SIIM23 ANNUAL MEETING + InformaticsTECH Expo JUNE 14-16 | AUSTIN, TX
Learning. Innovating. Belonging.

Development of An AI Model to Classify Shoulder Radiographs to Build Large-scale Image Registries

Linjun Yang, PhD, Data Science Analyst, Mayo Clinic
Rodrigo de Marinis, MD; Jacob Oeding; Erick Marigi, MD; Joaquin Sanchez-Sotelo, MD

Introduction

Linking shoulder radiographs and related imaging information to shoulder arthroplasty registries would tremendously increase the clinical and scientific values of the registries. Although DICOM tags provide abundant imaging information, they typically do not provide shoulder-arthroplasty-specific imaging features. Manually identifying and recording those features on a large scale can be laborious and error-prone. AI algorithms can be trained to annotate problem-specific imaging features at an expert level. This study aims to develop an AI algorithm to automatically analyze shoulder radiographs and identify key imaging features related to shoulder arthroplasty.

Hypothesis

An AI classification algorithm can be developed to analyze shoulder radiographs and identify shoulder sides, imaging views, the presence or absence of implants, and implant types.

Methods

A set of 920 shoulder radiographs were identified. As shown in figure 1, each image was labeled manually of shoulder sides (right/left), views [anterior-posterior (AP)/auxiliary], and implant information [preoperative/anatomic shoulder arthroplasty (ASA)/reverse shoulder arthroplasty (RSA)]. The set was divided into training, validation, and testing sets using a stratified split. The EfficientNet-b3 model pre-trained using hip radiographs was fitted and saved with the highest average F1 score on the validation set. The performance of the saved model of identifying different sides, views, and implants was evaluated on the testing set.

Results

It took the model 5.2 seconds to analyze 241 testing images. As a highlight of table 1, the F1 score in predicting the left and right sides was 0.970 and 0.981, respectively. The F1 score in predicting AP and auxiliary views was 0.975 and 0.976, respectively. The F1 score in predicting preoperative, ASA, and RSA implant types was 0.982, 0.972, and 0.972, respectively.

Conclusion

An AI model was developed to accurately identify key features in shoulder radiographs. The model can be further improved using a larger training set and used to efficiently build shoulder radiograph registries.

Figure(s)

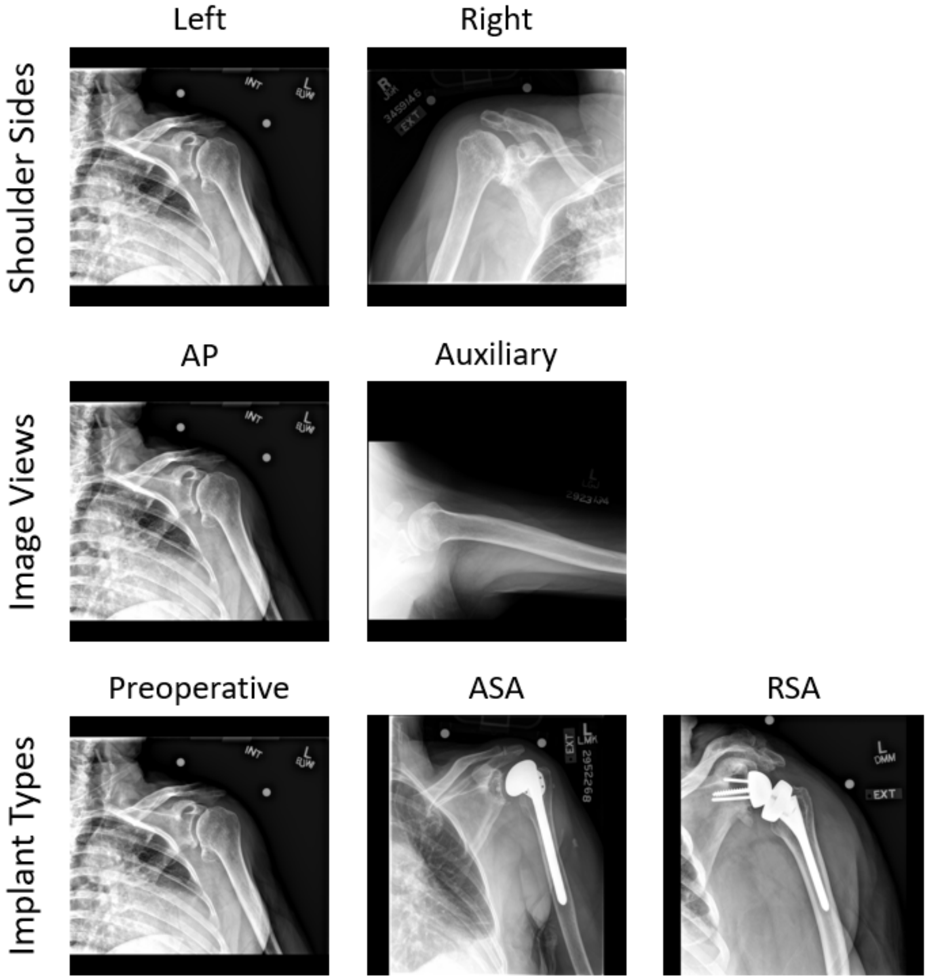

Figure 1. Illustration of key imaging features of shoulder arthroplasty radiographs: shoulder sides, imaging views, and implant types.

	Shoulder Side		**Imaging View**		**Implant Types**		
	left	right	AP	Auxiliary	Preop	ASA	RSA
Accuracy	0.977	0.981	0.969	0.989	0.974	0.963	1.000
F1 Score	0.970	0.981	0.975	0.976	0.982	0.972	0.972
Recall	0.967	0.981	0.981	0.967	0.991	0.981	0.950
Precision	0.977	0.981	0.969	0.989	0.974	0.963	1.000

Table 1. Result of the model evaluation on the testing set for each category reported in accuracy, f1 score, precision, and recall.

Keywords

Artificial Intelligence/Machine Learning; Clinical Workflow & Productivity; Imaging Research