Automated Classification of Aortic Valve Morphology from Phase-Contrast Cardiac MRI Using an Augmented CNN

Anonymous Author(s)
Affiliation
Address
email

Abstract

Bicuspid aortic valve (BAV) is most common congenital malformation of the heart occurring in 1-2% of all individuals, and is associated with a variety of poor health outcomes. Exploring BAV in large unlabeled imaging datasets such as the UK Biobank is complicated by the absence of functional models for classification of aortic valve morphology. Fully automated identification of BAV from MRI images was not previously studied in the medical literature. In this paper, we show that with an extremely limited set of 112 samples, subjected to simple pre-processing steps along with data augmentation, we can train a CNN classification model which achieves an AUC of 0.76, a 0.14 lift from non-augmented data, thus establishing a baseline for automatic BAV identification from cardiac MRIs.

1 Introduction

Bicuspid aortic valve (BAV) is a highly prevalent cardiac malformation present in 1-2% of individuals. As it may go undiagnosed even in adulthood [1], its early detection can help pinpoint other conditions such as congenital aortic stenosis, hypoplastic left heart, and coarctation of the aorta [2, 3]. In 2006-2010, the UK Biobank (UKBB) recruited 502,638 participants aged 37-73 years and made publicly available comprehensive health-related data including genome-wide data, and medical imaging data in the form of cardiac MRI [4]. The availability of large unstructured medical imaging datasets allows the possibility of performing a variety of genetic and epidemiological studies related to anatomical features within the dataset. However there are no extant software tools for unsupervised classification of anatomical structures from cardiac MRI data. Here we describe a computer vision approach incorporating a small labeled training dataset and data augmentation to classify individual participants with bicuspid aortic valve (BAV) within the cardiac MRI data in the UK Biobank.

2 Methods

The data consists of dicom files of cardiac MRI sequences [5] including eight sets of cardiac images – coronal and transverse scout imaging, cine-images from 2-chamber, 3-chamber, 4-chamber, left-ventricular-outflow tract, short axis at 9 separate planes, aortic valve phase-contrast imaging, and measures of regional wall motion and diastolic function. Phase-contrast images processed for magnitude of flow across the aortic valve taken in the short axis plane were selected for analysis of aortic valve morphology [6]. Individual images consisted of 30 frames of 192 by 192 pixels. Training data was imbalanced, reflecting the prevalence of BAV within the general population, at 100 healthy subjects and 12 with bicuspid aortic valves. As frame brightness is proportional to the magnitude of flow through the aortic valve, we selected a sequence of the six brightest frames from each series.

Submitted to 31st Conference on Neural Information Processing Systems (NIPS 2017). Do not distribute.

The primary model included base CNN layers of the VGG 16 model [7], added with 5-layer feed-forward deep neural network model with one node in the output layer. Stacked convolutional neural network layers can capture high level concepts such as edge detection, so we employed pre-trained CNN layers to extract features from gray scale cardiac MRI data. The FNN layers of VGG 16 model were constructed for classification of 1000 common objects, therefore we employed a separate FNN top model for classification of aortic valve morphology. For the input of 192 x 192 pixel cardiac MRI data used in this project, the CNN layers of the VGG 16 model output 512 layers of feature images with the size of 6 x 6 pixels. The feature image size gets smaller, in concordance with the size reduction of the input image, but the number of feature image layers (512) will stay the same.

Due to the small number of MRI samples, the CNN layers may easily overfit when using a complex FNN model. Preliminary experiments suggested a relatively simple FNN structure; 5 feed-forward layers with the number of each layer being 128, 64, 32, 16, 1. We used binary cross entropy as the loss function and the ADAM optimizer. Finally, to deal with the training data imbalance, and to penalize false negatives more than false positives, we introduced class weights in the loss function.

Basic pre-processing of the images to eliminate extraneous information led to increased performance of the deep learning model. A black border surrounding the frames due to acquisition or processing artifact differed between patients and contained no relevant information or features. Therefore we eliminated the 14 pixels on the periphery of each image yielding a borderless image of 164x164 pixels. Additionally the range of pixel intensity values varied from patient to patient, and thus were normalized to give a pixel value range of 0-255 per frame using the Pillow-4.3.0 package PIL.Image module. After normalization, we performed Gamma-correction [8] (Gray Level Transformation) to enhance contrast of high intensities ($f(x) = cx^r$, r=2) (255.0 x (pixels/255.0)^2).

We tested two different augmentations: translations of 2-4 pixels yielding 36 to 49 additional images, and shearing/stretching of up to 20% of the image size yielding 495 additional augmented images.

3 Results

We tested the model using 4-fold cross validation, where, within each fold, 50% of the data was used for training, 25% for tuning and the remaining 25% for testing. Table 1 shows the performance of the models obtained through each transformation. We report the area under the precision-recall curve (AUC), accuracy (acc), weighted accuracy (w-acc), negative weighted binary cross entropy (neg_ce), the F3 measure (f3), precision (prec), recall (rec), and the confusion matrix.

Table 1: Performance metrics for each of the tested augmentations. Original – model uses untransformed data. OTA – translation augmentation of original data. B – border removal. BTA – border removal and translation. GC – gamma correction. GCTA – gamma correction with translation augmentation. GCSA – gamma correction with stretch augmentation.

	auc	acc	w-acc	neg_ce	f3	prec	rec	tp	tn	fp	fn
O	0.62	0.90	0.3275	-11.04	0.26	0.5	0.25	0.75	24.5	0.5	2.25
OTA	0.58	0.89	0.255	-12.23	0.17	0.33	0.17	0.5	24.5	0.5	2.5
B	0.61	0.90	0.3275	-11.04	0.26	0.38	0.25	0.75	24.5	0.5	2.25
BTA	0.65	0.89	0.3975	-10.13	0.35	0.69	0.33	1	24	1	2
GC	0.70	0.92	0.4775	-8.65	0.43	0.58	0.42	1.25	24.5	0.5	1.75
GCTA	0.75	0.88	0.6175	-7.11	0.56	0.48	0.58	1.75	23	2	1.25
GCSA	0.76	0.89	0.62	-6.97	0.58	0.613	0.58	1.75	23.25	1.75	1.25

We described a method to classify BAV from phase contrast images from MRI, using a small training dataset and applying simple methods of augmentation to improve the predictive capabilities of the classifier. Pre-processing steps reduced inter-subject variability (border removal and normalization) and trimmed information irrelevant to the classification task (border removal and gamma correction). The VGG16/FNN model on the original data displayed a modest performance (AUC=0.6175). Pre-processing steps alone improved performance (AUC=0.70), suggesting that low-level intuitive steps to clean the input data are key to optimizing performance in medical imaging classification. When translation and stretch augmentations were applied to pre-processed data, the model characteristics were additionally improved (AUC=0.76), obtaining a baseline for automated BAV classification. Interestingly, improvements to the predictive characteristics of the VGG16/FNN model by both translation and stretching augmentation tasks appeared to be dependent on simple pre-processing.

References

[1] William C Roberts and Jong M Ko. Frequency by decades of unicuspid, bicuspid, and tricuspid aortic valves in adults having isolated aortic valve replacement for aortic stenosis, with or without associated aortic regurgitation. *Circulation*, 111(7):920–925, 2005.

[2] Ilenia Foffa, Lamia Ait Alì, Paola Panesi, Massimiliano Mariani, Pierluigi Festa, Nicoletta Botto, Cecilia Vecoli, and Maria Grazia Andreassi. Sequencing of notch1, gata5, tgfbr1 and tgfbr2 genes in familial cases of bicuspid aortic valve. *BMC medical genetics*, 14(1):44, 2013.

[3] Marc-Phillip Hitz, Louis-Philippe Lemieux-Perreault, Christian Marshall, Yassamin Feroz-Zada, Robbie Davies, Shi Wei Yang, Anath Christopher Lionel, Guylaine D'Amours, Emmanuelle Lemyre, Rebecca Cullum, et al. Rare copy number variants contribute to congenital left-sided heart disease. *PLoS genetics*, 8(9):e1002903, 2012.

[4] Naomi E Allen, Cathie Sudlow, Tim Peakman, Rory Collins, et al. Uk biobank data: come and get it, 2014.

[5] Steffen E Petersen, Paul M Matthews, Fabian Bamberg, David A Bluemke, Jane M Francis, Matthias G Friedrich, Paul Leeson, Eike Nagel, Sven Plein, Frank E Rademakers, et al. Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of uk biobank-rationale, challenges and approaches. *Journal of Cardiovascular Magnetic Resonance*, 15(1):46, 2013.

[6] Steffen E Petersen, Paul M Matthews, Jane M Francis, Matthew D Robson, Filip Zemrak, Redha Boubertakh, Alistair A Young, Sarah Hudson, Peter Weale, Steve Garratt, et al. Uk biobank's cardiovascular magnetic resonance protocol. *Journal of cardiovascular magnetic resonance*, 18(1):8, 2016.

[7] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[8] Jan Erik Solem. *Programming Computer Vision with Python: Tools and algorithms for analyzing images*. " O'Reilly Media, Inc.", 2012.