic changes and each type possesses varying correlation with the degradation of the vertebral bodies (VB) with Modic type 1 having the highest correlation with clinical pain scores [1]. Classification of these changes in vertebral endplates is highly beneficial as it gives a measure of health of a vertebra which would help in the diagnosis of lower back pain. To our knowledge, this is the first system to automatically classify Modic changes and vertebral endplates degeneration in general. An example of the vertebral regions associated with the task is shown in Fig. 1.

The system is trained and its performance validated using a clinical dataset which is heterogeneous i.e. the scans are sourced from different clinical centres using different machines and protocols. The scans in the dataset possess a variation of field-of-view, field strength, resolution, and are susceptible to bias field corruption. Since it is a necessity for Modic changes classification to use both T1 and T2 scans, we also have to localize the vertebrae of the spine in the two scans. The advantages of such an automated system are improvement of radiological score consistency, which varies from one radiologist to another, and ease of pathological detection of Modic changes.

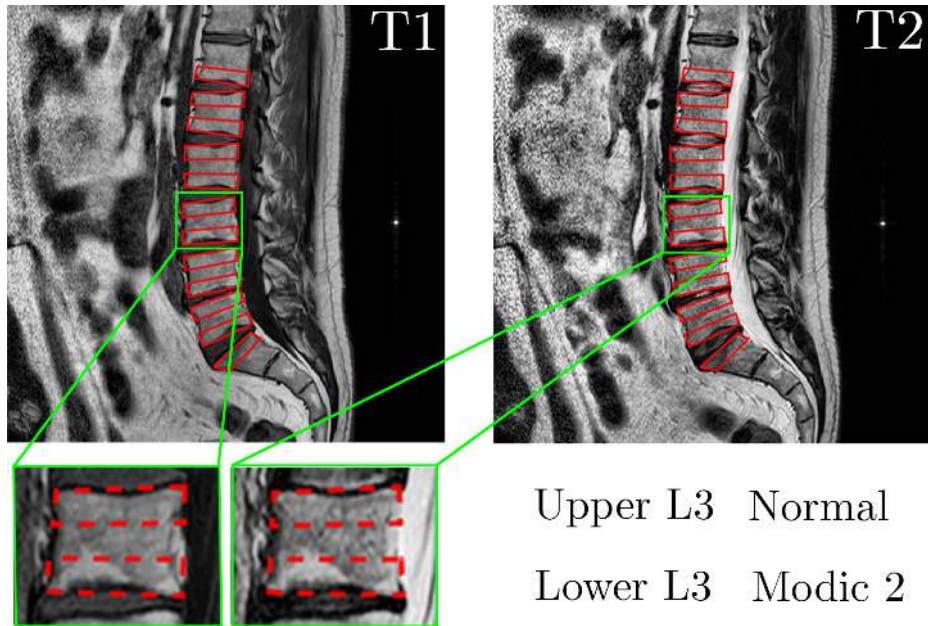


Fig. 1: **The task:** Given T1 and T2 weighted lumbar sagittal MR volumes, detect, localise, label, and predict the states of the endplates of the vertebrae from the lower endplate of T12 to the upper endplate of S1. In the example, only single slices of T1 and T2 are shown but in practice, the system operates on every slice in the scans. Bounding boxes, shown in red, represent the regions used in the classification of the endplates.

### 1.1 Modic Changes

The variation of voxel intensities in an MR scan can be said to be caused by the variation of proton densities of different organs. One good assumption is that a specific organ would roughly consist of the same material hence possess a narrower range of proton densities unique to that organ. By extension, a healthy or normal vertebra would exhibit homogeneous intensity distributions in its T1 and T2 scans. The opposite is true for some abnormal vertebrae where visible discolourations can be seen in both their T1 and T2 scans, as first discovered by Modic *et al.* [12], aptly named Modic changes and at times might also be referred to as marrow or vertebral body changes.

Modic postulated that such discolourations in the form of visible intensity changes of the vertebrae might be caused by the evolution of marrow of the vertebral endplate in response to the degeneration of the corresponding intervertebral disc [4, 12]. This response is hypothesised to be either mechanical or bacterial in cause but which or both has yet to be determined [1]. There are three different types of Modic changes, each characterized by a change in intensity of the vertebral endplates in both the T1 and T2 scans. See Fig. 2 and Table 1.

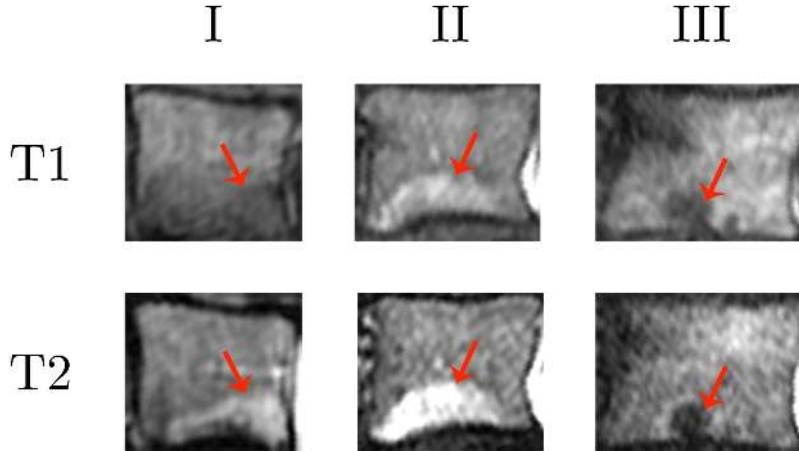


Fig. 2: Each column pair of images is from an individual vertebra scanned with T1 and T2 weightings and each pair represents an example of the type of Modic changes they belong to. These discolourations only appear on the endplates of the vertebrae adhering to the definition set forth by Modic *et al.* [12].

Modic Type	T1 Endplate Intensity	T2 Endplate Intensity
I	<i>Hypointense</i>	<i>Hyperintense</i>
II	<i>Hyperintense</i>	<i>Isointense/Hyperintense</i>
III	<i>Hypointense</i>	<i>Hypointense</i>

Table 1: An overview of the definition of the 3 types of Modic changes where we follow a standard radiological intensity terminology where a hyperintense area refers to an area with higher intensity in comparison to its surrounding and vice versa.

Several studies have been conducted to find the relationship between Modic changes and lower back pain. Modic type 1 has the strongest correlation with lower back pain and type 1 endplates may stabilise into Modic type 2 over time, which has a lower correlation with lower back pain [11, 14, 16]. This suggests that patients with lower back pain might possibly be monitored by observing just the state of their endplates.

## 1.2 Related Work

This work is the first fully automated system to classify Modic changes but there exists another system proposed by Vivas *et al.* [17] that is semi-automatic, requiring inputs for disc detections, and only does binary classification i.e. a simple Modic/non-Modic classification instead of a multi-class classification scheme into the three types or normal.

Most research in analysis of spine imaging has focussed on the intervertebral discs rather than the VB. Typically, Disc classifications methods use both intensity and shape information of the discs [7, 10]. Existing work on VB analyses

mainly focussed on vertebral fractures [15] and sclerotic metastases [2]. Vertebral fracture analysis classification requires a highly accurate per vertebra region of interest (ROI) fit because fractures are correlated with the height of the vertebra [15]. This is unlike Modic changes which are independent of the shape of the vertebra, focusing only on intensity. In this paper, we compare our feature with the normalised intensity histogram, Hist+, by Lootus *et al.* [10]. Our proposed features are also comparable to the spatial binned ROI intensity features by Ghosh *et al.* [7].

## 2 Approach Overview

Our method has five stages: (1) 2D vertebrae detection and labelling in all slices of a given scan, (2) corner localisation, (3) vertebrae alignment, (4) 3D vertebrae extent detection, and (5) classification. See Fig. 3.

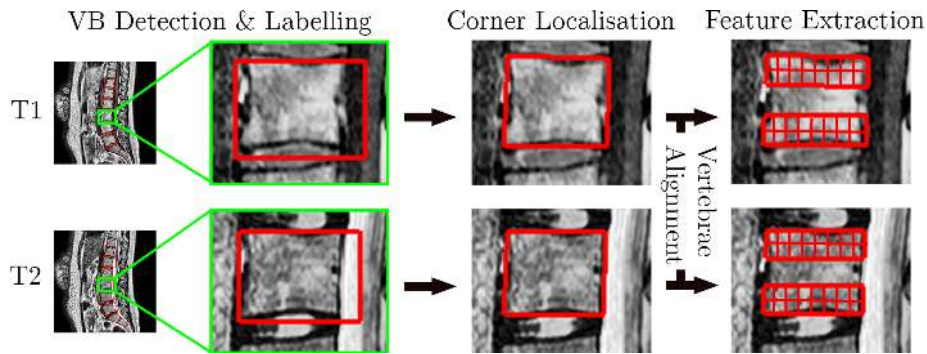


Fig. 3: An overview of our approach. The arrows indicate the processes while the images show a close up view of the vertebra in both scans. This is repeated in all of the slices of the volume.

### 2.1 2D Vertebrae Detection and Labelling – Stage 1

To detect and label the vertebrae, we follow and adapted the detection and labelling scheme proposed by Lootus *et al.* [9] which uses a combination of a Deformable Part Model (DPM) detector [5] and labelling via graphical model. The input to this stage is a 3D MR volume and the output is a series of approximate bounding boxes with the vertebrae labels from T12 to the combined sacrum (S1 and S2). The detector and graphical model is trained using scans with annotated ground truth bounding boxes with labels as described in [9].

### 2.2 Corner Localisation – Stage 2

There is an issue with using the loose bounding boxes as our classification features which is: the variation of fit of the bounding boxes, both intra-scan and inter-scan, which can be seen in Fig. 4.

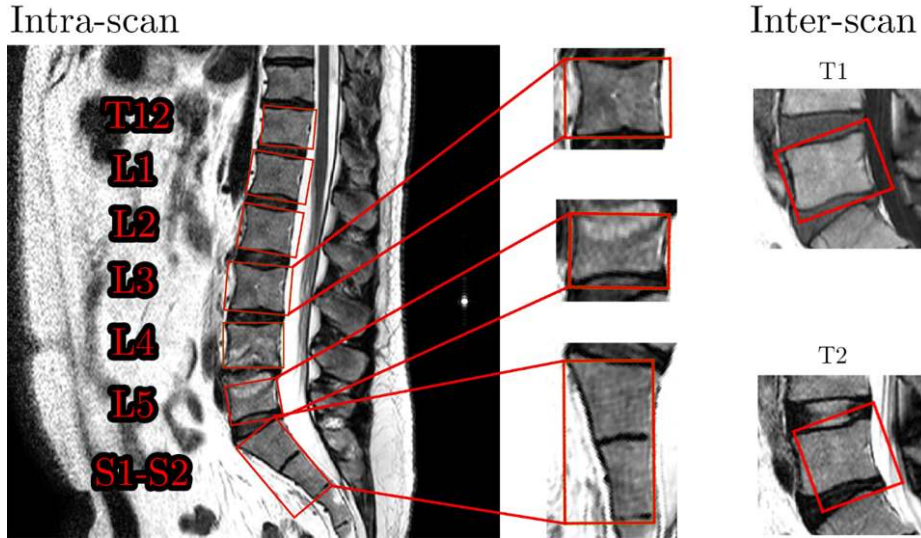


Fig. 4: An example output using the detection and labelling system described in [9]. The enlarged set of images of the vertebrae show the variation of fit of the bounding boxes to the vertebrae. **Intra-scan variability:** Note the sacrum bounding box contains both S1 and S2. The L3 bounding box contains all of the vertebra but is slightly loose while the L5 bounding box is missing the upper endplate of the vertebra. **Inter-scan variability:** Both images show the same L5 vertebra at the different contrasts and their respective bounding boxes.

We propose a finer localisation post-processing of these bounding boxes such that the resulting bounding boxes are more consistent and tightly aligned with the vertebrae. Tighter alignment leads to improvements in localizing the regions used to extract the features, and also helps in the alignment of the vertebrae in the T1 and T2 scans (Section 2.3).

We adapt the supervised descent method (SDM) by Xiong *et al.* [18] originally developed for the detection of facial landmarks to improve the localisation accuracy of the bounding boxes. The input to this stage is the image together with its bounding box from the vertebrae detection and labelling stage and the output is an irregular quadrilateral, with a tighter fit around the vertebra.

At test time, the algorithm works by first regressing the corner points of the loose bounding boxes to a learned mean vertebrae corner points. Then, these intermediate corner points are iteratively updated via regression based on SIFT features around the points. The regression can be solved iteratively and can be represented as:

$$x_{k+1} = x_k + \Delta x_k \quad (1)$$

where  $\lim_{k \rightarrow \infty} x_k = x_*$ .  $x_0$  are the 4 points of the loose bounding boxes and  $x_*$  are the vertebral corner points. After several iterations,  $k = 10$  works well in our dataset, the regression is stopped and the final points become the new corner points describing the best quadrilateral fit for the vertebra.

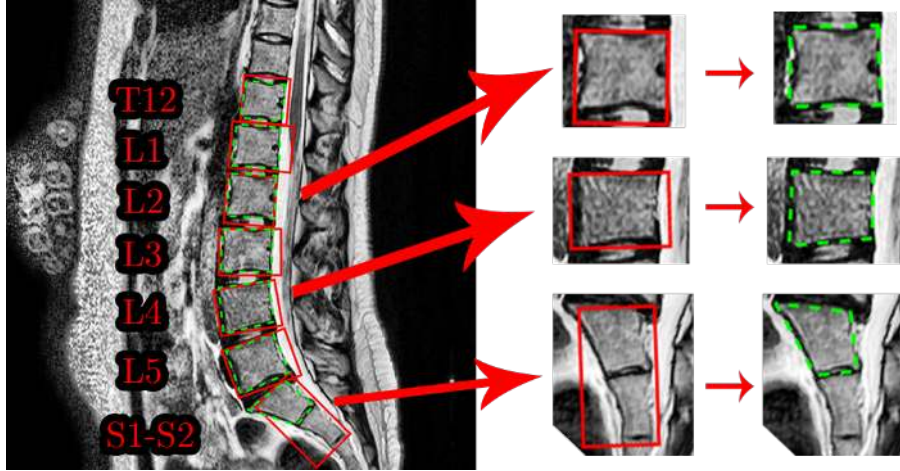


Fig. 5: Examples of inputs, shown in red, and outputs, shown in green, of the corner localisation. Note the sacrum bounding box is now more specific, containing only S1 instead of both S1 and S2, and the variation of bounding boxes fit is reduced.

Training the regressor is posed as a minimisation task of  $\Delta x$ :

$$\operatorname{argmin}_{\Delta x} f(x_0 + \Delta x) = \| \mathbf{h}(\mathbf{d}(x_0 + \Delta x)) - \phi_* \|_2^2 \quad (2)$$

where  $\mathbf{h}$  is the feature transformation, which in this case is SIFT, of that point and  $\phi_* = \mathbf{h}(\mathbf{d}(x_*))$  represents the SIFT features at the ground truth. To overcome overfitting, ridge regression was used. The regularizer,  $\lambda$ , is a general singular value penalty imposed per iteration.

Since the size of the vertebrae in the dataset vary, a normalisation step has to be conducted prior to regression. An individual vertebra is resized according to the height,  $V_h$ , of its bounding box and the image is translated such that the centre of the bounding box is the point of origin. Examples of corner localised vertebrae given its raw bounding boxes can be seen in Fig. 5.

### 2.3 Vertebrae Alignment – Stage 3

This section describes the vertebrae alignment stage which is necessary because there is no guarantee of a good alignment between the T1 and T2 scans. The inputs to this stage are two tight bounding boxes, one each from the T1 and T2 scans, that contain one vertebra and the outputs are rigid transforms that describe the motion between the two scans specifically for that single vertebra.

A good alignment of the vertebra in both T1 and T2 scans is important since features extracted from both scans are jointly used for classification. The two main reasons for misalignments of the vertebrae are: (1) the movements of the patients in between scans and (2) the inter-scan difference in bounding boxes, both in terms of position and shape, detected in the T1 and T2 scans as shown in Fig. 4.

For the alignment, the vertebrae are assumed to be rigid bodies and any motion related to them is assumed to be only in-slice pitch and translation i.e. slice correspondence is assumed to be valid. As such, any extreme movement especially any yaw and roll would result in failure in alignment. These assumptions work well since scanning procedures dictate that patients should lie in the same orientation for both the T1 and T2 scans. Thus, the solution for the motion or transformation between the scans of the same vertebra can be expressed as:

$$\mathbf{V}_{T1} = \mathbf{R}(\theta)\mathbf{V}_{T2} + \mathbf{T} \quad (3)$$

which describes both rotational,  $\mathbf{R}$  ( $2 \times 2$  matrix), and translational,  $\mathbf{T}$  ( $2 \times 1$  vector), motions between T1 and T2. Both,  $\mathbf{V}_{T1}$  and  $\mathbf{V}_{T2}$  are 2D coordinate feature points which are detected in both images.

We use the 4 corner points obtained in the previous stages in both T1 and T2 to be the feature points for alignment. However, the regressed corner points are not without mistakes. This is because vertebral corners tend to be smooth making it hard to pinpoint the exact locations of the corner points. Thus, a mechanism for identifying these mistakes and being tolerant to them is needed. We use a RANSAC-like approach to estimate the transformation and identify inliers/outliers simultaneously [6]. Detected outliers are removed from  $\mathbf{V}_{T1}$  and  $\mathbf{V}_{T2}$ . After the rigid transformation of each vertebra has been obtained, we transform one scan to the other, aligning them, and use these as the inputs for Modic classification.

#### 2.4 3D Vertebrae Extent Detection – Stage 4

Our aim in this stage is to determine the 3D extent of the vertebrae from the 2D quadrilaterals in each sagittal slice; this requires determining where the original detections should start and end slice-wise. This is important since the positions of the vertebrae in a scan are initially unknown and there exist slices which contain only partial volumes of the vertebrae, mostly containing tissue. These partial vertebrae are problematic if they are selected as ROIs for feature extraction. To this end we utilise a classifier to distinguish non-vertebrae and vertebrae quadrilaterals.

We follow a standard image classification scheme, discussed by Chatfield *et al.* in [3], where the sequential steps are: (1) dense SIFT feature extraction over the quadrilaterals, (2) Fisher Vector (FV) encoding of the features, (3) spatial tiling of the features in the image and (4) classification via linear SVM. This is done on a per slice basis on every slice. Quadrilaterals classified as vertebrae are passed through to the Modic classification stage.

### 3 Classification

The classification of the vertebral endplates starts with feature extractions of the endplate regions which then are fed into the classifier. There are four different types of classes an endplate might be classified as, namely: normal and the three different types of Modic changes.

### 3.1 Feature Extraction

Prior to classification, the right features have to be extracted such that they best separate the different classes which in this case are the different types of Modic changes. Since the very definition of the different Modic types is dependent on the joint intensity of the images, we propose a feature that captures this information between T1 and T2: a spatially-binned joint histogram of intensities, SJT. As a baseline measure, our feature is compared with a histogram based scheme by Lootus *et al.* [10] tested on radiological disc grading, Hist+.

The ROIs for feature extractions are the upper and lower thirds of the vertebra-aligned corner-localised tight quadrilaterals of T1 and T2. These ROIs are essentially shorter quadrilaterals, one each for upper and lower endplates shown in Fig. 1 and Fig. 6, which cover the vertebral endplates. We have also experimented on using vertebrae segmentation proposed in [10] but since the nature of the problem itself is dependent on intensity, segmentation of the vertebrae proved to be unhelpful due to the intensity variation of endplate with Modic changes. The intensity of each endplate ROI is median normalised using both the intensity distributions of the vertebra and its neighbouring vertebra. This reduces the effect of bias field and protocol intensity variation while preserving actual per vertebra variation which is crucial for Modic classification.

**SJT:** Each ROI is spatially-binned,  $2 \times 8$  bins, so that there is a measure of spatial statistics of the Modic changes and from each corresponding T1-T2 pair of spatial cells we construct a  $16 \times 16$  joint intensity histogram. The joint histogram is constructed by binning every pixel pair (T1 and T2 in the ROI) according to its joint-intensity; see Fig. 6. The overall dimension of the feature vector is 4096. Since the histogram is quite sparse, images are upsampled three times with respect to their original sizes prior to feature extraction [8].

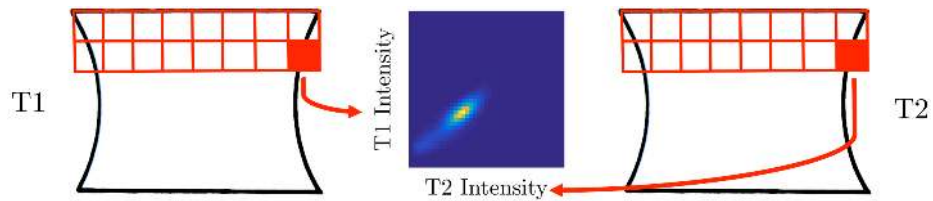


Fig. 6: An example of a joint histogram of one cell in the whole region. By introducing spatial bins, we capture the localisation of the intensity changes of the vertebral endplates.

**Hist+:** For each ROI, an intensity histogram, 16 bins, with its moments (mean, standard deviation, kurtosis, skewness, and entropy) is constructed. This forms a feature vector with 21 dimensions (16 intensity bins + 5 moments) for each ROI. The T1 and T2 features are concatenated forming the final feature vector.



### 3.2 Learning

For the classification task, a linear Support Vector Machine (SVM) is used. The classification is assessed with a 50:50 split of the dataset described in Section 4.1. To obtain the statistical variation of the classifier, the assessment is done 20 times, each time with a randomly selected 50:50 split according to patients. The optimal value for the parameter C is found via a 5-fold cross validation of the training set. A uni-slice classifier is trained using vertebral endplate labels marked by a radiologist where the best slice is used per endplate. The best slice is manually selected using the endplate labels from the radiologist as reference i.e. if the endplate is labelled to show Modic 1 change, the slice which best represent type 1 change is used. We use the one-versus-rest approach for the multi-class classification task e.g. for a Modic 1 classifier, the positive examples are Modic 1 endplates while the negative examples are the opposite and vice versa. To handle classification in multiple slices (3D extent of the endplates), the detection of which is discussed in Section 2.4, we use mean pooling of the classifier scores.

### 3.3 Data Augmentation

To further enhance the accuracy of our classification task, we applied several different augmentation transforms, five in total, to the data. To automatically choose the five best transforms i.e. five extra augmented samples in both train and test sets that help capture the invariances of the vertebrae in our dataset the Image Transformation Pursuit (ITP) algorithm by Paulin *et al.* [13] was used. The transforms, 43 overall, we experimented on are as follows:

- 1 Flip/mirror
- 20 Rotations:  $\theta = -10^\circ$  to  $10^\circ$  at  $1^\circ$  increment
- 16 Shifts:  $\uparrow, \downarrow, \leftarrow, \rightarrow, \nearrow, \searrow, \swarrow, \nwarrow$  with two pixel distances (5 and 10 pixels)
- 6 Bounding Box Scales: 85% to 115% in 5% increment

The five chosen augmentations samples are used as extra samples in training and we pool (mean) the classifier scores from the unaugmented samples with its corresponding five augmentations at test time.

## 4 Results & Discussion

### 4.1 Dataset

The dataset used to train and validate our system came from a range of different MRI machines and protocols. In all, the dataset consists of 785 patients of which 683 possess radiological scores (graded by a single radiologist) e.g. Modic changes, Pfirrmann grading etc. These patients are scanned on the basis that they are already diagnosed with back pain and no patient can be classified as healthy.

Of the 785 patients, 341 were used to train stages (1) to (4) described in Section 2 while 444 were used for stage (5), classification. There is no overlap of patients in training stages (1) to (4) and the patients in the classification. Only 388 patients out of the 444 possess labels for Modic changes, see Table 2.

Endplate Class	Patients	Endplates
Normal	144	3921
Modic 1	111	249
Modic 2	195	551
Modic 3	18	42

Table 2: There is a total of 388 patients and 4656 endplates where 194 patients are used for training and 194 for testing i.e. 50:50 random splits with no overlap of patients in training and testing. Most patients with Modic changes possess more than one Modic endplates and some of the patients possess more than one Modic types resulting in the total of the numbers shown in the table to be different than the actual total.

## 4.2 Corner Localisation

Two sets of models are trained for corner localisation, each set having a single regressor for the standard VB from the T12 to the L5 vertebrae, and a more specific S1 regressor. The first set, termed the 1<sup>st</sup> stage, is for more coarse localisation, and the second set, termed the 2<sup>nd</sup> stage is for a precise corner regression. Both sets of models are trained with the same ground truth used in training the detection and labelling system. Overall, 4274 T12-L5 and 720 S1 vertebrae were used in training the VB and S1 models respectively. The results of the trained models on an unseen validation set is shown in Fig. 7.

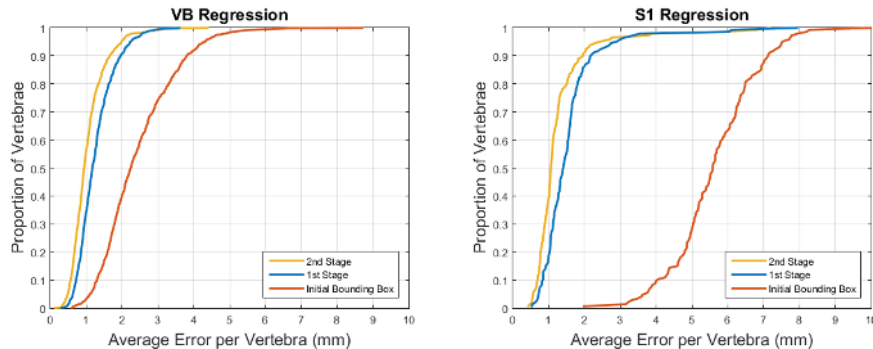


Fig. 7: (Left) Error per vertebra of the VB regression. (Right) Error per vertebra of the S1 regression. The red, blue, and yellow plots represent the errors of the raw bounding boxes, 1<sup>st</sup> stage and 2<sup>nd</sup> stage respectively. Both stages outperform the initial bounding box considerably. This is especially true for S1 which has a larger fit variance of the bounding boxes compared to VB.

There are only 4 parameters that have to be optimised; they are: the normalised height of the VB,  $V_h$ , the size of the SIFT patch, the regularizer of the ridge regression,  $\lambda$ , and the number of iterations,  $k$ . They are optimized such that they minimize the error on a hold-out validation set. In general, 91.1% of the S1 vertebrae and 95.3% of the VB have errors less than 2mm which is a considerable improvement over the original bounding box detections with none of the S1 vertebrae and 41.6% of the VB at the same error threshold.

### 4.3 3D Vertebrae Extent Detection

For classification of vertebrae and non-vertebrae, the same patients that trained the corner regressors were used; a total of 66556 VB quadrilaterals (39309 vertebrae and 27247 non-vertebrae) with 50:50 train and test split. In general, this stage of the system performs well with an accuracy of 95.6%.

### 4.4 Classification

For the classification task, the results can be separated into two parts: AUC of ROC of the one-versus-rest classifiers and overall accuracy of the multi slice classification task. Results are shown in Table 3.

Features	AUC of ROC				Accuracy
	Normal	Modic 1	Modic 2	Modic 3	
Hist+	83.8 $\pm$ 1.4	81.6 $\pm$ 2.7	81.2 $\pm$ 1.4	89.3 $\pm$ 8.7	85.8 $\pm$ 0.7
SJT	88.8 $\pm$ 0.9	85.5 $\pm$ 2.6	86.1 $\pm$ 1.0	90.1 $\pm$ 6.5	<b>87.8 <math>\pm</math> 0.6</b>

Table 3: Automatic endplate classification: (Left) AUC of the ROC curve of the one-versus-rest classifiers. (Right) Accuracy of the classification.

It can be seen that the suggested SJT features outperform the Hist+ features by a considerable margin with a 2.0% difference in accuracy and roughly between 4-5% difference in AUC of ROC. Furthermore, the standard deviation of the SJT features is consistently slightly less than that of Hist+ which suggests SJT to be more robust in learning. Note that the vertebral endplate as a whole is classified but this does not localise the slices that the Modic changes occur in. The result of the classification, 87.8% in accuracy, is better than the accuracy without augmentation at 87.4% and the chance accuracy at 69.4%.

Up this point we have assessed a fully automatic classification of the endplates by pooling the uni-slice classifications across the vertebra. In comparison, if only a single hand-picked best slice for each vertebra was used at test time i.e. manually select the slice that most clearly show the Modic change, we see a slightly better performance of 88.3%, a minor difference of 0.5%.

## 5 Conclusion

This paper has presented the first system to classify Modic changes automatically and validated using a large dataset. Since the ROIs extracted by the system are highly accurate and consistent, they can be used in other vertebral tissue classification problems which we hope to explore, alongside segmentation and localisation of Modic changes, in the near future.

**Acknowledgments.** We are grateful for discussions with Prof. Jeremy Fairbank, Dr. Meelis Lootus, and Dr. Jill Urban. This work was supported by the RCUK CDT in Healthcare Innovation (EP/G036861/1). The data used in this research was obtained during the EC FP7 project HEALTH-F2-2008-201626.

## References

1. Albert, H.B., Kjaer, P., Jensen, T.S., Sorensen, J.S., Bendix, T., Manniche, C.: Modic changes, possible causes and relation to low back pain. *Med. Hypotheses* 70(2), 361–368 (2008)
2. Burns, J.E., Yao, J., Wiese, T.S., Munoz, H.E., Jones, E.C., Summers, R.M.: Automated detection of sclerotic metastases in the thoracolumbar spine at CT. *Radiology* 268(1), 69–78 (2013)
3. Chatfield, K., Lempitsky, V., Vedaldi, A., Zisserman, A.: The devil is in the details: an evaluation of recent feature encoding methods. In: *Proc. BMVC.* (2011)
4. Emch, T.M., Modic, M.T.: Imaging of lumbar degenerative disk disease: history and current state. *Skeletal Radiol.* 40(9), 1175–1189 (2011)
5. Felzenszwalb, P., Girshick, R., McAllester, D.: Cascade object detection with deformable part models. In: *Proc. CVPR.* pp. 2241–2248 (2010)
6. Fischler, M.A., Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM* 24(6), 381–395 (1981)
7. Ghosh, S., Alomari, R.S., Chaudhary, V., Dhillon, G.: Computer-aided diagnosis for lumbar mri using heterogeneous classifiers. In: *Proc. ISBI* (2011)
8. Kadir, T., Brady, M.: Estimating statistics in arbitrary regions of interest. In: *Proc. BMVC.* (2005)
9. Lootus, M., Kadir, T., Zisserman, A.: Vertebrae detection and labelling in lumbar mr images. In: *MICCAI Workshop: Computational Methods and Clinical Applications for Spine Imaging* (2013)
10. Lootus, M., Kadir, T., Zisserman, A.: Radiological grading of spinal MRI. In: *MICCAI Workshop: Computational Methods and Clinical Applications for Spine Imaging* (2014)
11. Mitra, D., Cassar-Pullicino, V.N., McCall, I.W.: Longitudinal study of vertebral type-1 end-plate changes on MR of the lumbar spine. *Eur Radiol* 14(9), 1574–1581 (2004)
12. Modic, M.T., Steinberg, P.M., Ross, J.S., Masaryk, T.J., Carter, J.R.: Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. *Radiology* 166, 193–199 (1988)
13. Paulin, M., Revaud, J., Harchaoui, Z., Perronnin, F., Schmid, C.: Transformation Pursuit for Image Classification. In: *Proc. CVPR* (2014)
14. Rahme, R., Moussa, R.: The modic vertebral endplate and marrow changes: pathologic significance and relation to low back pain and segmental instability of the lumbar spine. *AJNR* 29(5), 838–842 (2008)
15. Roberts, M.G., Pacheco, E.M., Mohankumar, R., Cootes, T.F., Adams, J.E.: Detection of vertebral fractures in DXA VFA images using statistical models of appearance and a semi-automatic segmentation. *Osteoporos Int* 21(12), 2037–2046 (2010)
16. Toyone, T., Takahashi, K., Kitahara, H., Yamagata, M., Murakami, M., Moriya, H.: Vertebral bone-marrow changes in degenerative lumbar disc disease. An MRI study of 74 patients with low back pain. *J Bone Joint Surg Br* 76(5), 757–764 (1994)
17. Vivas, E.L.A., Oliva, F.J.C., Aguilar, R.Q.F., González, A.G., Cruz, J.G., Ávila, J.M.J.: Application of a semiautomatic classifier for Modic and disk hernia changes in magnetic resonance. *Coluna/Columna* 14(1), 18–22 (2015)
18. Xiong, X., De la Torre, F.: Supervised descent method and its applications to face alignment. In: *Proc. CVPR* (2013)