Clinical Motivation

Aortic stenosis (AS) is a common cardiac valve condition, best detected using echocardiograms (ultrasound images of the heart).

Many patients are missed by current practice - Up to 66% of symptomatic AS patients may not be referred for care

Improved early detection of AS sorely needed - If undetected, severe AS is often fatal (higher mortality than some metastatic cancers)

With timely detection, severe AS is treatable with low mortality

Automating preliminary screening of echocardiograms for AS via machine learning may improve detection (and thus improve outcomes).

Contributions

1. New open-access dataset: TMED
2. Analysis of recent SSL classifiers - What works on medical images?
3. Methods for coherent patient diagnosis from many images

Future Work

Data represents only 1 site. External validation needed.

Try multi-task SSL

Go beyond AS: more diagnoses and detailed measurements

Open-Access Dataset Release: TMED

Authentic benchmark for vision methods that learn from limited labeled data

Two classification tasks relevant to automatic diagnosis of aortic stenosis
1. Classify the view type of each image
2. Diagnose AS severity of each patient

Existing public echo datasets (EchoNet or CAMUS) are great, but not suitable for AS diagnosis.

Solution: Semi-supervised Learning (SSL)

1000s of unlabeled studies (easy to acquire)

100s of labeled studies (expensive to acquire)

Number of Unlabeled
Patients Images Method
0 0 Basic WRN
380 ~41k Pseudo-Label ([Lee et al. ’13])
380 ~41k VAT (Miyato et al. ’18)
380 ~41k Augment-Only MixMatch (Berthelot et al. ’19)
380 ~41k MixMatch

View Task Balanced acc. on test set
81.97
84.23
87.31
88.75
91.11

Takeaways:
- Modern SSL can use a large uncurated unlabeled set to boost performance over using only the modest-size labeled set.
- Among several methods, MixMatch is particularly effective.

Solution: Prioritize Relevant Views

Aggregation across images
Simple average
Prioritize relevant view

Takeaways:
- Using view and diagnosis classifiers together can improve diagnosis.
- Manually curating relevant views is not necessary.

Background: Echocardiogram Workflow

Handheld transducer is used to capture different views of the heart’s anatomy. There are dozens of standard view types.

One study yields ~100 images of diverse view and quality. Images are not labeled with view type or diagnosis.

Challenge: Lack of labeled data

Most classifiers require large training sets of labeled images to be successful

Echocardiogram imagery is easy to collect from existing records
- View and diagnostic labels not recorded when imagery is captured
- Require post-hoc annotation by clinical experts

Recent SSL methods show promise on standard vision tasks (e.g. CIFAR-10)
But use class-balanced data and artificially forget labels to make unlabeled set

Can SSL methods handle an uncurated unlabeled set of real medical images?

Challenge: Predict diagnosis from many images

Most classifiers are designed to take in only one image and predict its class.

One echocardiogram study of one patient produces ~100 diverse images.
- Only some view types are relevant (e.g. PLAX and PSAX show the aortic valve)
- Labels identifying which images are relevant are not available

Clinicians can take in many uncurated images, identify which ones are relevant views, and aggregate information from relevant images to make a diagnosis.

Can we automate diagnosis from many images?

Takeaways:
- Using view and diagnosis classifiers together can improve diagnosis.
- Manually curating relevant views is not necessary.

Solution: Prioritize Relevant Views

Aggregation across images
Simple average
Prioritize relevant view

Takeaways:
- Using view and diagnosis classifiers together can improve diagnosis.
- Manually curating relevant views is not necessary.

Future Work

Data represents only 1 site. External validation needed.

Try multi-task SSL

Go beyond AS: more diagnoses and detailed measurements

Open-Access Dataset Release: TMED

Authentic benchmark for vision methods that learn from limited labeled data

Two classification tasks relevant to automatic diagnosis of aortic stenosis
1. Classify the view type of each image
2. Diagnose AS severity of each patient

Existing public echo datasets (EchoNet or CAMUS) are great, but not suitable for AS diagnosis.

Solution: Semi-supervised Learning (SSL)

1000s of unlabeled studies (easy to acquire)

100s of labeled studies (expensive to acquire)

Number of Unlabeled
Patients Images Method
0 0 Basic WRN
380 ~41k Pseudo-Label ([Lee et al. ’13])
380 ~41k VAT (Miyato et al. ’18)
380 ~41k Augment-Only MixMatch (Berthelot et al. ’19)
380 ~41k MixMatch

View Task Balanced acc. on test set
81.97
84.23
87.31
88.75
91.11

Takeaways:
- Modern SSL can use a large uncurated unlabeled set to boost performance over using only the modest-size labeled set.
- Among several methods, MixMatch is particularly effective.