Engineering and Physical Sciences Research Council

Getting Accurate Full Spine Segmentations Without Ground Truths

Rhydian Windsor, Amir Jamaludin, Timor Kadir, Andrew Zisserman

CANCER RESEARCH UK

EPSRC Centre for Doctoral Training in Autonomous Intelligent Machines & Systems

TL;DR

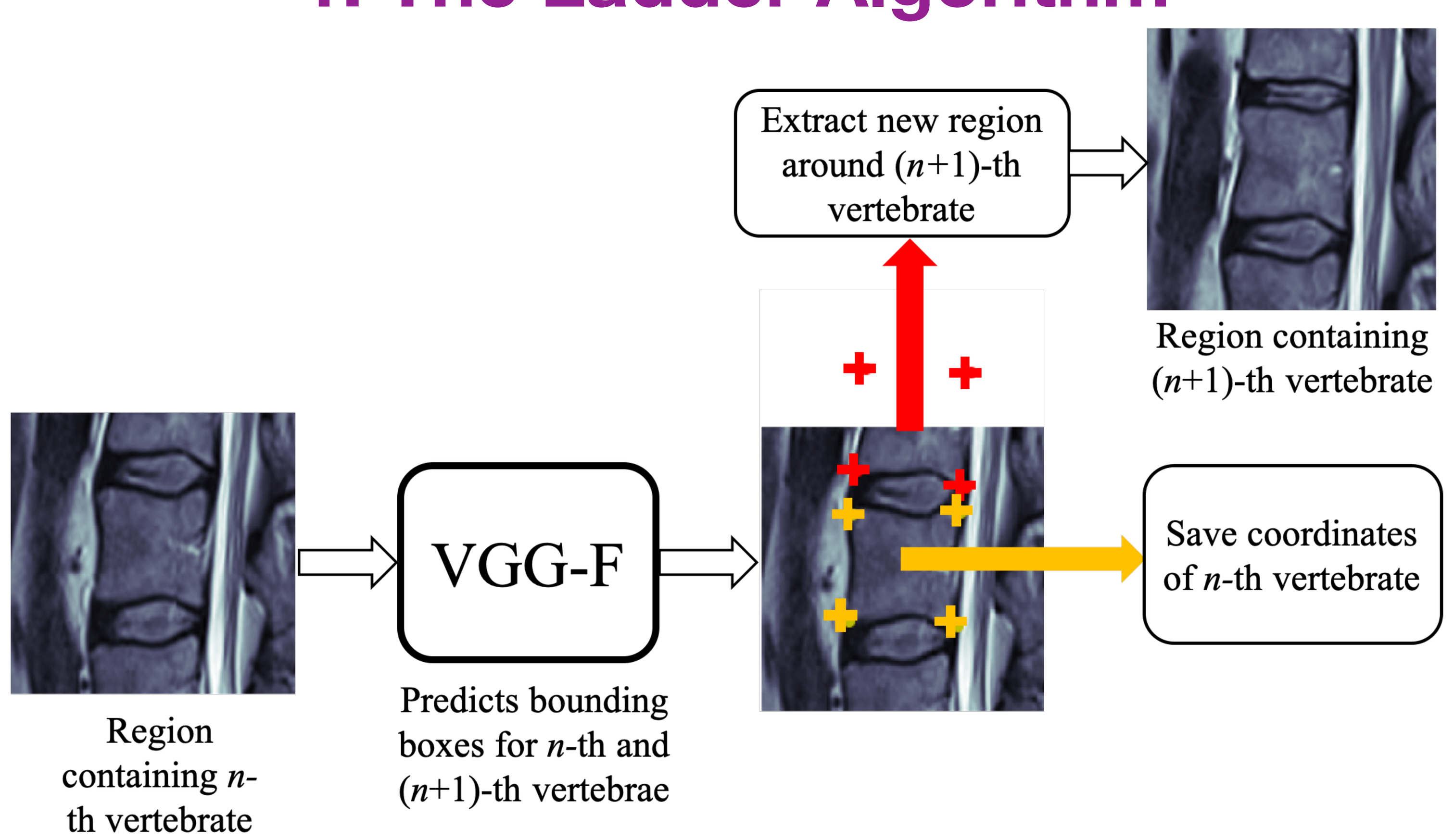
Labelling medical images is *boring* & training data is hard to come by.

It's *even worse* when we have repeating similar objects like spinal vertebrae, teeth and ribs.

Here, a new method reduces the amount of data needed to train classifiers to detect these repetitive structures by inductive reasoning.

We test the algorithm on localising vertebrae in full spine MR scans and achieve SOTA results.

1. The Ladder Algorithm

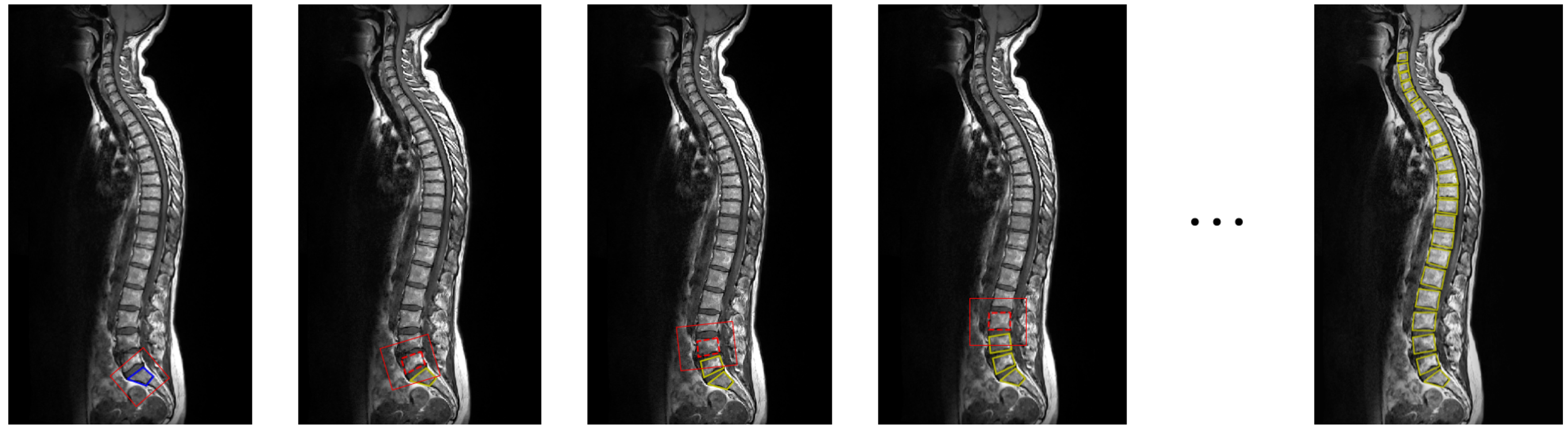


Given patch from a medical scan surrounding the n^{th} instance of a repetitive object in an image, we use the VGG-F network to predict bounding boxes for the n^{th} and $(n+1)^{th}$ instance of the object.

The coordinates of the boxes are transformed back into the full scan's coordinate frame. Those corresponding to the n^{th} detection are saved as a detection, whereas the coordinates corresponding to the $(n+1)^{th}$ scan are used to extract a new patch.

The algorithm iterates up the spine for a fixed number of iterations or until a separate 'stopping criterion' algorithm determine that we have reached the end of the sequence.

2. Application: Vertebrae Localisation



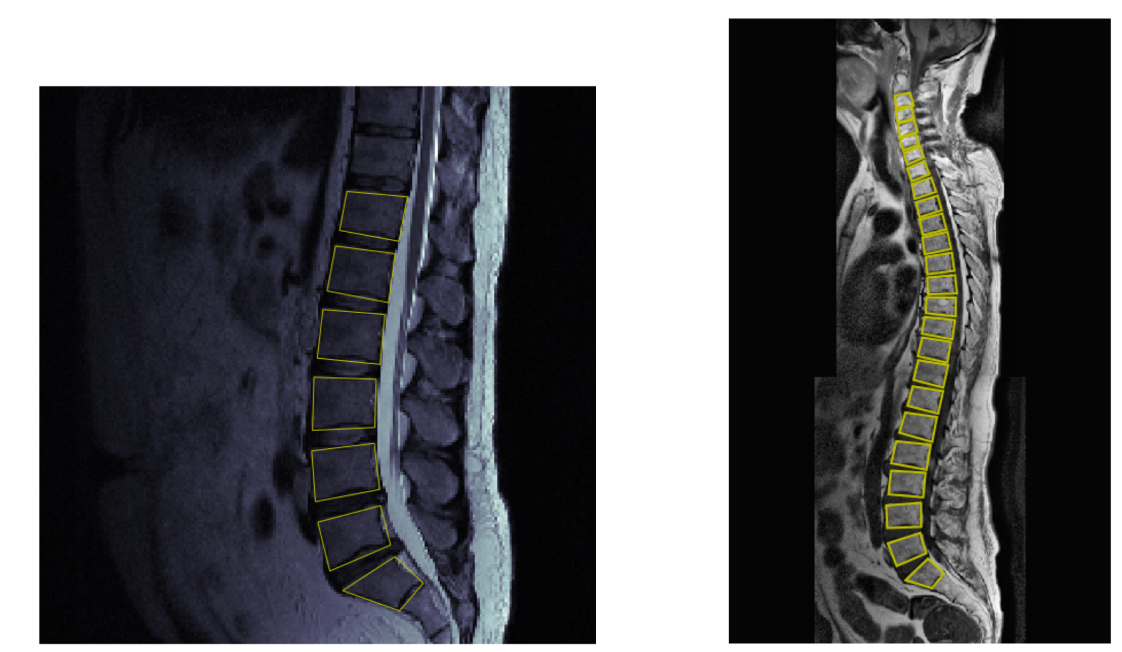
The algorithm is evaluated on the labelling of vertebrae in full spine MR scans. This task an important intermediate for diagnosing some spine diseases e.g. myeloma, AS. This is an interesting problem because there is a great deal of labelled data for lumbar (lower back) scans but very little exists for full spine scans.

Consequently, we set ourselves the following task: **Train a system to extract vertebrae from lumbar scans which will generalise 'up the spine' to full spine scans.**

Two datasets are used for this:

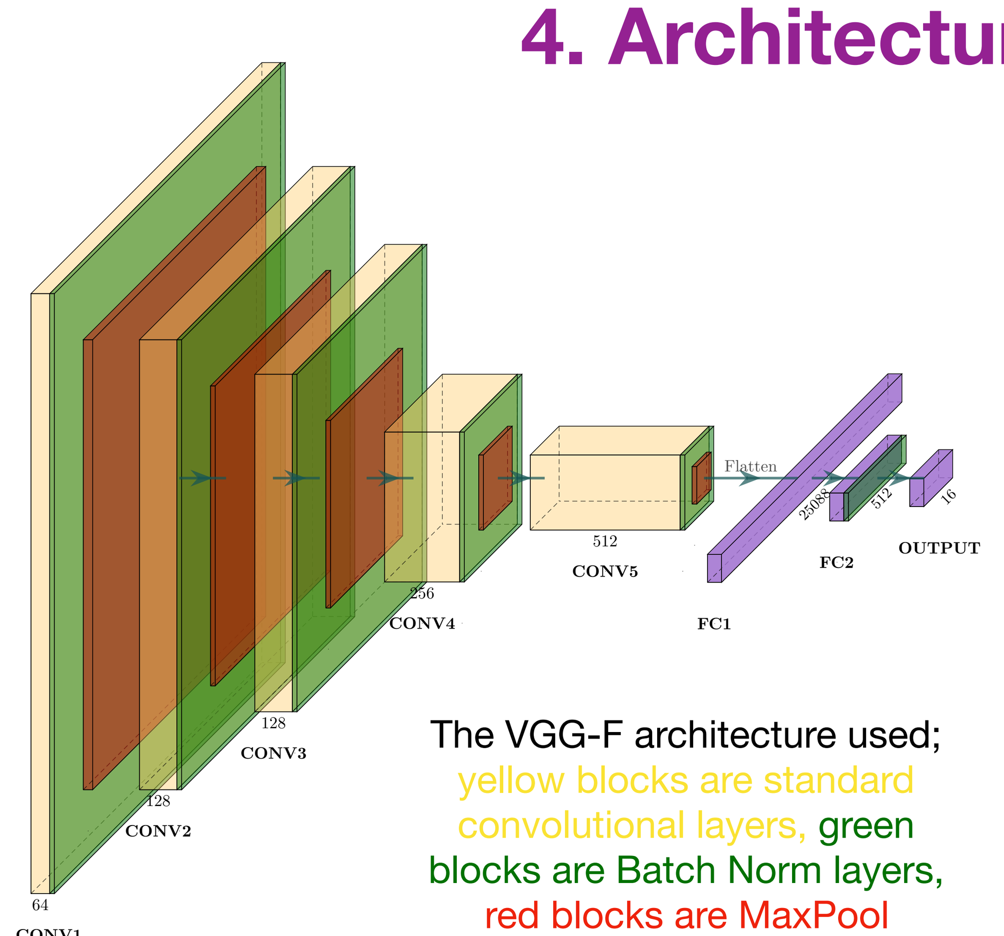
1) **Genodisc:** 5404 lumbar MR scans with ground truth bounding boxes for S1-T12 vertebrae (6 vertebrae, shown above). ***Used for training + evaluation***
2) **Oxford Whole Spine (OWS)**: 64 full spine MR scans with ground truth bounding boxes for S1-C2 vertebrae (23 vertebrae, shown above). ***Used only for evaluation***

The algorithm working recurrently up a full spine scan is shown above. The red box is the patch extracted and used as input to the network.



Lumbar Spine (from Genodisc) ***Full Spine*** (from OWS)

4. Architecture and Training



The VGG-F architecture used; yellow blocks are standard convolutional layers, green blocks are Batch Norm layers, red blocks are MaxPool operations and purple blocks are fully connected layers

The network is trained S1 to L1 vertebrae from Genodisc & applied from S1 to C3 vertebrae from OWS

Output of network are vectors of corner coordinates rather than response maps so we can predict vertebrae corners not visible in the patch

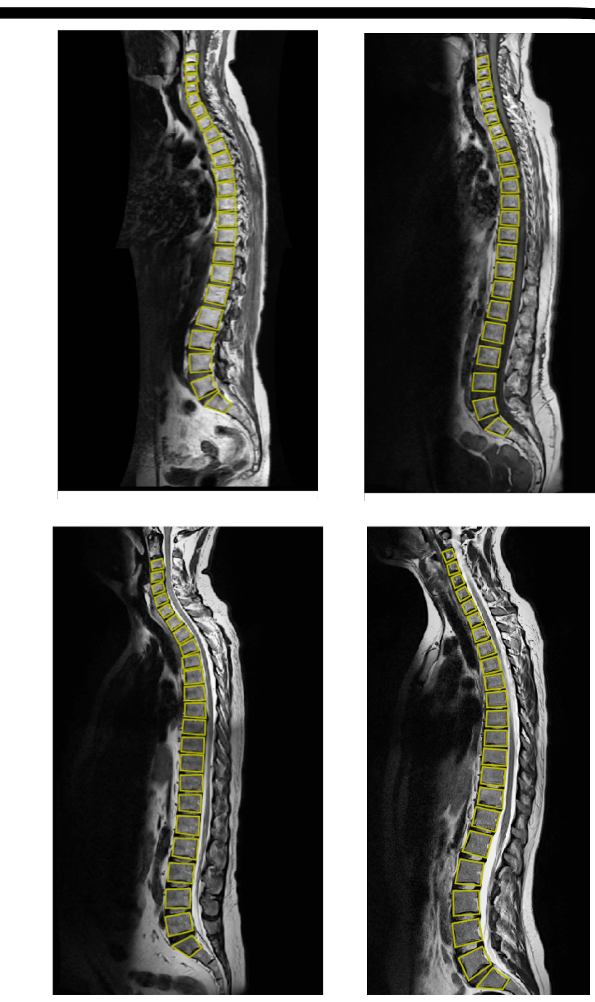
Each patch is resized by cubic interpolation to a 224×224 image. Random gaussian blur, rotations, crops and translations are used in augmentation.

4. Results

Current state of the art achieves 99.5% Dice score at vertebrate detection on lumbar scans [1]. We exceed this, achieving 99.8%.

More importantly, there is minimal drop inaccuracy when we apply same network to full spine scans with a Dice score of 99.4%.

Future work involves applying this to other imaging modalities (e.g. PET, CT) and tasks, (rib & teeth detection, temporal object tracking)



References: [1] Forsberg et al. *Journal of Digital Imaging 30, 2017*