Spondylolisthesis Grading in sagittal MRI scans using convolutional neural networks

Editors: Under Review for MIDL 2020

Abstract

Spondylolisthesis is a Latin term that is derived from "spondylos" meaning "spine" or "vertebra", and "listhesis" meaning "to slip or slide". Spondylolisthesis is a condition in which a component bone of spine (vertebrae) slips laterally out of place with respect to vertebra below it. This condition manifests clinically as lower back pain, stiffness or muscle spasms. Clinical treatment of Spondylolisthesis is guided by analysis provided by Radiologist on MRI (Magnetic Resonance Imaging) scans of relevant portion of spine. In this paper we propose an approach that uses CNNs to automatically detect Listhesis in spinal MRIs and provides extent of slippage both geometrically and in terms of medical grading of Listhesis. To do this, we use YOLO object detector to create loose bounding boxes around each vertebrae. Human vertebrae have a roughly rectangular shape. Rectangular clips of images, that loosely contain individual vertebrae are extracted and fed into a custom network to identify 4 corners of each vertebra. These corner points are then used to estimate extent of slippage between consecutive vertebrae. A commonly used medical criterion grades Listhesis in 4 classes. We use it to translate mathematical displacement calculated on images to medical grade for Listhesis.

1. Introduction

In this paper we use the term Listhesis to refer to vertebral displacements both in upper spine (the neck region) and the lower spine (lower back region of body). Vertebral displacements can happen due to several factors, including congenital conditions where bone is abnormally formed at birth; or from stress fractures, or due to age related degenerative changes of spine. Degenerative changes, by far are the most common reason for Listhesis and prevalence is estimated to be around 25% in women and 19.1% in men (Yi Xiang J. Wang, 2016) above the age of 50 years.

MRI scans of spine are modality of choice for identifying Listhesis and providing a grade that relates to the extent of displacement. This medical grading of Listhesis further guides the clinical procedure to be followed for remediation of the condition or its symptoms. For example, grade 1 and 2 slips do not generally require surgery and are treated medically, while grade 3 and above might require surgery in case of persistent pain experienced by the patient.

A MRI scan typically consists of parallel 2D images taken along 3 different planes i.e. Coronal (front to back), Sagittal (left to right sideways) and Axial (top to bottom). Series of images are taken across each plane which when stacked together create a 3D impression of the region. For diagnosis of Listhesis radiologists use sagittal series from MRI scans where they calculate the displacement of vertebrae with respect to vertebrae above it. In this paper, we use T2 weighted sagittal scan that lies approximately at the middle of the body.
We use this image because this is where the spinal canal and the central nerve root passing through it can be viewed most clearly. This is also the image that most Radiologists use to create diagnosis for vertebral slippage. Any slippage that results in pressure on the nerve passing through spinal canal is more likely to manifest clinical symptoms such as pain and disability.

Our approach for automatic detection and grading of Listhesis involves 3 technical stages that allow us to extract image features most relevant for medical interpretation. In the first stage we classify Cervical (neck region) MRI scans from Lumbar (lower back region) MRI scans. This is necessary because there are anatomical differences in vertebrae from these regions. Cervical vertebrae are smaller in size and vertebrae at the top are fused, whereas Lumbar vertebrae are larger and vertebrae at the bottom are fused while also having significant curvature towards the tail bone.

In the second stage, we use YOLO to detect individual vertebrae on mid-sagittal image and draw a loose bounding box around them. We use the same YOLO detector for both Cervical and Lumbar scans, since anatomical differences are not significant enough to require separate models when we only want to determine rough locations of vertebrae. From the bounding box, we clip rectangular images that roughly contain individual vertebra and process them separately. Since vertebrae from the same region are generally very similar looking in size and aspect ratio, this step greatly simplifies model structure that is needed for identification of corner points of vertebra.

In the third stage, we use a custom CNN that takes image clips that necessarily contain a single vertebra. This network is trained separately for image clips containing Cervical and Lumbar vertebrae, and outputs coordinates of 4 corners of vertebra contained in image clip. These coordinates are then used to calculate exact relative displacement of consecutive vertebrae and an established medical criterion is applied to arrive at Listhesis grading at each level.

2. Related Work

Deep Learning techniques are gaining increasing popularity for analysis of medical images. Analysis of spinal MRI images using deep Convolutional Networks has also been explored (Jamaludin et al., 2017). However, bulk of work is focused generally on coming up with technical classification, segmentation etc. without as much regard to clinical implications and usability. For instance, (Jamaludin, 2018) proposes an approach for detecting Listhesis at various vertebral levels by posing the problem as CNN based classification. Although this approach does provide insight into presence or absence of disease it cannot guide clinical course of action. Clinical decisions are based on more information pertaining to extent of displacement observed in the case. Smaller displacements might require simple measures such as dietary advice for weight loss and larger displacements might require administration of pain killers or surgery. This paper not only detect Listhesis but also provide actual extent of displacement along with medical grading for the displacement.

2.1. Deep Learning in MRI

Since the rise of deep learning and digitization of MRI scans, there have been numerous applications of deep learning in MRI scans. Typically deep CNNs are used for tasks such
Figure 1: Example of Cervical (top) and Lumbar (bottom) MRI scans. Often spinal MRI scans are not focused specifically on either Cervical or Lumbar regions but also include large portions of Thoracic spine. Fractures and different manifestations of spine degeneration also add to the variety. Hence a relatively deep CNN is required to differentiate Cervical from Lumbar scans.

as classification, segmentation, regression (Akkus et al., 2017), (Guo, 201), (MRI, a), to achieve state of the art performances. Deep learning is being applied to MRI scans for analysis of various diseases or abnormalities such as tumour segmentation in MRI scans of brain (Mohsen et al., 2012), prediction of Alzheimer (Suhuai Luo1, 2017) and detection of Spinal Stenosis through image segmentation (U Upadhyay and Singh, 2019).

3. Diagnosis of Listhesis

A regular human spine is made of 7 Cervical, 12 Thoracic, 5 Lumbar, 5 Sacral and 4 Coccyx vertebrae. These vertebrae are separated by inter-vertebral discs that act as a ligament to hold vertebrae together and allow for slight movement. Vertebrae are named using numbers that increment from head towards the toe, for example all 7 Cervical vertebrae are named as C1, C2, ..., C7; Thoracic vertebrae as T1...T12 and so on. The C2 vertebra has an upward-facing long bony process called the dens. The C1 sits atop and rotates around C2 below. Because of its ring like structure C1 vertebra is not clearly visible in MRI scans. Generally Sacral and Coccyx vertebrae are either fully or partially fused and therefore vertebral displacement, hence Listhesis, is not a relevant diagnosis for these regions of spine.

Even though objective criteria for reporting Listhesis (MRI, b; McTimoney CA, 2003) have been studied and clinically validated, they are not used in practice. Common practice for reporting Listhesis on a MRI scan is to subjectively assess relative displacement of consecutive vertebra, factor in any clinical symptoms and provide a grading. This is error
Figure 2: (Left) Meyerding Classification is a commonly adopted method for grading Spondylolisthesis. (Right) Naming convention for corners of vertebra and a geometric approach for calculating relative displacement.

Figure 3: Precise keypoints cannot be marked for vertebra with deformities such as fractures, osteophytes or in presence of rounded corners

prone for cases with irregularities in spinal curvature, vertebral fracture or simple image rotation.

Another factor for consideration while analysing Listhesis is direction of vertebral displacement. Displacement towards posterior side is called Retrolisthesis, while that towards anterior side is Anterolisthesis. Medically, Listhesis is reported with confidence on MRI scans with sagittal view of the spine. Sample MRI scans are shown in Fig. 1. As can be observed, Cervical scans include a portion of the head and neck while Lumbar scans extend to include the tail bone. Both Cervical and Lumbar scans can also include some thoracic vertebrae along with Cervical and Lumbar vertebrae respectively.

4. Method
4.1. Approach
In this paper we present an approach to automatically calculate grading of Listhesis at every vertebral level. The approach also calculates the actual displacement of a vertebra with respect to the one below in percentage. To arrive at a grading we use Meyerding Classification as shown in Fig. 2(Left) that associates a grade with relative displacement of consecutive vertebrae.

First, we define nomenclature used in this paper to refer to corners of vertebrae as shown in Fig. 2(Right). The corners are named as Anterior Superior (AS), Posterior Superior (PS), Anterior Inferior (AI) and Posterior Inferior (PI).
Now, the original medical criteria as defined in Meyerding Classification does not pre-
scribe a geometric method for calculating relative displacements of vertebrae, but only
provides a grade for approximate displacements in a range. For geometric calculation of
displacement, we drop a perpendicular from AI (or PI) corner of vertebra above to line
joining AS and PS corners of vertebra below to arrive at anterior (or posterior) displace-
ment. For example, in vertebrae shown in Fig. 2(Right) displacement is calculated as (see
equation 1).

\[
D_{12} = \frac{|AS2P|}{|AS2PS2|}
\]  

4.2. Dataset

A team of 3 senior Radiologists who were asked to provide independent assessments for
40 cases of Spinal MRI scans. Displacements calculated by dropping perpendicular from
inferior end plate of vertebra above had the maximum agreement with Radiologists’ grading
of Listhesis.
712 are used in training while rest are used for testing. Dataset was annotated by a team of radiographers in supervision of professional radiologists. They marked 4 corner points of each vertebra that can be used for objective evaluation of vertebral displacements. An annotation protocol was agreed upon in consultation with senior practicing Radiologists. This was necessary because every vertebra is different in shape, size, orientation and some may have blunted / rounded edges, osteophytes (bony growths), vertebral fractures or schmorl’s nodes (irregular indentations in vertebrae) making identification of exact corner point a non-trivial task. Fig. 3 shows sample vertebra where marking exact corner points required subjective judgement.

4.3. Architecture

Approach proposed to this problem is split into two separate models. First model is a YOLO v3 object detector (Redmon and Farhadi, 2018) which helps detects all the vertebrae present in sagittal scan. Vertebrae detected by YOLO v3 object detector is then cropped and interpolated for further computations. A custom CNN based model was used which
outputs 4 points corresponding to the corners of the vertebrae as shown in figure 4. This custom made model was trained on 2 different dataset to give us 2 models (lumbar and cervical).

A 2-way classification network is trained for differentiating between Cervical and Lumbar sagittal MRI scan. Network used is Alex-Net and trained on 1200 unique sagittal scans as training data with 50% distribution between Cervical and Lumbar sagittal scans. This network is used to decide whether the cropped vertebrae image from YOLO v3 object detector is to be fed to custom made model trained on Cervical vertebrae images or Lumbar vertebrae images.

Annotations for YOLO v3 object detector was made using the annotations of radiologists. Calculation of displacements depends greatly on the output of custom model it is important that model generates points with high precision. The points generated are later used for calculation of displacements which in turn led to grading of Listhesis. Paper compare the ratios calculated by the network proposed with the ratios calculated using the annotated points. Fig. 5 shows the full architecture of the proposed approach.

4.4. Training

Mean Squared Error loss function was used for custom model whereas Binary Cross Entropy loss function was used for training of 2-way classification network.

Each network was trained independently with normalized and resized and also with data augmentation by applying 7 different contrast levels. YOLO v3 was trained with image size of 416 × 416, custom model and classification model was trained on image size of 256 × 256. Learning rate for all three models was $1e^{-3}$ and to train Adam Optimizer (Kingma and Ba., 2015) was used. L2 regularization was used on weights to avoid over-fitting for custom and classification model.

5. Results and Conclusions

The paper measures performance of the networks using metrics such as mAP, precision, recall, accuracy. Performance of overall architecture is evaluated using two criteria - (1) Calculating distance between ground truth and output of vertebrae corner points (2) Comparing calculated displacements of vertebrae between ground truth and output. Table 1 shows performance of individual networks used. We calculated the accuracy of custom model by taking euclidean distance of ground truth vertebrae corner points and corre-
<table>
<thead>
<tr>
<th>Network</th>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>YOLO v3</td>
<td>mAP</td>
<td>70.12%</td>
</tr>
<tr>
<td>Alex-Net</td>
<td>Classification</td>
<td>99.6 %</td>
</tr>
<tr>
<td>Custom Model Cervical</td>
<td>Correct prediction</td>
<td>90.61%</td>
</tr>
<tr>
<td>Custom Model Lumbar</td>
<td>Correct prediction</td>
<td>94.67%</td>
</tr>
</tbody>
</table>

Table 1: Performance of individual networks on validation set

Spondylolysis: The slip. If the Euclidean distance between output and ground truth is less or equal to 2 units we consider it as correct prediction, otherwise we consider it as incorrect prediction. Model was trained on image size of 256 × 256. Validation set for custom Cervical model was 890 vertebrae images and for custom Lumbar model was 1019 vertebrae images. YOLO v3 was validated with 140 sagittal scans with distribution of 50% between Cervical and Lumbar scans. Classification model was validated with 586 sagittal scans with distribution of 40% and 60% between Cervical and Lumbar scans.

Proposed network was evaluated on test set containing 100 unique sagittal scans with 50% distribution between Cervical and Lumbar scans. Fig. 6 shows the distribution of predictions from the network on test set. Two histograms are plotted for each level of vertebrae - Cervical, Thoracic and Lumbar. The horizontal axis on left column represents the bin of distance values and vertical axis represents number of samples in test dataset that falls onto a particular bin. Plot on right column represents histogram for displacement value. Negative value indicates Retrolisthesis and positive value represents Anterolisthesis.

Distance histogram shows peak value at around 2 units whereas displacements histogram peaks between -5% to 5% for all 3 vertebrae levels. Fig. 7 shows performance of the network and its corresponding ground truth on sample vertebrae. Vertebrae with unclear corners, fractures or Osteophytes (bony lumps) are harder to detect. The network introduces standardization for the diagnosis of Listhesis unlike the existing measures taken by radiologists.

References

Deep-learning-based imaging-classification identified cingulate island sign in dementia with Lewy bodies. a.


