Semi-Supervised Detection of Chest Diseases using Generative Adversarial Networks

Noor Sajid

University College London
noor.sajid.18@ucl.ac.uk

Supervisers: Dr. Waty Lilaonitkul, Prof. Paul Taylor and Prof. Daniel Alexander

January 30, 2019

Overview

Background

- For many people, chest diseases can be an extremely serious health problem. These include tuberculosis, pulmonary disease, pneumonia, lung diseases, etc. [1]
- Chest radiography has become an important tool used for diagnosing these heart and lung diseases [2, 3].
- Automatic and timely detection of these with high accuracy could greatly enhance real world diagnosis processes.
- However, lack of standard publicly available labeled data-sets and benchmark studies make this difficult.

How?

"We approached the problem by employing a **semi-supervised detection of chest diseases** using generative adversarial networks."

Generative Adversarial Networks

Generative adversarial networks are an unsupervised deep learning algorithm introduced by [4].

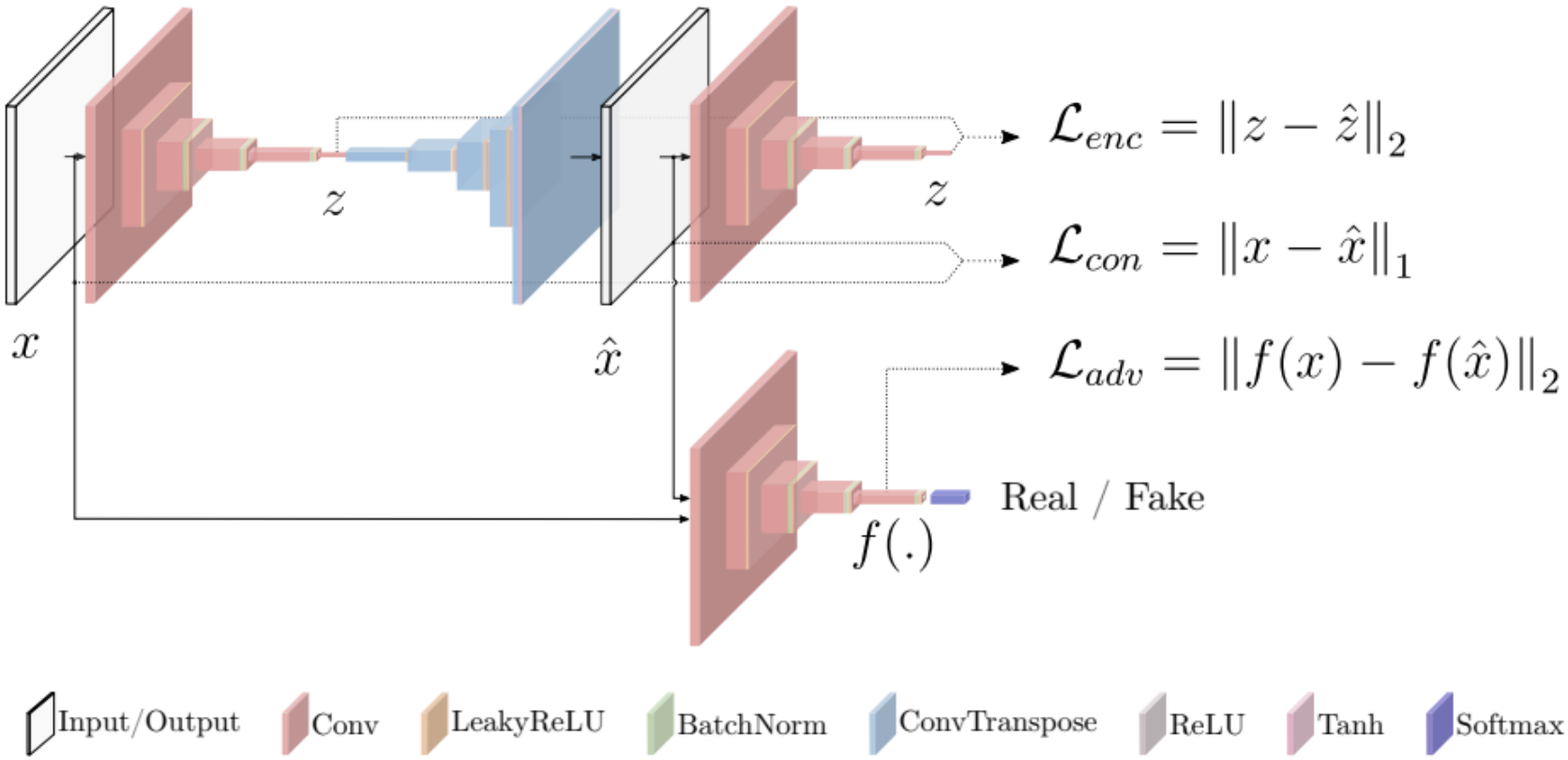
- Primary objective was to create realistic images [5].
- Two competing networks; the generator and discriminator
- During training the generator creates an image and the discriminator determines whether that image is real or fake.
- From this, the GAN could learn the true distribution of the input images from a latent space and the input data could then be mapped to this latent space using a pre-trained model

Generative Adversarial Networks with Inference

[6, 7, 8] make use of the latent space representation from GANs to detect anomalies. For our case, this can be achieved by using a data-set that is highly biased towards **normal lung chest x-rays**.
Propose using a conditional-GAN that learns both the high dimensional image space generation and the latent space inference [8]:

- The **encoder-decoder-encoder** structure within the generator network allows the model to input the image to a lower dimension vector. This can be used to reconstruct the output image.
- The second encoder allows for the image to be generated to it's latent space.
- By reducing the distance between the image and the latent space during training helps learn a true normal lung distribution
- The larger the distance from the learned distribution at inference, the more indicative that the incoming data is an anomaly.

What does this look like?



Learn both the normal data distribution and minimises the output anomaly score

Anomaly Detection

Encoder loss used for evaluating abnormality. The overall anomaly performance is computed via the anomaly score for each patch image. Feature scaling is performed post to ensure a probabilistic range between $[0, 1]$ [8]. *

$$s_i = \frac{s_i - min(S)}{max(S) - min(S)} \tag{1}$$

The final anomaly detection score for scan is inferred by combining the z patches together via Eq.2.

$$score = \frac{\sum_{i=1}^{z} X_i}{z} \tag{2}$$

$$x_i = \begin{cases} 1 & \text{if } abnormal \\ 0 & \text{otherwise} \end{cases}$$

* If the patch is above a certain threshold it is classified as anomalous (1) or normal (0). Threshold is based on the proportion of anomalous patches fed in during inference stage.

Set-up

The experiments were carried out on $3 * 3$ image patches of the same size from the original image. This is motivated by radiologists looking for multiple different symptoms in the x-ray that might be indicative of chest diseases.

- **Training:** The normal lung image patches are fed to the network during the training phase. Please note that only the healthy chest x-rays from Treat and Find data are using for the training stage.
- **Inference:** At this stage, both abnormal and normal patches are fed to the network. The patches reconstructed and anomaly score generated for each individual patch.
- **Output:** The patches are grouped together post the inference phase to reconstruct the original scan.
- **Code:** Model was trained in PyTorch (using ADAM [9] with an initial learning rate of 0.0002, and momentum of 0.5 and 0.999. The threshold for the anomaly score was set to 0.2.

Data-sets [1/2]

Training & Inference: UCLH's Treat and Find Data

UCLH developed a screening program to help combat the rise of TB detection in the UK. This initiative has led to an accumulated database of approx 93,000 digital chest x-rays. Within the dataset there are 92,000 healthy images and 811 TB lung images.

Inference: NIH Kaggle Data

Over 100,000 anonymized chest x-ray images and their corresponding data from National Institute of Health. These include 14 different chest abnormalities: ateectasis, cardiomegaly, consolidation, edema, effusion, emphysema, fibrosis, hernia, infiltration, mass, nodule, pleural thickening, pneumonia and pneumothorax.

Data-sets [2/2]

Inference: Pneumonia Data

Chest X-ray images (anterior-posterior) were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Childrens Medical Center, Guangzhou. All chest X-ray imaging was performed as part of patients routine clinical care.There are 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal)

Inference: Fibrosis Data

Fifty Chest X-rays which are of patients with fibrosis and have concurrent CT scans so information is known about the abnormality present.

Inference: Belarus Tuberculosis Data

Belarus TB is one of the leading scientific resources of annotated high resolution X-ray and Computer Tomography lung images of tuberculosis patients. There are 310 patient chest x-rays.

Bone Suppression for Chest X-rays [10]

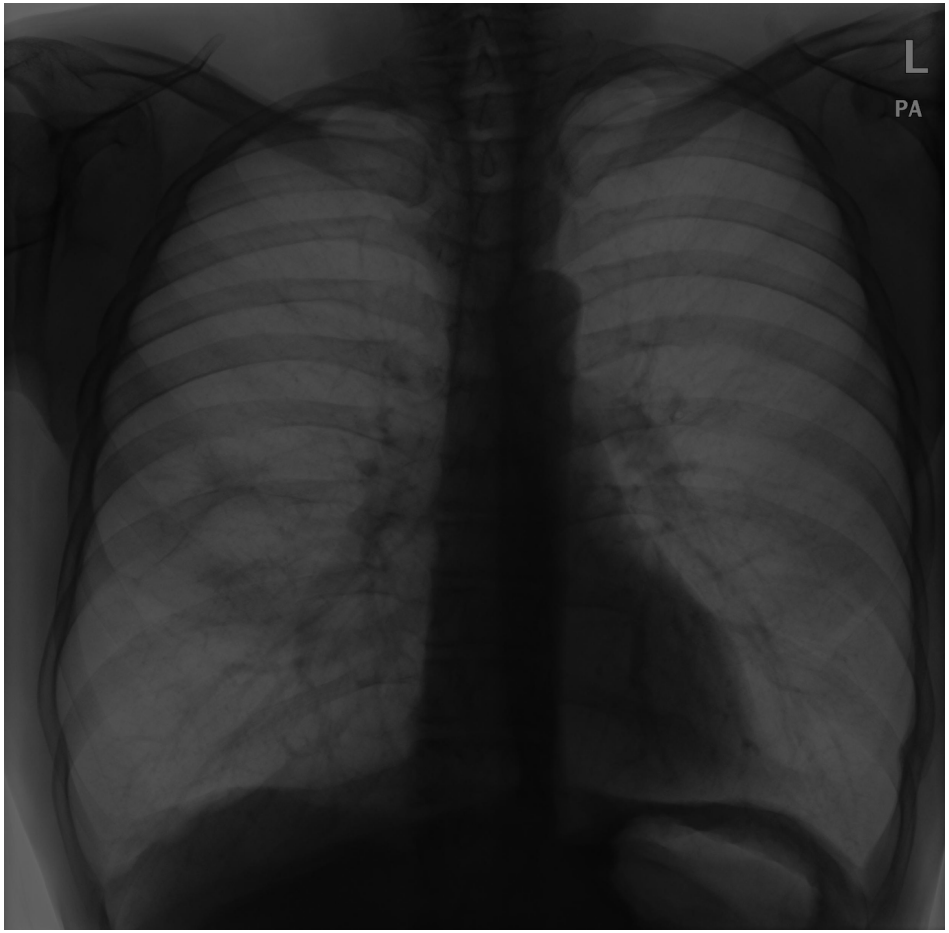


Figure: Original Scan

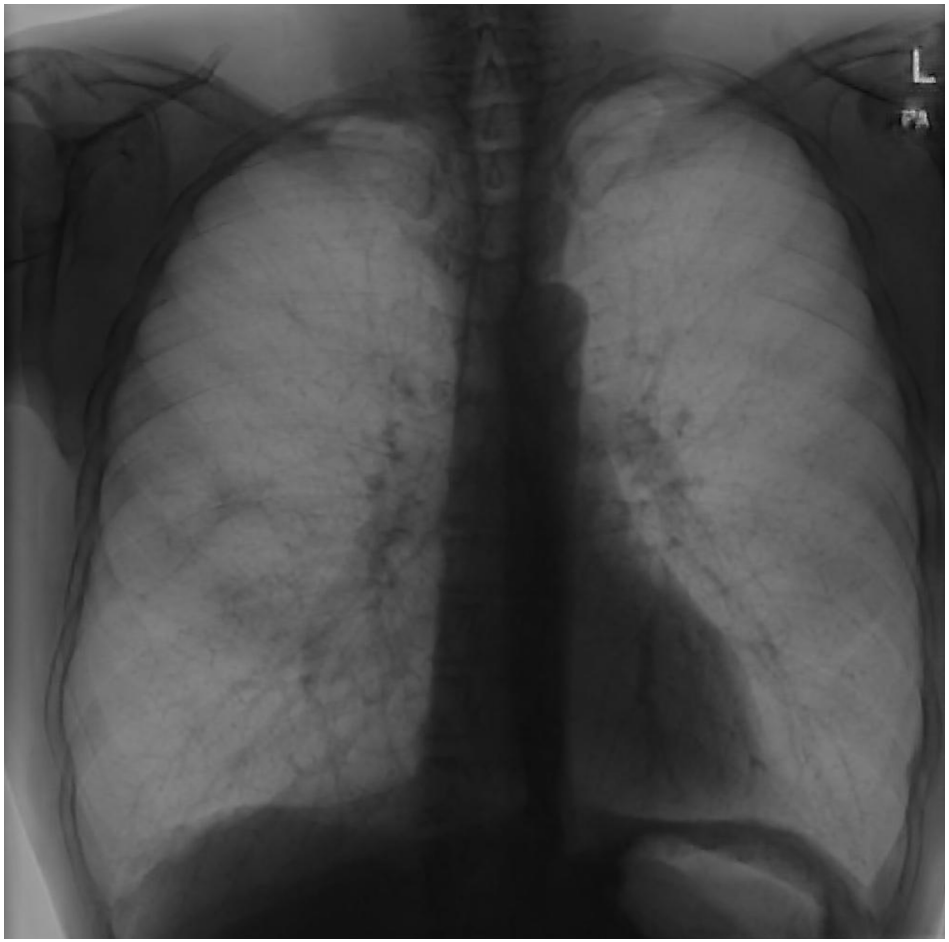


Figure: Bone Suppressed Scan

Slicing Chest X-rays

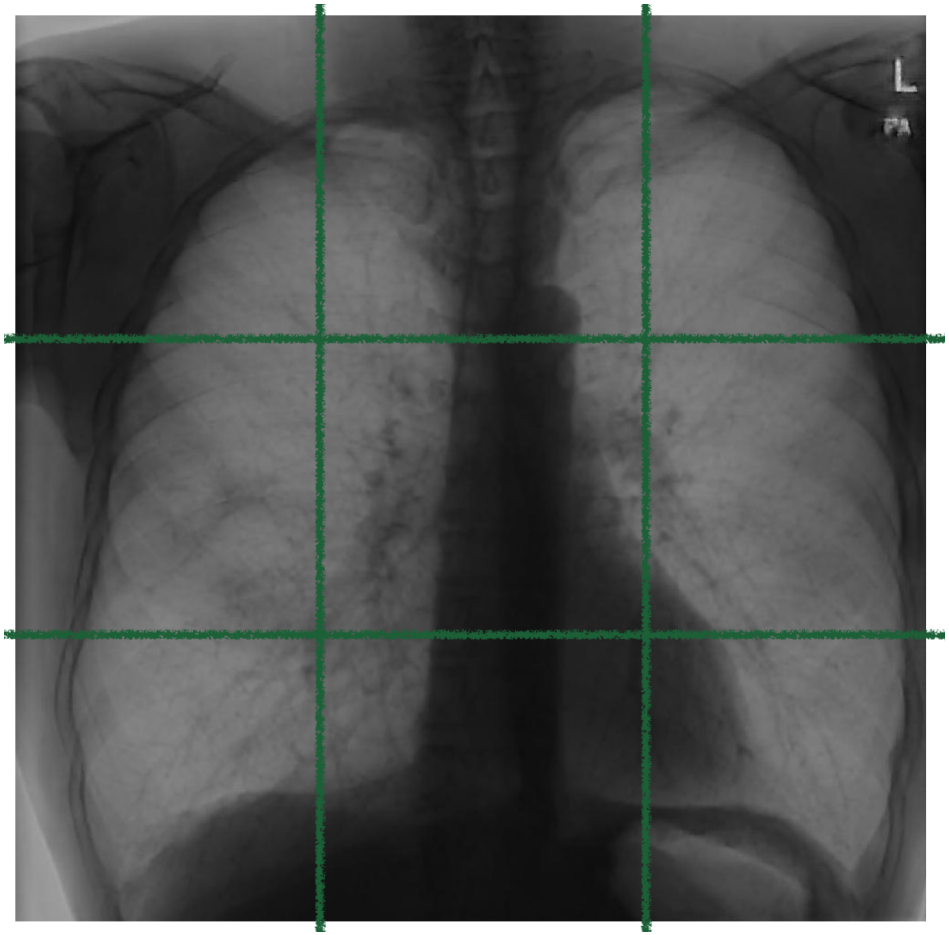


Figure: $3 * 3$ Scan Slice

Generated Images

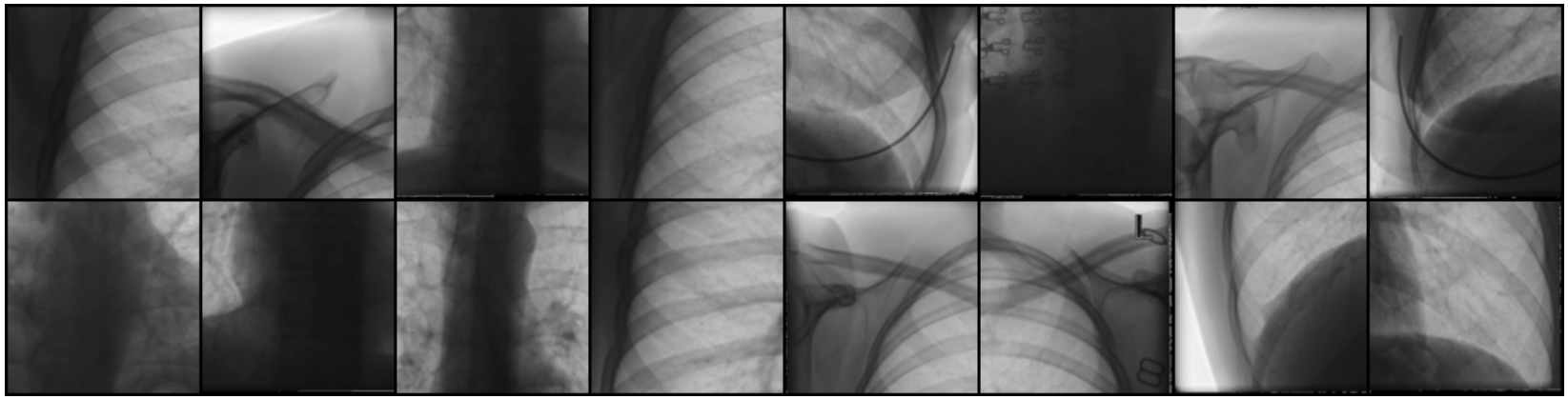


Figure: Example of sliced bone suppression images

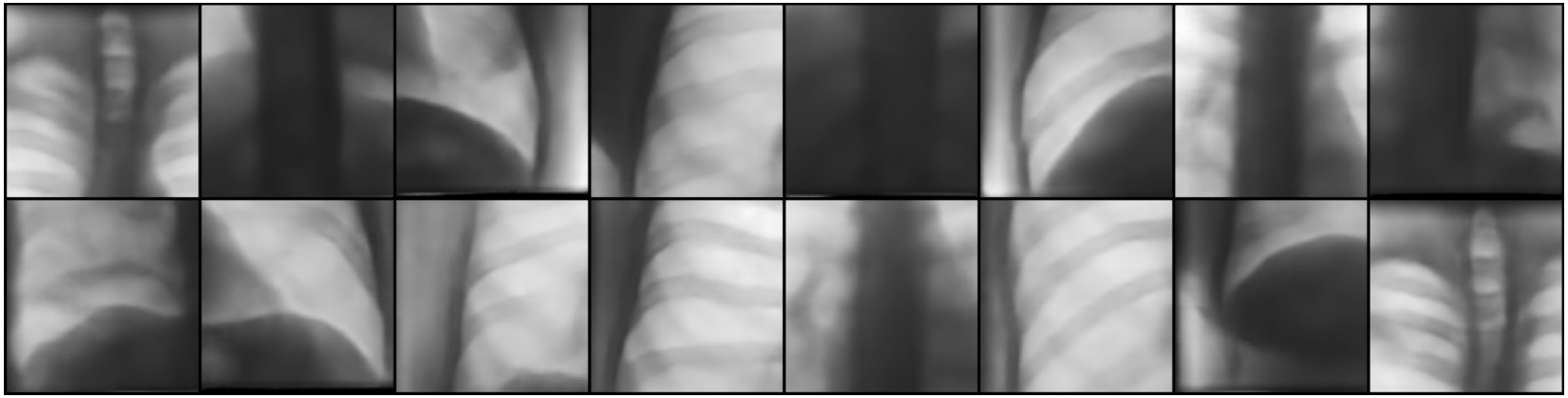


Figure: Example of generated chest x-ray slices

Results

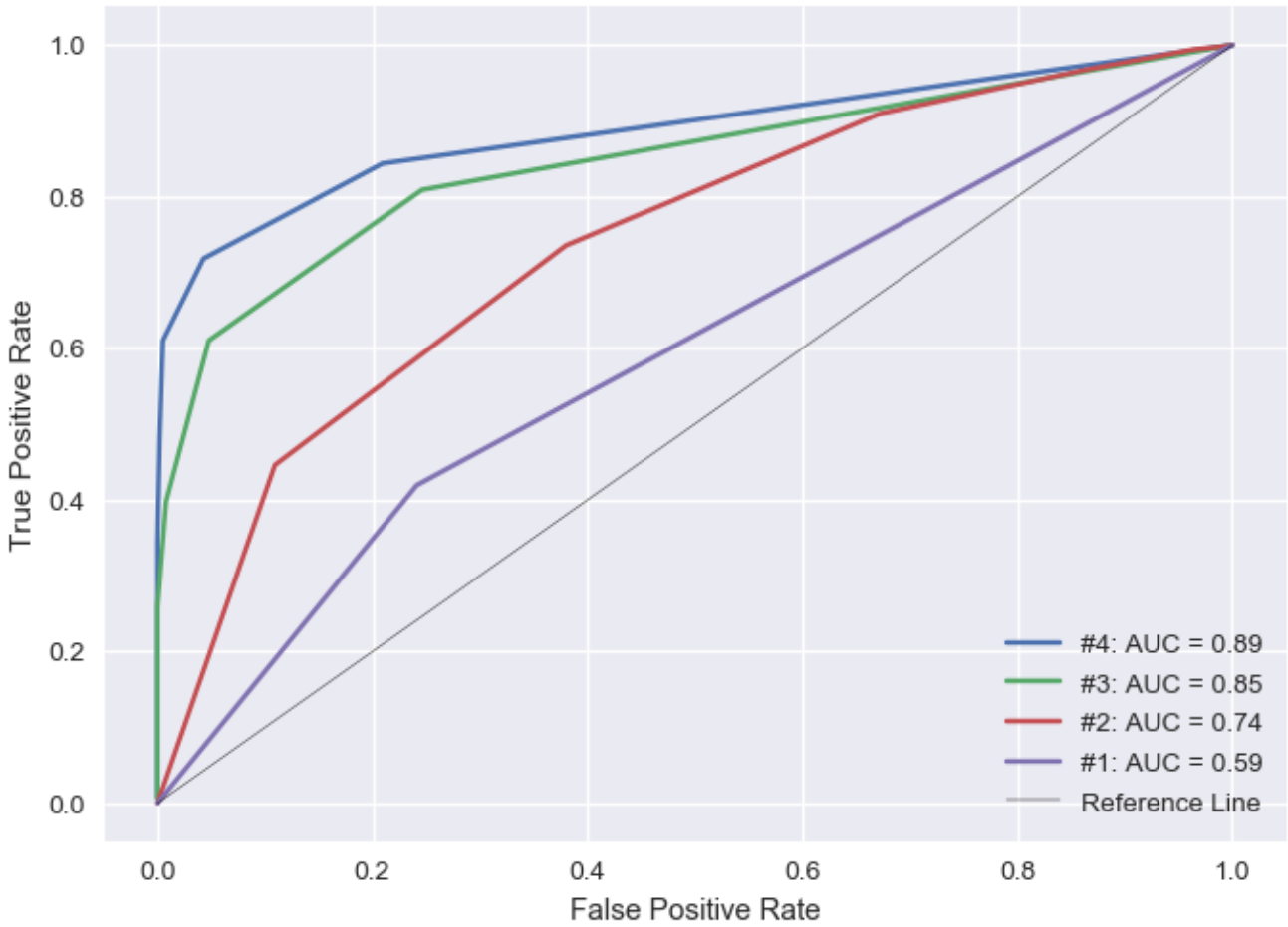


Figure: Results for the Treat and Find data

Results

#	Method	Data	Image Size	Batch-Size	Epochs	AUC
1	Proposed	TF	$32 * 32$	32	50	0.591
2	Proposed	TF	$64 * 64$	32	50	0.743
3	Proposed	TF	$128 * 128$	16	100	0.852
4	Proposed	TF	$256 * 256$	4	100	0.890
5	Proposed	Pneu	$256 * 256$	-	-	0.798
6	[11]	NIH	-	-	-	0.885
7	[12]	NIH	-	-	-	0.949

Table: Detailed breakdown of the results

NB: $256 * 256$ trained model will be used for inference on remaining data.

Results and Discussion

Current

- We have proposed a semi-supervised method that makes use of the learnt latent representation during the training phase to identify chest x-rays with *abnormalities*. These could be nodules, masses, etc.
- We are excited by the possibility that the technique might provide a new pathway towards creating a generalisable solution that is not constrained by the caveats associated with the data used for training.

Future Work

- Better understanding of what features the GAN is unable to generate could be misconstrued as anomalies
- Continue the experimentation on additional chest diseases
- Hyper-parameter optimisation to improve model performance

Citations I

Rahib Hidayat Abiyev and Mohammad Khaleel Sallam Ma'aitah.

Deep convolutional neural networks for chest diseases detection.
Journal of healthcare engineering, 2018.

Mohammad Tariqul Islam, Md Abdul Aowal, Ahmed Tahseen Minhaz, and Khalid Ashraf.

Abnormality detection and localization in chest x-rays using deep convolutional neural networks.
ArXiv, 2017.

Chunli Qin, Demin Yao, Yonghong Shi, and Zhijian Song.

Computer-aided detection in chest radiography based on artificial intelligence: a survey.
Biomedical Engineering, 2018.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio.

Ganomaly: Semi-supervised anomaly detection via adversarial training.
Advances in neural information processing systems, pages 2672–2680, 2014.

T. Schlegl, P. Seebock, S.M. Waldstein, U. Schmidt-Erfurth, and G. Langs.

Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics).
LNCS, pages 146–147, 2017.

H. Zenati, C.S. Foo, B. Lecouat, G. Manek, and V.R. Chandrasekhar.

Efficient gan based anomaly detection.
arXiv:1805.06222, 2018.

J. Donahue, P. Krahenbuhl, and T. Darrell.

Adversarial Feature Learning.
2017.

Citations II

S. Akcay, A. Atapour-Abarghouei, and T. P Breckon.

Ganomaly: Semi-supervised anomaly detection via adversarial training.
arXiv:1805.06725, 2018.

D. Kinga and J.B. Adam.

Adam: A method for stochastic optimization.
In: International Conference on Learning Representations (ICLR), 5, 2015.

Hunh Minh Chng.

Ml-bonesuppression.
`https://github.com/hmchuong/ML-BoneSuppression`, 2018.

S Jaeger, A Karargyris, S Candemir, L Folio, J Siegelman, F Callaghan, Xue Zhiyun, K Palaniappan, RK Singh, S Antani, G Thoma, Wang Yi-Xiang, Lu Pu-Xuan, and CJ McDonald.

Automatic tuberculosis screening using chest radiograph.
IEEE transactions on medical imaging, 33:233–245, 2014.

Mrinal Haloi, Raja K Rajalakshmi, and Pradeep W Artelus.

Towards radiologist-level accurate deep learning system for pulmonary screening.
International Journal of Computer Applications, 2018.