Scientific Abstract Template

2018 SIIM Conference on
Machine Intelligence in Medical Imaging
September 9-10, 2018 | Hilton SF Financial District | San Francisco, CA

First Author Name (Include Credentials)
Tony Lindsey, PhD, antonia.e.lindsey@nasa.gov

Additional Authors (Include First, Middle Initial, Last, Credentials & Email)
Rebecca Sawyer Lee, PhD, rebecca.l.lee-1@nasa.gov

First Author Institution
NASA Ames Research Center

Topic
☐ Datasets
☒ Machine Learning Algorithms and Toolkits
☐ Clinical Applications
☐ Non-pixel-based Machine Learning in Medical Imaging Projects
☐ Infrastructure to Support Machine Learning

Preferred Method of Presentation
☐ Paper Presentation (8 minutes; suitable for most papers)
☒ Poster (4'x8' poster; front illuminated for detailed math and/or fine image comparison)

Title
Automated Semantic Segmentation for Volumetric Cardiovascular Feature Quantification and Pathology Assessment

Introduction
We present a pipeline method that curtails the expense and observer bias of manual cardiac evaluation by combining semantic segmentation and disease classification as a fully automatic processing pipeline, see Figure 1. The initial element consists of a 2D U-Net convolutional neural network architecture for voxel-wise segmentation of the myocardium and ventricular cavities. The results of the segmentation were used to compute a comprehensive volumetric feature matrix that captured diagnostic clinical procedure data and that was used to model a cardiac pathology classifier.

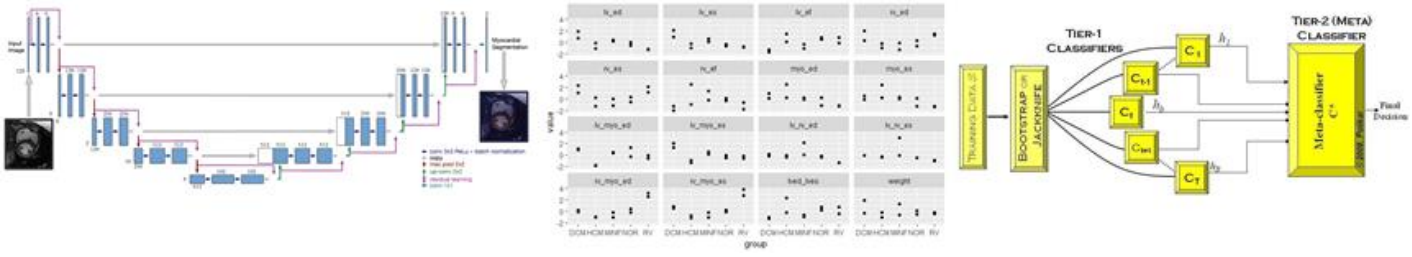

Figure 1: Voxel-wise segmentation quantifies volumetric cardiac features assessed for pathology by ensemble meta-classifier.

Our approach evaluated anonymized parasternal MRI cardiac images from a database of 100 patients (4 pathology groups, 1 healthy group, 20 patients per group) examined at the University Hospital of Dijon. We achieved top average Dice index scores of 0.939, 0.849, 0.886 for structure segmentation of the left ventricle (LV), right ventricle (RV) and myocardium respectively. A 5-ary pathology classification accuracy of 90% was recorded on an independent test set using our trained model.

Methods

The dataset of images was split into equal halves as train and validation sets for the segmentation phase. Data augmentation with affine transforms during each training epoch was performed to improve network localization and mitigate overfitting. The Adam optimizer with L2-regularization, and Xavier initialization were used for model training. Pairs of cardiac short axis cine-MRI end diastolic and systolic training images were processed for 10,200 iterations over 30 hours on a Tesla K80 GPU to generate a model. The network was optimized with mini-batches containing 16 images and default learning rate of 0.01 during each iteration.

The segmentation information obtained from our deep learning model was used to compute volumetric features for the left ventricle, right ventricle, and the myocardium. Feature selection was performed in order to determine the most discriminatory attributes for classification of the 5 categories. Features were ranked by importance based on the amount of mean Gini impurity decrease. Breiman's 1-SE rule was utilized to select 5 features for building the parsimonious pathology classification model, see Figure 2.

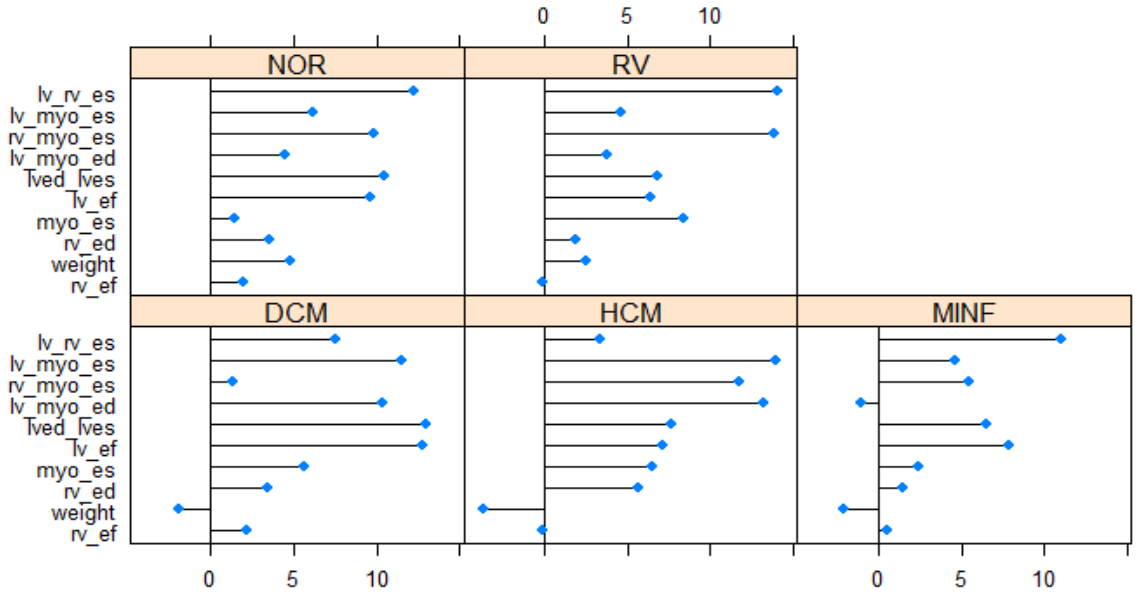

Figure 2: Feature importance ranked by Gini impurity decrease and **lv_rv_es**, **lv_myo_es**, **rv_myo_es**, **lved_lves**, **lv_ef** selected.

Results

Clinical evaluation measures resulting from the 2D U-Net architecture are shown in Table 1.

2D U-Net	LV (ED)	LV (ES)	RV (ED)	RV (ES)	MYO (ED)	MYO (ES)
ASSD	0.359 (0.414)	0.885 (0.896)	0.987 (1.324)	2.816 (1.647)	0.536 (0.281)	0.713 (0.450)
DICE	**0.962 (0.021)**	**0.917 (0.063)**	**0.914 (0.075)**	**0.783 (0.120)**	**0.881 (0.037)**	**0.891 (0.034)**
HD	6.072 (3.529)	9.703 (5.144)	14.899 (9.158)	21.383 (9.868)	11.053 (7.991)	13.316 (9.019)

Table 1: Segmentation measures for 2D U-Net architecture. Each entry represents the mean (std) value accuracy measure obtained for specific structure and cardiac phase.

A resulting feature matrix contained 13 cardiac volume features including permutation ratios, and 2 ejection fractions plus weight for each of 50 patients. The parsimonious model was built and trained using ensemble stacking of several learning algorithms (SVM, Random Forests, PCA and Bagging). The average performance measures across all classes include 90% accuracy, 90% sensitivity and 97.5% specificity for the test set data.

Discussion

One limitation to our current study is the size of our dataset, additional image data can provide a more accurate assessment of our system's performance. Several methods will be explored to develop a more robust

pipeline including Bayesian optimization, contrast enhancement preprocessing and GANs for training data augmentation.

Conclusion

In this project report we presented a fully automatic processing pipeline for pathology classification on cardiac cine-MRI. The system achieved higher accuracy on 5-ary classification compared to a baseline robust pre-trained CNN architecture with transfer learning.

Keywords

Deep learning, Cardiac cine-MRI, Semantic Segmentation, Feature Selection, Classification