frontiers

UNIVERSITY OF ABERDEEN · PHILIPS · Canon CANON MEDICAL RESEARCH EUROPE LTD.

MathWorks · SINAPSE · aidence

27th Conference on Medical Image Understanding and Analysis 2023

Foresterhill, Aberdeen, Scotland

27th Conference on Medical Image Understanding and Analysis 2023

ISBN
9782832512319

DOI
10.3389/978-2-8325-1231-9

Citation
Waiter, G., Lambrou, T., Leontidis, G., Oren, N., Morris, T., Gordon, S., McGowan, J., Nicolson, C. (2023)

July 19–21, 2023, Foresterhill, Aberdeen, Scotland

Towards automatic scoring of spinal X-ray for ankylosing spondylitis

Author
Yuanhan Mo – Big Data Institute, University of Oxford, Oxford, UK
Yao Chen – Novartis Pharmaceutical Company, East Hanover, NJ, USA
Aimee Readie – Novartis Pharmaceutical Company, East Hanover, NJ, USA
Gregory Ligozio – Novartis Pharmaceutical Company, East Hanover, NJ, USA
Thibaud Coroller – Novartis Pharmaceutical Company, East Hanover, NJ, USA
Bart–lomiej W. Papiez – Big Data Institute, University of Oxford, Oxford, UK

Citation
Mo, Y., Chen, Y., Readie, A., Ligozio, G., Coroller, T., Papiez, B.W. Towards automatic scoring of spinal X-ray for ankylosing spondylitis.

Abstract
Manually grading structural changes with the modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS) on spinal X-ray imaging is costly and time-consuming due to bone shape complexity and image quality variations. In this study, we address this challenge by prototyping a 2-step auto-grading pipeline, called VertXGradeNet, to automatically predict mSASSS scores for the cervical and lumbar vertebral units (VUs) in X-ray spinal imaging. The VertXGradeNet utilizes VUs generated by our previously developed VU extraction pipeline (VertXNet) as input and predicts mSASSS based on those VUs. VertXGradeNet was evaluated on an in-house dataset of lateral cervical and lumbar X-ray images for axial spondylarthritis patients. Our results show that VertXGradeNet can predict the mSASSS score for each VU when the

data is limited in quantity and imbalanced. Overall, it can achieve a balanced accuracy of 0.56 and 0.51 for 4 different mSASSS scores (i.e., a score of 0, 1, 2, 3) on two test datasets. The accuracy of the presented method shows the potential to streamline the spinal radiograph readings and therefore reduce the cost of future clinical trials.

Introduction

X-ray imaging is one of the imaging modalities, utilized to monitor the structural progression of ankylosing spondylitis (AS), and this progression can be quantified by the modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS) [6]. Applying the mSASSS to each VU in an X-ray is usually performed manually by expert readers (radiologists) due to its complexity, making it a costly and tedious process. Moreover, such a manual process can introduce inter- and intra-reader variability in the final dataset. Therefore, to address these problems, we aim to propose an automatic pipeline for mSASSS grading.

In this paper, we propose a 2-step auto-grading pipeline, VertXGradeNet, for estimating mSASSS scores from the given spinal X-ray images. The proposed auto-grading pipeline was built on top of the previously developed VUs extraction pipeline [2] and predicts mSASSS scores based on extracted VUs. Then the proposed pipeline was validated utilizing clinical trial data from radiographic and non-radiographic axial spondyloarthritis patients.

Method

Inspired by work conducted by Koo et al. [5], a two-step grading pipeline is proposed (Figure 1) which uses a deep neural network called ResNet [4] as the backbones. In the test phase, given a VU extracted by our in-house VU extraction pipeline, the first step of the proposed pipeline is to detect if a given VU has a bony bridge or not (i.e., mSASSS = 3 versus others). If the given VU does not have a bony bridge, then the VU will proceed to the second step for further mSASSS prediction (i.e., mSASSS = 0,1 and 2). In the second step, the pipeline will produce two predictions on both the upper and lower corners of the given VU (e.g. mSASSS = 0, 1, or 2) that is not classified

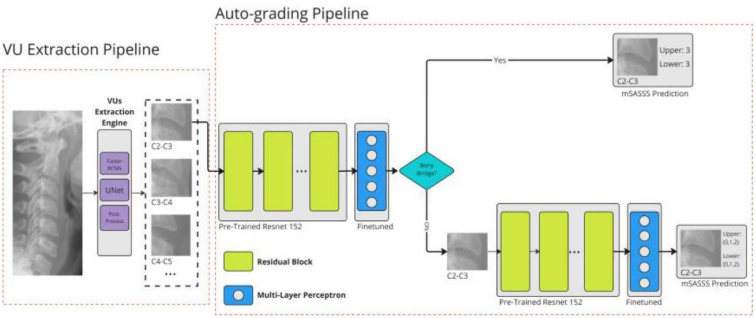

FIGURE 1
The initial step of vertebral units extraction Pipeline was developed previously [2]. The auto-grading pipeline was incorporated into the initial pipeline using the VUs as the input to predict the mSASSS. The auto-grading pipeline consists of two steps: the first step is to detect the bony bridge (mSASSS = 3) in the given VU, and the second step is to perform further mSASSS grading (i.e., mSASSS = 0,1 and 2) for a given VU without a bony bridge

as a bony bridge. Finally, each given VU will have the estimated mSASSS given by the auto-grading pipeline.

The ResNet152 [4] is used twice in the two-step auto-grading pipeline. In the first step, the ResNet 152 takes a VU as input and performs a binary classification, namely if the VU does or does not have a bony bridge. In the second step, Another ResNet 152 takes the remaining VUs (mSASSS $\leq$ 2) as the inputs and predicts the rest of the mSASSS scores for both the upper and lower corners of a VU.

In the training stage, a pre-trained ResNet 152 (on ImageNet) is fine-tuned on the extracted VUs/mSASSS scores from the MEASURE1 dataset. The training process of the two stages of the auto-grading pipeline is not end-to-end. Therefore, each step is trained independently.

Experiments and Results

Data. The method was developed based on the anonymized datasets of secukinumab radiographic and non-radiographic axial spondyloarthritis clinical studies (MEASURE 1 [1] and PREVENT [3]). A total of 7239 and 9313 VUs were successfully extracted from the two studies.

Experiments. Extracted VUs were randomly split into 5 folds at the patient level for MEASURE 1 and performed 5-fold cross-validation. The performance of the model, trained on MEASURE 1 data only, was evaluated on the PREVENT dataset. The ground truth mSASSS scores for the training VUs were provided by the clinical trial. The detailed results are shown in Table 1 and 2.

TABLE 1: Results of 5-fold cross-validation for MEASURE 1 dataset

mSASSS	Precision	Recall	AUC(ROC)	F1-score
0	0.934(0.010)	0.918(0.007)	0.897(0.011)	0.926(0.010)
1	0.200(0.097)	0.240(0.103)	0.809(0.076)	0.218(0.096)
2	0.390(0.069)	0.300(0.020)	0.849(0.021)	0.332(0.026)
3	0.654(0.067)	0.800(0.023)	0.959(0.009)	0.718(0.034)
Micro average	0.544(0.033)	0.564(0.032)	0.898(0.011)	0.548(0.032)
Macro average	0.860(0.014)	0.856(0.014)	0.879(0.020)	0.856(0.014)

TABLE 2: Results for PREVENT dataset

mSASSS	Precision	Recall	AUC(ROC)	F1-score	Support
0	0.99	0.95	0.97	0.825	15201
1	0.01	0.12	0.02	0.759	25
2	0.15	0.23	0.18	0.857	244
3	0.14	0.73	0.23	0.958	64
Micro average	0.32	0.51	0.35	0.826	15534
Macro average	0.97	0.93	0.95	0.850	15534

Conclusion

We have prototyped a 2-step auto-grading pipeline for automatic mSASSS scoring. The current approach, which now can be considered as a benchmark, improves the grading performance compared to the preliminary results. However, limited training samples and class imbalance issues still limit the current performance of the auto-grading pipeline. Thus, further analysis is required to address the aforementioned problems.

References

[1] Baeten, D., et al.: Secukinumab, an interleukin-17a inhibitor, in ankylosing spondylitis. New England journal of medicine **373**(26), 2534–2548 (2015)

[2] Chen, Y., Mo, Y., Readie, A., Ligozio, G., Coroller, T., Papiez, B.W.: Vertxnet: Automatic segmentation and identification of lumbar and cervical vertebrae from spinal x-ray images (2022). https://doi.org/10.48550/ARXIV.2207.05476, https://arxiv.org/abs/2207.05476

[3] Deodhar, A., Blanco, R., Dokoupilová, E., Hall, S., Kameda, H., Kivitz, A.J., Pod- dubnyy, D., van de Sande, M., Wiksten, A.S., Porter, B.O., et al.: Improvement of signs and symptoms of nonradiographic axial spondyloarthritis in patients treated with secukinumab: primary results of a randomized, placebo-controlled phase iii study. Arthritis & Rheumatology **73**(1), 110–120 (2021)

[4] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. CoRR **abs/1512.03385** (2015), http://arxiv.org/abs/1512.03385

[5] Koo, B.S., Lee, J.J., Jung, J.W., Kang, C.H., Joo, K.B., Kim, T.H., Lee, S.: A pilot study on deep learning-based grading of corners of vertebral bodies for assessment of radiographic progression in patients with ankylosing spondylitis. Therapeutic Advances in Musculoskeletal Disease **14**, 1759720X221114097 (2022)

[6] Van Der Heijde, D., Braun, J., Deodhar, A., Baraliakos, X., Landewé, R., Richards, H.B., Porter, B., Readie, A.: Modified stoke ankylosing spondylitis spinal score as an outcome measure to assess the impact of treatment on structural progression in ankylosing spondylitis. Rheumatology **58**(3), 388–400 (2019)