A Simple but Effective Training Process for the Few-shot Prediction Task of Early Rheumatoid Arthritis from MRI

Yanli Li[1] Y.LI2@LUMC.NL
[1] *Division of Image Processing, Dept of Radiology, Leiden University Medical Center (LUMC)*
Denis Shamonin[1] D.P.SHAMONIN@LUMC.NL
Tahereh Hassanzadeh[1] T.HASSAN_ZADEH_KOOHI@LUMC.NL
Monique Reijnierse[2] M.REIJNIERSE@LUMC.NL
[2] *Department of Radiology, Leiden University Medical Center (LUMC)*
Annette van der Helm-van Mil[3] A.H.M.VAN_DER_HELM@LUMC.NL
[3] *Department of Rheumatology, Leiden University Medical Center (LUMC)*
Berend C. Stoel[*1,2] B.C.STOEL@LUMC.NL

Editors: Under Review for MIDL 2022

Abstract

Predicting rheumatoid arthritis (RA) in an early-stage based on MRI can help initiate timely treatment and therefore halt the progression of the disease and increase the possibility of recovery. Deep learning methods are in general highly suitable for this type of labeling tasks. However, applying this approach to RA detection faces challenges from the lack of a large number of samples, difficulty in distinguishing patterns of RA from imaging artifacts, and a wide anatomical variation, leading to the failure of transfer learning based on pre-trained models. In this paper, a pre- and post-training method for this few-shot task is proposed. Based on the clinical MRI data, this method was validated through cross-validation, achieving a significant improvement in AUC, F1 score, and accuracy to the baseline deep learning models. Since these pre- and post-training strategies are intuitive, effective and easy to implement, they can also contribute to other challenging few-shot medical tasks.

Keywords: Deep learning, few-shot learning, rheumatoid arthritis diagnosis, MRI

1. Introduction

Predicting rheumatoid arthritis (RA) through deep learning methods based on magnetic resonance imaging (MRI) faces serious overfitting, which requires numerous labeled data. However, this is not realistic for most clinical tasks. Moreover, pre-training models based on natural images also fails in this few-shot task.

The purpose is therefore to provide a framework to train neural network models for early RA prediction. As a first step, the aim of this paper is to distinguish patients with clinically suspect arthralgia (CSA) from normal controls. Within this framework, a masked auto-encoder reconstruction is introduced to stimulate the model to learn the underlying patterns. Besides, an augmentation-based consistency loss function is proposed to make the model independent of irrelevant information in the data.

* Corresponding author

2. Methodology

The MRI scans contain both coronal and transversal 3D scans from the wrists of patients from the CSA group, and the control group. Considering the total number of MRI scans (789 patients and 177 controls), the sample size, and low resolution in the Z-direction (512 × 512 × 20), the central slices of coronal scans were used, which can also benefit from a pre-trained model.

2.1. Masked autoencoder-based reconstruction

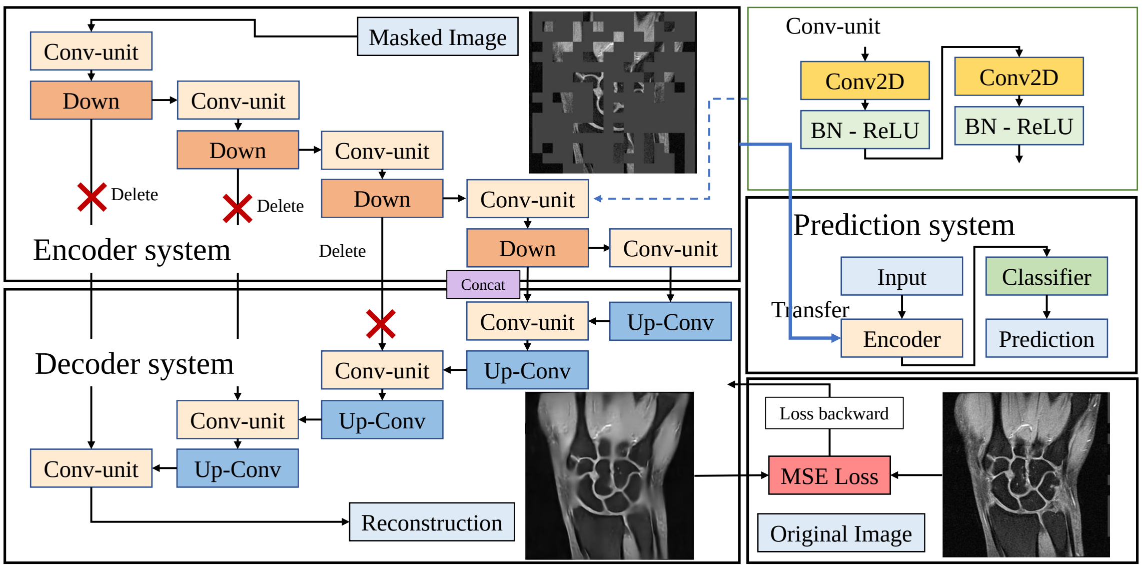

Figure 2.1: The workflow of the reconstruction agent task design.

Inspired by the masked auto-encoder (He et al., 2021) and the model genesis (Zhou et al., 2019), as shown in Figure 2.1, we defined an agent task that requires the auto-encoder model to reconstruct the original images with randomly masked images as input.
The model can achieve a good reconstruction on the masked images from the test set with a mask ratio of 75%. After this pre-training process, we transferred the parameters of the encoder of this reconstruction model to the RA prediction model.

2.2. Augmentation-based consistency loss post-training process

While the reconstruction pre-training process is trying to get the model well-initialized, this post-training process is aimed at improving the consistency of output predictions of images before and after using augmentation methods.
This method is based on an observation that there is a significant inconsistency between the prediction output probabilities of an image and its augmented image, which reveals that much irrelevant spatial information remains in the model output.

Instead of inputting the original image and augmented image separately in two epochs, we propose to individually input both of them into the model in one epoch, and then calculate a mean squared error loss between the output prediction of these two images. The procedure is shown in Figure 2.2.

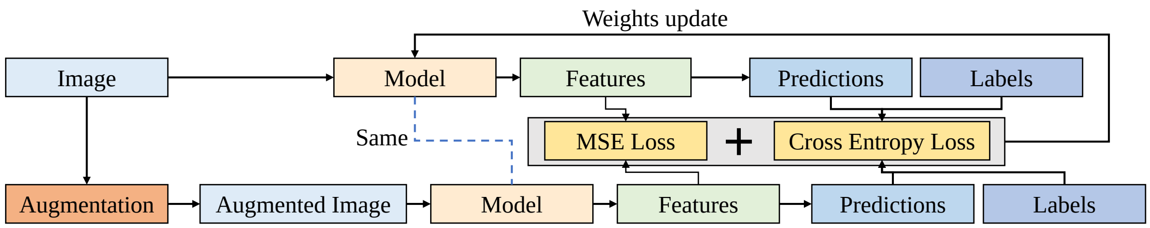

Figure 2.2: The workflow of the post-training procedure.

3. Experimental Results

The results shown in Table 1 are the average performances from the standard 5-fold cross-validation. The baseline model is the VGG16 with normalization, which has a similar encoder to U-Net that can be easily modified to an autoencoder structure. Besides, after some early attempts, we found that VGG-based models have overall better performances than ResNet-based models.

Table 1: the average performances

Spatial augmentation	Masked reconstruction	ImageNet pretrain	Consistency loss	AUC	F1 score	Acc
				0.72	0.62	66.7%
Y				0.75	0.58	69.2%
Y		Y		0.69	0.64	65.5%
	Y			0.79	0.73	72.4%
Y	Y			0.80	0.72	73.2%
Y			Y	0.79	0.73	74.5%
Y	Y		Y	0.81	0.75	76.2%

According to Table1, the model with masked reconstruction pre-training and augmentation-based post-training shows a significantly better performance than the baseline model and ImageNet pre-trained model, which illustrates the effectiveness of this framework.

4. Conclusion

We proposed a framework to train neural network models for early RA diagnosis, which can also be easily extended to other few-shot challenging diagnosis tasks. According to the AUC, F1 score, and accuracy of all the models, this proposed framework significantly improves the prediction results of the deep learning models.

Acknowledgments

MRI scans were part of the Early Signs in MRI, Identifying Rheumatoid Arthritis (ESMIRA) project (NWO-TTW-HTSM 13329). This work is supported by the China Scholarship Council No.202108510012.

References

Kaiming He, Xinlei Chen, and et. al Xie, Saining. Masked autoencoders are scalable vision learners. *arXiv preprint arXiv:2111.06377*, 2021.

Zongwei Zhou, Vatsal Sodha, and et. al Rahman Siddiquee. Models genesis: Generic autodidactic models for 3d medical image analysis. In *International conference on medical image computing and computer-assisted intervention*, pages 384–393. Springer, 2019.