Deep Learning for Chest X-ray Analysis: A Survey

Ecem Sogancioglu*[a], Erdi Çallı*[a], Bram van Ginneken[a], Kicky G. van Leeuwen[a], Keelin Murphy[a]

[a]*Radboud University Medical Center, Institute for Health Sciences, Department of Medical Imaging, Nijmegen, The Netherlands*

Abstract

Recent advances in deep learning have led to a promising performance in many medical image analysis tasks. As the most commonly performed radiological exam, chest radiographs are a particularly important modality for which a variety of applications have been researched. The release of multiple, large, publicly available chest X-ray datasets in recent years has encouraged research interest and boosted the number of publications. In this paper, we review all studies using deep learning on chest radiographs, categorizing works by task: image-level prediction (classification and regression), segmentation, localization, image generation and domain adaptation. Commercially available applications are detailed, and a comprehensive discussion of the current state of the art and potential future directions are provided.

Keywords:
Deep Learning, chest radiograph, chest X-ray analysis, survey

1. Introduction

A cornerstone of radiological imaging for many decades, chest radiography (chest X-ray, CXR) remains the most commonly performed radiological exam in the world with industrialized countries reporting an average 238 erect-view chest X-ray images acquired per 1000 of population annually (United Nations, 2008). In 2006, it is estimated that 129 million CXR images were acquired in the United States alone (Mettler et al., 2009). The demand for, and availability of, CXR images may be attributed to their cost-effectiveness and low radiation dose, combined with a reasonable sensitivity to a wide variety of pathologies. The CXR is often the first imaging study acquired and remains central to screening, diagnosis, and management of a broad range of conditions (Raoof et al., 2012).

Chest X-rays may be divided into three principal types, according to the position and orientation of the patient relative to the X-ray source and detector panel: posteroanterior, anteroposterior, lateral. The posteroanterior (PA) and anteroposterior (AP) views are both considered as frontal, with the X-ray source positioned to the rear or front of the patient respectively. The AP image is typically acquired from patients in the supine position, while the patient is usually standing erect for the PA image acquisition. The lateral image is usually acquired in combination with a PA image, and projects the X-ray from one side of the patient to the other, typically from right to left. Examples of these image types are depicted in Figure 1.

The interpretation of the chest radiograph can be challenging due to the superimposition of anatomical structures along the projection direction. This effect can make it very difficult to detect abnormalities in particular locations (for example, a nodule posterior to the heart in a frontal CXR), to detect small or subtle abnormalities, or to accurately distinguish between different pathological patterns. For these reasons, radiologists typically show high inter-observer variability in their analysis of CXR images (Quekel et al., 2001; Balabanova et al., 2005; Young, 1994).

The volume of CXR images acquired, the complexity of their interpretation, and their value in clinical practice have long motivated researchers to build automated algorithms for CXR analysis. Indeed, this has been an area of research interest since the 1960s when the first papers describing an automated abnormality detection system on CXR images were published (Lodwick et al., 1963; Becker et al., 1964; Meyers et al., 1964; Kruger et al., 1972; Toriwaki et al., 1973). The potential gains from automated CXR analysis include increased sensitivity for subtle findings, prioritization of time-sensitive cases, automation of tedious daily tasks, and provision of analysis in situations where radiologists are not available (e.g., the developing world).

In recent years, deep learning has become the technique of choice for image analysis tasks and made a tremendous impact in the field of medical imaging (Litjens et al., 2017). Deep learning is notoriously data-hungry and the CXR research community has benefited from the publication of numerous large labeled databases in recent years, predominantly enabled by the generation of labels through automatic parsing of radiology reports. This trend began in 2017 with the release of 112,000 images from the NIH clinical center (Wang et al., 2017b). In 2019 alone, more than 755,000 images were released in 3 labelled databases (CheXpert (Irvin et al., 2019), MIMIC-CXR (Johnson et al., 2019), PadChest (Bustos et al., 2020)). In this work, we demonstrate the impact of these data releases on the number of deep learning publications in the field.

There have been previous reviews on the field of deep learning in medical image analysis (Litjens et al., 2017; van Gin-

⋆Ecem Sogancioglu and Erdi Çallı contributed equally.
*Corresponding author: E-mail: erdi.calli@radboudumc.nl

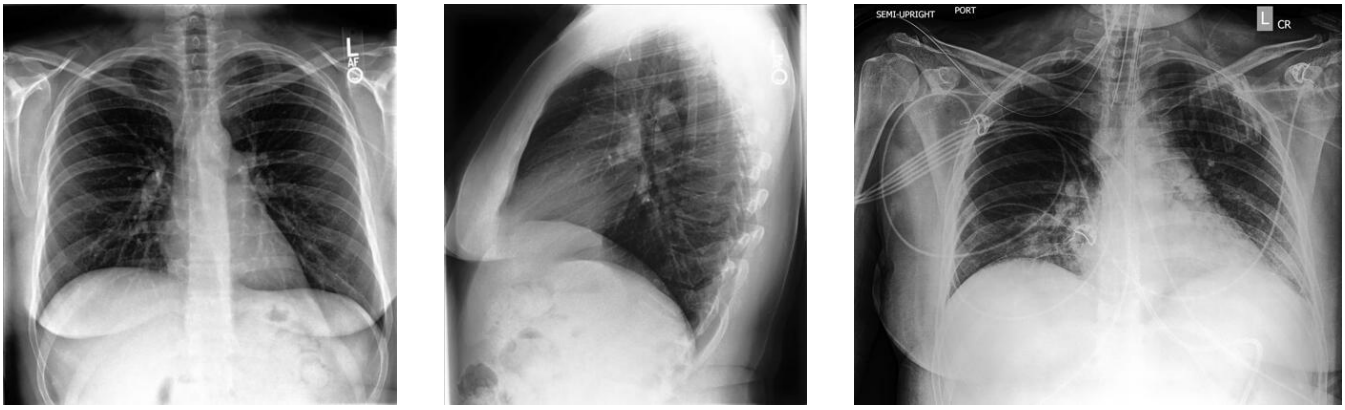

Figure 1: Left: posterior-anterior (PA) view frontal chest radiograph. Middle: lateral chest radiograph. Right: Anterior-posterior (AP) view chest radiograph. All three CXRs are taken from the CheXpert dataset (Irvin et al., 2019), patient 184.

neken, 2017; Sahiner et al., 2018; Feng et al., 2019) and on deep learning or computer-aided diagnosis for CXR (Qin et al., 2018; Kallianos et al., 2019; Anis et al., 2020). However, recent reviews of deep learning in chest radiography are far from exhaustive in terms of the literature and methodology surveyed, the description of the public datasets available, or the discussion of future potential and trends in the field. The literature review in this work includes 295 papers, published between 2015 and 2021, and categorized by application. A comprehensive list of public datasets is also provided, including numbers and types of images and labels as well as some discussion and caveats regarding various aspects of these datasets. Trends and gaps in the field are described, important contributions discussed, and potential future research directions identified. We additionally discuss the commercial software available for chest radiograph analysis and consider how research efforts can best be translated to the clinic.

The initial selection of literature to be included in this review was obtained as follows: A selection of papers was created using a PubMed search for papers with the following query.

```
chest and ("x-ray" or xray or radiograph) and
("deep learning" or cnn or "convolutional" or
"neural network")
```

A systematic search of the titles of conference proceedings from SPIE, MICCAI, ISBI, MIDL and EMBC was also performed, searching paper titles for the same search terms listed above. In the case of multiple publications of the same paper, only the latest publication was included. Relevant peer-reviewed articles suggested by co-authors and colleagues were added. The last search was performed on March 3rd, 2021.

This search strategy resulted in 767 listed papers. Of these, 61 were removed as they were duplicates of others in the list. A further 261 were excluded as their subject matter did not relate to deep learning for CXR, they were commentary or evaluation papers or they were not written in English. Publications that were not peer-reviewed were also excluded (8). Finally, during the review process 142 papers were excluded as the scientific content was considered unsound, as detailed further in Section 6, leaving 295 papers in the final literature review.

The remainder of this work is structured as follows: Section 2 provides a brief introduction to the concept of deep learning and the main network architectures encountered in the current literature. In Section 3, the public datasets available are described in detail, to provide context for the literature study. The review of the collected literature is provided in Section 4, categorized according to the major themes identified. Commercial systems available for chest radiograph analysis are described in Section 5. The paper concludes in Section 6, with a comprehensive discussion of the current state of the art for deep learning in CXR as well as the potential for future directions in both research and commercial environments.

2. Overview of Deep Learning Methods

This section provides an introduction to deep learning for image analysis, and particularly the network architectures most frequently encountered in the literature reviewed in this work. Formal definitions and more in-depth mathematical explanations of fully-connected and convolutional neural-networks are provided in many other works, including a recent review of deep learning in medical image analysis (Litjens et al., 2017). In this work, we provide only a brief overview of these fundamental details and refer the interested reader to previous literature.

Deep learning is a branch of machine learning, which is a general term describing learning algorithms. The algorithm underpinning all deep learning methods is the neural network, in this case, constructed with many hidden layers (‘deep’). These networks may be constructed in many ways with different types of layers included and the overall construction of a network is referred to as its ‘architecture’. Sections 2.3 to 2.6 describe commonly used architectures categorized by types of application in the CXR literature.

2.1. Convolutional Neural Networks

In the 1980s, networks using convolutional layers were first introduced for image analysis (Fukushima and Miyake, 1982), and the idea was formalized over the following years (LeCun and Bengio, 1998). These convolutional layers now form the

basis for all deep learning image analysis tasks, almost without exception. Convolutional layers use neurons that connect only to a small 'receptive field' from the previous layer. These neurons are applied to different regions of the previous layer, operating as a sliding window over all regions, and effectively detecting the same local pattern in each location. In this way, spatial information is preserved and the learned weights are shared.

2.2. *Transfer Learning*

Transfer learning investigates how to transfer knowledge extracted from one domain (source domain) to another (target) domain. One of the most commonly used transfer learning approaches in CXR analysis is the use of pre-training.

With the pre-training approach, the network architecture is first trained on a large dataset for a different task, and the trained weights are then used as an initialization for the subsequent task for fine-tuning (Yosinski et al., 2014). Depending on data availability from the target domain, all layers can be re-trained, or only the final (fully connected) layer can be re-trained. This approach allows neural networks to be trained for new tasks using relatively smaller datasets since useful low-level features are learned from the source domain data. It has been shown that pre-training on the ImageNet dataset (for classification of natural images) (Baltruschat et al., 2019b) is beneficial for chest radiography analysis and this type of transfer learning is prominently used in the research surveyed in this work. ImageNet pre-trained versions of many architectures are publicly available as part of popular deep learning frameworks. The pre-trained architectures may also be used as feature extractors, in combination with more traditional methods, such as support vector machines or random forests. Domain adaptation is another subfield of transfer learning and is discussed thoroughly in Section 2.7.

2.3. *Image-level Prediction Networks*

In this work we use the term 'image-level prediction' to refer to tasks where prediction of a category label (classification) or continuous value (regression) is implemented by analysis of an entire CXR image. These methods are distinct from those which make predictions regarding small patches or segmented regions of an image. Classification and regression tasks are grouped together in this work since they typically use the same types of architecture, differing only in the final output layer. One of the early successful deep convolutional architectures for image-level prediction was AlexNet (Krizhevsky et al., 2012), which consists of 5 convolutional layers followed by 3 fully connected layers. AlexNet became extremely influential in the literature when it beat all other competitors in the ILSVRC (ImageNet) challenge (Deng et al., 2009) by a large margin in 2012. Since then many deep convolutional neural network architectures have been proposed. The VGG family of models (Simonyan and Zisserman, 2014) use 8 to 19 convolutional layers followed by 3 fully-connected layers. The Inception architecture was first introduced in 2015 (Szegedy et al., 2015) using multiple convolutional filter sizes within layered blocks known as Inception modules. In 2016, the ResNet family of models (He et al., 2016) began to gain popularity and improve upon previous benchmarks. These models define residual blocks consisting of multiple convolution operations, with skip connections which typically improve model performance. After the success of ResNet, skip connections were widely adopted in many architectures. DenseNet models (Huang et al., 2017), introduced in 2017, also use skip connections between blocks, but connect all layers to each other within blocks. A later version of the Inception architecture also added skip connections (Inception-Resnet) (Szegedy et al., 2017). The Xception network architecture (Chollet, 2017) builds upon the Inception architecture but separates the convolutions performed in the 2D image space from those performed across channels. This was demonstrated to improve performance compared to Inception V3.

The majority of works surveyed in this review use one or more of the model architectures discussed here with varying numbers of hidden layers.

2.4. *Segmentation Networks*

Segmentation is a task where pixels are assigned a category label, and can also be considered as a pixel classification. In natural image analysis, this task is often referred to as 'semantic segmentation' and frequently requires every pixel in the image to have a specified category. In the medical imaging domain these labels typically correspond to anatomical features (e.g., heart, lungs, ribs), abnormalities (e.g., tumor, opacity) or foreign objects (e.g., tubes, catheters). It is typical in the medical imaging literature to segment just one object of interest, essentially assigning the category 'other' to all remaining pixels.

Early approaches to segmentation using deep learning used standard convolutional architectures designed for classification tasks (Chen et al., 2018b). These were employed to classify each pixel in a patch using a sliding window approach. The main drawback to this approach is that neighboring patches have huge overlap in pixels, resulting in inefficiency caused by repeating the same convolutions many times. It additionally treats each pixel separately which results in the method being computationally expensive and only applicable to small images or patches from an image.

To address these drawbacks, fully convolutional networks (FCNs) were proposed, replacing fully connected layers with convolutional layers (Shelhamer et al., 2017). This results in a network which can take larger images as input and produces a likelihood map output instead of an output for a single pixel. In 2015, a fully convolutional architecture known as the U-Net was proposed (Ronneberger et al., 2015) and this work has become the most cited paper in the history of medical image analysis. The U-Net consists of several convolutional layers in a contracting (downsampling) path, followed by further convolutional layers in an expanding (upsampling) path which restores the result to the input resolution. It additionally uses skip connections between the same levels on the contracting and expanding paths to recover fine details that were lost during the pooling operation. The majority of image segmentation works in this review employ a variant of the FCN or the U-Net.

2.5. Localization Networks

This survey uses the term localization to refer to identification of a specific region within the image, typically indicated by a bounding box, or by a point location. As with the segmentation task, localization, in the medical domain, can be used to identify anatomical regions, abnormalities, or foreign object structures. There are relatively few papers in the CXR literature reviewed here that deal specifically with a localization method, however, since it is an important task in medical imaging, and may be easier to achieve than a precise segmentation, we categorize these works together.

In 2014, the RCNN (Region Convolutional Neural Network) was introduced (Girshick et al., 2014), identifying regions of interest in the image and using a CNN architecture to extract features of these regions. A support vector machine (SVM) was used to classify the regions based on the extracted features. This method involves several stages and is relatively slow. It was later superseded by fast-RCNN (Girshick, 2015) and subsequently by faster-RCNN (Ren et al., 2017) which streamlined the processing pipeline, removing the need for initial region identification or SVM classification, and improving both speed and performance. In 2017, a further extension was added to faster-RCNN to additionally enable a precise segmentation of the item identified within the bounding box. This method is referred to as Mask R-CNN (He et al., 2017). While this is technically a segmentation network, we mention it here as part of the RCNN family. Another architecture which has been popular in object localization is YOLO (You Only Look Once), first introduced in 2016 (Redmon et al., 2016) as a single-stage object detection method, and improved in subsequent versions in 2017 and 2018 (Redmon and Farhadi, 2017, 2018). The original YOLO architecture, using a single CNN and an image-grid to specify outputs was significantly faster than its contemporaries but not quite as accurate. The improved versions leveraged both classification and detection training data and introduced a number of training improvements to achieve state of the art performance while remaining faster than its competitors. A final localization network that features in medical imaging literature is RetinaNet (Lin et al., 2017). Like YOLO, this is a single stage detector, which introduces the concept of a focal loss function, forcing the network to concentrate on more difficult examples during training. Most of the localization works included in this review use one of the architectures described above.

2.6. Image Generation Networks

One of the tasks deep learning has been commonly used for is the generation of new, realistic images, based on information learned from a training set. There are numerous reasons to generate images in the medical domain, including generation of more easily interpretable images (by increasing resolution, or removal of projected structures impeding analysis), generation of new images for training (data augmentation), or conversion of images to emulate appearances from a different domain (domain adaptation). Various generative schemes have also been used to improve the performance of tasks such as abnormality detection and segmentation.

Image generation was first popularized with the introduction of the generative adversarial network (GAN) in 2014 (Goodfellow et al., 2014). The GAN consists of two network architectures, an image generator, and a discriminator which attempts to differentiate generated images from real ones. These two networks are trained in an adversarial scheme, where the generator attempts to fool the discriminator by learning to generate the most realistic images possible while the discriminator reacts by progressively learning an improved differentiation between real and generated images.

The training process for GANs can be unstable with no guarantee of convergence, and numerous researchers have investigated stabilization and improvements of the basic method (Salimans et al., 2016; Heusel et al., 2017; Karras et al., 2018; Arjovsky et al., 2017). GANs have also been adapted to conditional data generation (Chen et al., 2016; Odena et al., 2017) by incorporating class labels, image-to-image translation (conditioned on an image in this case) (Isola et al., 2017), and unpaired image-to-image translation (CycleGAN Zhu et al. (2017)).

GANs have received a lot of attention in the medical imaging community and several papers were published for medical image analysis applications in recent years (Yi et al., 2019b). Many of the image generation works identified in this review employed GAN based architectures.

2.7. Domain Adaptation Networks

In this work we use the term ‘Domain Adaptation’, which is a subfield of transfer learning, to cover methods attempting to solve the issue that architectures trained on data from a single ‘domain’ typically perform poorly when tested on data from other domains. The term ‘domain’ is weakly defined; In medical imaging it may suggest data from a specific hardware (scanner), set of acquisition parameters, reconstruction method or hospital. It could, less frequently, also refer to characteristics of the population included, for example the gender, ethnicity, age or even strain of some pathology included in the dataset.

Domain adaptation methods consider a network trained for an image analysis task on data from one domain (the source domain), and how to perform this analysis accurately on a different domain (the target domain). These methods can be categorized as supervised, unsupervised, and semi-supervised depending on the availability of labels from the target domain and they have been investigated for a variety of CXR applications from organ segmentation to multi-label abnormality classification. There is no specific architecture that is typical for domain adaptation, but rather architectures are combined in various ways to achieve the goal of learning to analyze images from unseen domains. The approaches to this problem can be broadly divided into three classes (following the categorization of (Wang and Deng, 2018)); discrepancy-based, reconstruction-based and adversarial-based.

Discrepancy-based approaches aim to induce alignment between the source and target domain in some feature space by fine-tuning the image analysis network and optimizing a measurement of discrepancy between the two domains. Reconstruction-based approaches, on the other hand, use an auxiliary encoder-decoder reconstruction network that aims to

learn domain invariant representation through a shared encoder. Adversarial-based approaches are based on the concept of adversarial training from GANs, and use a discriminator network which tries to distinguish between samples from the source and target domains, to encourage the use of domain-invariant features. This category of approaches is the most commonly used in CXR analysis for domain adaptation, and consists of generative and non-generative models. Generative models transform source images to resemble target images by operating directly on pixel space whereas non-generative models use the labels on the source domain and leverage adversarial training to obtain domain invariant representations.

3. Datasets

Deep learning relies on large amounts of annotated data. The digitization of radiological workflows enables medical institutions to collate and categorize large sets of digital images. In addition, advances in natural language processing (NLP) algorithms mean that radiological reports can now be automatically analyzed to extract labels of interest for each image. These factors have enabled the construction and release of multiple large labelled CXR datasets in recent years. Other labelling strategies have included the attachment of the entire radiology report and/or labels generated in other ways, such as radiological review of the image, radiological review of the report, or laboratory test results. Some datasets include segmentations of specified structures or localization information.

In this section we detail each public dataset that is encountered in the literature included in this review as well as any others available to the best of our knowledge. Details are provided in Table 1. Each dataset is given an acronym which is used in the literature review tables (Tables 2 to 7) to indicate that the dataset was used in the specified work.

1. ChestX-ray14 (C) is a dataset consisting of 112,120 CXRs from 30,805 patients (Wang et al., 2017b). The CXRs are collected at the (US) National Institute of Health. The images are distributed as 8-bit grayscale images scaled to 1024×1024 pixels. The dataset was automatically labeled from radiology reports, indicating the existence of 14 types of abnormality.
2. CheXpert (X) is a dataset consisting of 224,316 CXRs from 65,240 patients (Irvin et al., 2019). The CXRs are collected at Stanford Hospital between October 2002 and July 2017. The images are distributed as 8-bit grayscale images with original resolution. The dataset was automatically labeled from radiology reports using a rule-based labeler, indicating the presence, absence, uncertainty, and no-mention of 12 abnormalities, no findings, and the existence of support devices.
3. MIMIC-CXR (M) is a dataset consisting of 371,920 CXRs from and 64,588 patients (Johnson et al., 2019). The CXRs are collected from patients admitted to the emergency department of Beth Israel Deaconess Medical Center between 2011 and 2016. In version 1 (V1) the images are distributed as 8-bit grayscale images in full resolution. The dataset was automatically labeled from radiology reports using the same rule-based labeler system (described above) as CheXpert. A second version (V2) of MIMIC-CXR was later released including the anonymized radiology reports and DICOM files.
4. PadChest (P) is a dataset consisting of 160,868 CXRs from 109,931 studies and 67,000 patients (Bustos et al., 2020). The CXRs are collected at San Juan Hospital (Spain) from 2009 to 2017. The images are stored as 16-bit grayscale images with full resolution. 27,593 of the reports were manually labeled by physicians. Using these labels, an RNN was trained and used to label the rest of the dataset from the reports. The reports were used to extract 174 findings, 19 diagnoses, and 104 anatomic locations. The labels conform to a hierarchical taxonomy based on the standard Unified Medical Language System (UMLS) (Bodenreider, 2004).
5. PLCO (PL) is a screening trial for prostate, lung, colorectal and ovarian (PLCO) cancer (Zhu et al., 2013). The lung arm of this study has 185,421 CXRs from 56,071 patients. The NIH distributes a standard set of 25,000 patients and 88,847 frontal CXRs. This dataset contains 22 disease labels with 4 abnormality levels and the locations of the abnormalities.
6. Open-i (O) is a dataset consisting of 7,910 CXRs from 3,955 studies and 3,955 patients (Demner-Fushman et al., 2012). The CXRs are collected from the Indiana Network for Patient Care (McDonald et al., 2005). The images are distributed as anonymized DICOMs. The radiological findings obtained by radiologist interpretation are available in MeSH format[1].
7. Ped-Pneumonia (PP) is a dataset consisting of 5,856 pediatric CXRs (Kermany, 2018). The CXRs are collected from Guangzhou Women and Children's Medical Center, Guangzhou, China. The images are distributed in 8-bit grayscale images scaled in various resolutions. The labels include bacterial and viral pneumonia as well as normal.
8. JSRT dataset (J) consists of 247 images with a resolution of 2048×2048, 0.175mm pixel-size and 12-bit depth (Shiraishi et al., 2000). It includes nodule locations (on 154 images) and diagnosis (malignant or benign). The reference standard for heart and lung segmentations of these images are provided by the SCR dataset (van Ginneken et al., 2006) and we group these datasets together in this work.
9. RSNA-Pneumonia (RP) is a dataset consisting of 30,000 CXRs with pneumonia annotations (RSNA, 2018). These images are acquired from ChestX-ray14 and are 8-bit grayscale with 1024×1024 resolution. Annotations are added by radiologists using bounding boxes around lung opacities and 3 classes indicating normal, lung opacity, not normal.
10. Shenzhen (S) is a dataset consisting of 662 CXRs (Jaeger et al., 2014). The CXRs are collected at Shenzhen No.3

[1] https://www.nlm.nih.gov/mesh/meshhome.html

Table 1: CXR datasets available for research. Values above 10,000 are rounded and shortened using K, indicating thousand (such as 10K for 10,000).
Labeling Methods: RP=Report Parsing, RIR=Radiologist Interpretation of Reports, RI=Radiologist Interpretation of Chest X-Rays, RCI=Radiologist Cohort agreement on Chest X-Rays, LT=Laboratory Tests.
Annotation Types: BB=Bounding Box, CL=Classification, CLoc=Classification with Location label, R=Report, SE=Segmentation.
Gold Standard Data: This refers to the number of images labeled by methods other than Report Parsing

	Patients (P) Studies (S) Images (I)	View Positions	Annotation Types	 Labels	 Studies	Image Format	Labeling method	Gold Standard Data
ChestX-ray14 (C) Wang et al. (2017b)	P: 31K I: 112K	PA: 67K AP: 45K	CL BB	14 8	112K 983	PNG	RP RI	 984
CheXpert (X) Irvin et al. (2019)	P: 65K S: 188K I: 224K	PA: 29K AP: 162K LL: 32K	CL	14	224K	JPEG	RCI RP	235
MIMIC-CXR (M) Johnson et al. (2019)	P: 65K S: 224K I: 372K	PA+AP: 250K LL: 122K	CL (V1) R (V2)	14	372K 372K	JPEG (V1) DICOM (V2)	RP	
PadChest (P) Bustos et al. (2020)	P: 67K S: 110K I: 160K	PA: 96K AP: 20K LL: 51K	CL R	193	110K 110K	DICOM	RIR RP	27593
PLCO (PL) Zhu et al. (2013)	P: 25K I: 89K	PA: 89K	CL CLoc	22 17	89K 89K	TIFF	RI	All
Open-i (O) Demner-Fushman et al. (2012)	P: 3,955 I: 7,910	PA: 3,955 LL: 3,955	R		3,955	DICOM	RI	All
Ped-pneumonia (PP) Kermany (2018)	I: 5,856		CL	2	5,856	JPEG	RI	All
JSRT+SCR (J) Shiraishi et al. (2000)	I: 247	PA: 247	SE	3	247	DICOM	RI	All
RSNA-Pneumonia (RP) RSNA (2018)	I: 30K	PA: 16K AP:14K	BB CL	1	30K	DICOM	RI	All
Shenzhen (S) Jaeger et al. (2014)	I: 340	PA: 340	CL	2	340 340	DICOM	RI	All
Montgomery (MO) Jaeger et al. (2014)	I: 138	PA: 138	CL SE	2	138 138	PNG	RI	All
BIMCV (B) Vayá et al. (2020)	P: 1,305 S: 2,391 I: 3,293	PA: 1,132 AP: 1,346 LL: 815	CL	1	2,391	DICOM	LT	All
COVIDDSL (CD) HMHospitales	P: 1,725 S: 4,943	PA AP (most) LL	CL	1	4,943	DICOM	LT	All
COVIDGR (CG) Tabik et al. (2020)	I: 852	PA:852	CL	2	852	JPEG	RI	All
SIIM-ACR (SI) ACR (2019)	I: 16K P:16K	PA: 11K AP: 4,799	SE	1	16K	DICOM	RI	All
CXR14-Rad-Labels (CR) Majkowska et al. (2019)	P: 1,709 I: 4,374	AP: 3,244 PA: 1,132	CL	4	4,374	PNG	RCI	All
COVID-CXR (CC) Cohen et al. (2020c)	I:866 P:449	PA:344 AP:438 LL:84	CL BB SE			PNG+JPEG	Various	
NLST (N) National Lung Screening Trial Research Team et al. (2011)	I: 5493	No public information available Number of images is reported by Lu et al. (2019).						
Object-CXR (OB)	I:10K	No longer at original download location						
Belarus (BL)	I: 300	No longer at original download location						

Hospital in Shenzhen, Guangdong providence, China in September 2012. The images, including some pediatric images, are distributed as 8-bit grayscale with full resolution and are annotated for signs of tuberculosis.

11. Montgomery (MO) is a dataset consisting of 138 CXRs (Jaeger et al., 2014). The CXRs are collected by the tuberculosis control program of the Department of Health and Human Services of Montgomery County, MD, USA. The images are distributed as anonymized DICOMs, annotated for signs of tuberculosis and additionally include lung segmentation masks.

12. BIMCV (B) is a COVID-19 dataset released by the Valencian Region Medical ImageBank (BIMCV) in 2020 (Vayá et al., 2020). It includes CXR images as well as CT scans and laboratory test results. The dataset includes 3,293 CXRs from 1,305 COVID-19 positive subjects. CXR images are 16-bit PNG format with original resolution.

13. COVIDDSL (CD) is a COVID-19 dataset released by the HM Hospitales group in Spain (HMHospitales). It includes CXR images for 1,725 patients as well as detailed results from laboratory testing, vital signs etc. All subjects are stated to be confirmed COVID-19 positive.

14. COVIDGR (CG) is a dataset consisting of 852 PA CXR images where half of them are labeled as COVID-19 positive based on corresponding RT-PCR results obtained within at most 24 hours (Tabik et al., 2020). This dataset was collected from Hospital Universitario Clínico San Cecilio, Granada, Spain, and the level of severity of positive cases is provided.
15. SIIM-ACR (SI) This dataset was released for a Kaggle challenge on pneumothorax detection and segmentation (ACR, 2019). Researchers have determined that at least some (possibly all) of the images are from the ChestX-ray14 dataset although the challenge organizers have not confirmed the data sources. They are supplied in 1024 × 1024 resolution as DICOM files. Pixel segmentations of the pneumothorax in positive cases are provided.
16. CXR14-Rad-Labels (CR) supplies additional annotations for a subset of ChestX-ray14 data (Majkowska et al., 2019). It consists of 4 labels for 4,374 studies and 1,709 patients. These labels are collected by the adjudicated agreement of 3 radiologists. These radiologists were selected from a cohort of 11 radiologists for the validation split (2,412 studies from 835 patients), and 13 radiologists for the test split (1,962 studies from 860 patients). The individual labels from each radiologist as well as the agreement labels were provided.
17. COVID-CXR (CV) is a dataset consisting of 930 CXRs at the time of writing (the dataset remains in continuous development) (Cohen et al., 2020c). The CXRs are collected from a large variety of locations using different methods including screenshots from papers researching COVID-19. Available labels vary accordingly, depending on what information is available from the source where the image was obtained. Images do not have a standard resolution and are published as 8-bit PNG or JPEG files.
18. NLST (N) is a dataset of publicly available CXRs collected during the NLST screening trial National Lung Screening Trial Research Team et al. (2011). This trial aimed to compare the use of low-dose computed tomography (CT) with CXRs for lung cancer screening in smokers. The study had 26,732 participants in the CXR arm and a part of this data is available upon request.
19. Object-CXR (OB) is a dataset of 10,000 CXR images from hospitals in China with foreign objects annotated on the images. The download location (https://jfhealthcare.github.io/object-CXR/) is no longer available at the time of writing. Further detail is not provided since it cannot be verified from the image source.
20. Belarus (BL) This dataset is included since it is used in a number of reviewed papers however the download location (http://tuberculosis.by) is no longer available at the time of writing. The dataset consisted of approximately 300 frontal chest X-rays with confirmed TB. Further detail is not provided since it can no longer be verified from the image source.

The rapid increase in the number of publicly available CXR images in recent years has positively impacted the number of deep learning studies published in the field. Figure 2 illustrates the cumulative number of publicly available CXR images and the number of publications on deep learning with CXR per year.

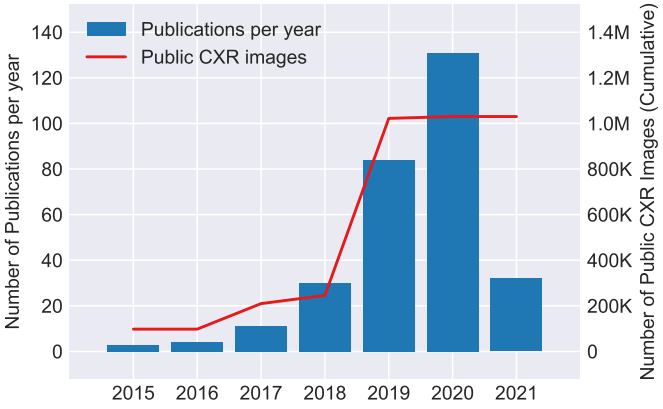

Figure 2: Number of publications that were reviewed in this work, by year, compared with the number of publicly available CXR images. Data for 2021 is until March 3rd of that year.

3.1. Public Dataset Caution

Publication of medical image data is extremely important for the research community in terms of advancing the state of the art in deep learning applications. However, there are a number of caveats that should be considered and understood when using the public datasets described in this work. Firstly, many datasets make use of Natural Language Processing (NLP) to create labels for each image. Although this is a fast and inexpensive method of labeling, it is well known that there are inaccuracies in labels acquired this way (Irvin et al., 2019; Oakden-Rayner, 2020, 2019). There are a number of causes for such inaccuracies. Firstly, some visible abnormalities may not be mentioned in the radiology report, depending on the context in which it was acquired (Olatunji et al., 2019). Further, the NLP algorithm can be erroneous in itself, interpreting negative statements as positive, failing to identify acronyms, etc. Finally, many findings on CXR are subtle or doubtful, leading to disagreements even among expert observers (Olatunji et al., 2019). Acknowledging some of these issues, Irvin et al. (2019) includes labels for uncertainty or no-mention in the labels on the CheXpert dataset. One particular cause for concern with NLP labels is the issue of systematic or structured mislabeling, where an abnormality is consistently labeled incorrectly in the same way. An example of this occurs in the ChestX-ray14 dataset where subcutaneous emphysema is frequently identified as (pulmonary) 'emphysema' (Calli et al., 2019; Oakden-Rayner, 2020).

It has been demonstrated that deep neural networks can tolerate reasonable levels of label inaccuracy in the training set without a significant effect on model performance (Calli et al., 2019; Rolnick et al., 2018). Although such labels can be used for training, for an accurate evaluation and comparison of models it is desirable that the test dataset is accurately labelled. In

the literature reviewed in this work, many authors rely on labels from NLP algorithms in their test data, while others use radiologist annotations, laboratory tests and/or CT verification for improved test set labelling. We refer to data that uses these improved labelling techniques as gold standard data (Table 1).

The labels defined in the public datasets should also be considered carefully and understood by the researchers using them. Many labels have substantial dependencies between them. For example, some datasets supply labels for both 'consolidation' and 'pneumonia'. Consolidation (blocked airspace) is an indicator of a patient with pneumonia, suggesting there will be significant overlap between these labels. A further point for consideration is that, in practice, not all labels can be predicted by a CXR image alone. Pneumonia is rarely diagnosed by imaging alone, requiring other clinical signs or symptoms to suggest that this is the cause for a visible consolidation.

Many public datasets release images with a lower quality than is used for radiological reading in the clinic. This may be a cause for decreased performance in deep learning systems, particularly for more subtle abnormalities. The reduction in quality is usually related to a decrease in image size or bit-depth prior to release. This is typically carried out to decrease the overall download size of a dataset. However, in some cases, CXR data has been collected by acquiring screenshots from online literature, which results in an unquantifiable degradation of the data. In the clinical workflow, DICOM files are the industry standard for storing CXRs, typically using 12 bits per pixel and with image dimensions of approximately 2 to 4 thousand pixels in each of the X and Y directions. In the event that the data is post-processed before release it would be desirable that a precise description of all steps is provided to enable researchers to reproduce them for dataset combination.

4. Deep Learning for Chest Radiography

In this section we survey the literature on deep learning for chest radiography, dividing it into sections according to the type of task that is addressed (Image-level Prediction, Segmentation, Image Generation, Domain Adaptation, Localization, Other). For each of these sections a table detailing the literature on that task is provided. Some works which have equal main focus on two tasks may appear in both tables. For Segmentation and Localization, only studies that quantitatively evaluate their results are included in those categories. Figure 3 shows the number of studies for each of the tasks.

4.1. Image-level Prediction

Image-level prediction refers to the task of predicting a label (classification) or a continuous value (regression) by analyzing an entire image. Classification labels may relate to pathology (e.g. pneumonia, emphysema), information such as the subject gender, or orientation of the image. Regression values might, for example, indicate a severity score for a particular pathology, or other information such as the age of the subject.

We classified 187 studies, fully detailed in Table 2 as image-level predictions. Most of these studies make use of off-the-shelf deep learning models to predict a pathology, metadata information or a set of labels provided with a dataset. The number of studies for each label are provided in Figure 4.

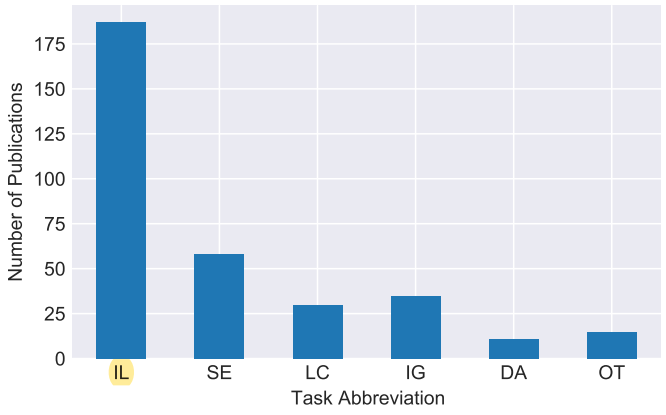

Figure 3: Number of publications reviewed for each task. 295 studies are included, each study may perform at most two tasks.
Tasks: IL=Image-level Predictions, SE=Segmentation, LC=Localization, IG=Image Generation, DA=Domain Adaptation, OT=Other.

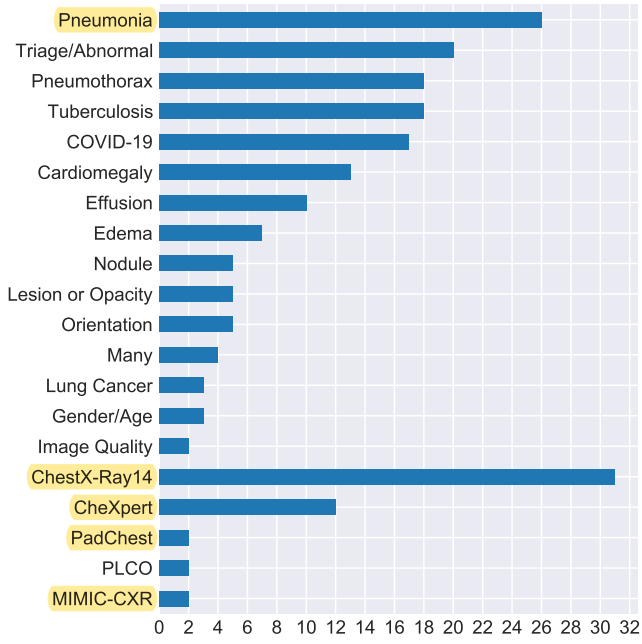

Figure 4: Number of studies for the Image-level Prediction labels. The studies that specifically work on a dataset and its labels are grouped together at the bottom. 187 papers are included, each may study more than one label.

The most commonly studied image-level prediction task is predicting the labels of the ChestX-ray14 dataset (31 studies). For example, Baltruschat et al. (2019a) compares the performance of various approaches to classify the 14 disease labels provided by the ChestX-ray14 dataset. Rajpurkar et al. (2018) compares the performance of an ensemble of deep learning models to board-certified and resident radiologists, showing that their models achieve a performance comparable to ex-

Table 2: Image-Level Prediction Studies (Section 4.1).
Tasks: AA=Adversarial Attack, DA=Domain Adaptation, IC=Interval Change, IG=Image Generation, IR=Image Retrieval, LC=Localization, OT=Other, PR=Preprocessing, RP=Report Parsing, SE=Segmentation, WS=Weak Supervision. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: C=ChestX-Ray14, CM=Cardiomegaly, CV=COVID, E=Edema, GA=Gender/Age, L=Lung, LC=Lung Cancer, LO=Lesion or Opacity, M=MIMIC-CXR, MN=Many, ND=Nodule, OR=Orientation, P=PadChest, PE=Effusion, PL=PLCO, PM=Pneumonia, PT=Pneumothorax, Q=Image Quality, T=Triage/Abnormal, TB=Tuberculosis, TU=Catheter or Tube, X=CheXpert, Z=Other.
Datasets: BL=Belarus, C=ChestX-ray14, CC=COVID-CXR, CG=COVIDGR, J=JSRT+SCR, M=MIMIC-CXR, MO=Montgomery, O=Open-i, P=PadChest, PL=PLCO, PP=Ped-pneumonia, PR=Private, RP=RSNA-Pneumonia, S=Shenzen, SI=SIIM-ACR, SM=Simulated CXR from CT, X=CheXpert.

Citation	Method	Other Tasks	Labels	Datasets
Liu et al. (2019)	Combines lung cropped CXR model and a CXR model to improve model performance	**SE**,LC,PR	C,L	C,J
Sogancioglu et al. (2020)	Comparison of image-level prediction and segmentation models for cardiomegaly	**SE**	CM	C
Que et al. (2018)	A network with DenseNet and U-Net for classification of cardiomegaly	**SE**	CM	C
Li et al. (2019)	U-Net based model for heart and lung segmentation for cardiothoracic ratio	**SE**	CM	PR
Moradi et al. (2020)	Combines lung cropped CXR model and a CXR model using the segmentation quality	**SE**	E,LO,PE,PT,Z	M
E et al. (2019)	Pneumonia detection is improved by use of lung segmentation	**SE**	PM	J,MO,PP,PR
Hurt et al. (2020)	U-Net based model to segment pneumonia	**SE**	PM	RP
Wang et al. (2020c)	Multi-scale DenseNet based model for pneumothorax segmentation	**SE**	PT	PR
Liu et al. (2020b)	DenseNet based U-Net for segmentation of the left and right humerus of the infant	**SE**	Z	PR
Owais et al. (2020)	Uses a database of the intermediate ResNet-50 features to find similar studies	**OT**,IR	TB	MO,S
Ouyang et al. (2020)	Uses activation and gradient based attention for localization and classification	**LC**	C,X	C
Rajaraman et al. (2020b)	Detects and localizes COVID-19 using various networks and ensembling	**LC**	CV	C,CC,PP,RP,X
Samala et al. (2021)	GoogleNet trained with CXR patches, correlates with COVID-19 severity score	**LC**	CV,PM	C,PR
Park et al. (2020)	Proposes a segmentation and classification model compares with radiologist cohort	**LC**	LO,ND,PE,PT	PR
Nam et al. (2019)	Trains a semisupervised network on a large CXR dataset with CT-confirmed nodule cases	**LC**	ND	PR
Pesce et al. (2019)	Defines a loss that minimizes the saliency map errors to improve model performance	**LC**	ND	PR
Taghanaki et al. (2019b)	A weakly supervised localization with variational model, leverages attention maps	**LC**	PM	C
Li et al. (2019)	Attention guided CNN for pneumonia detection with bounding boxes	**LC**	PM	RP
Hwang et al. (2019b)	A CNN for identification of abnormal CXRs and localization of abnormalities	**LC**	T	PR
Lenis et al. (2020)	Introduces a visualization method to identify regions of interest from classification	**LC**	TB	PR
Hwang and Kim (2016)	Weakly supervised framework jointly trained with localization and classification	**LC**	TB	PR
Wang et al. (2018)	Combines classification loss and autoencoder reconstruction loss	**IG**,SE	T	J,MO,O,S
Seah et al. (2019)	Wasserstein GAN to permute diseased radiographs to appear healthy	**IG**,LC	Z	PR
Wolleb et al. (2020)	Novel GAN model trained with healthy and abnormal CXR to predict difference map	**IG**	PE	SM,X
Tang et al. (2019c)	GANs with U-Net autoencoder and CNN discriminator and encoder for one-class learning	**IG**	T	C
Mao et al. (2020)	Autoencoder uses uncertainty for reconstruction error in one-class learning setting	**IG**	T	PP,RP
Lenga et al. (2020)	Continual learning methods to classify data from new domains	**DA**	C,M	C,M
Tang et al. (2019a)	CycleGAN model to adapt adult to pediatric CXR for pneumonia classification	**DA**	PM	PP,RP
Gyawali et al. (2019)	Trains a Variational Autoencoder, uses encoded features to train models	WS	X	X
Gyawali et al. (2020)	Predicts labels for unlabeled data using latent space similarity for semisupervision	WS	X	X
McManigle et al. (2020)	Y-Net to normalize image geometry for preprocessing	SE,PR	OR	C,M,X
Blain et al. (2020)	COVID-19 opacity localization and severity detection on CXRs	SE,LC	CV	PR
Oh et al. (2020)	ResNet-18 backbone for Covid-19 classification with limited data availability	SE,LC	CV,PM	CC,J,MO
Tabik et al. (2020)	Proposes a new dataset COVIDGR and a novel method using transformations with GANs	SE,IG	CV	CG
Junior et al. (2021)	DenseNet for cardiomegaly detection given lung cropped CXR	SE	CM	O,P
Tartaglione et al. (2020)	Multiple models and combinations of CXR datasets used for COVID-19 detection	SE	CV	C,CC,PR,RP
Kusakunniran et al. (2021)	ResNet-101 trained for COVID-19, heatmaps are generated for lung-segmented regions	SE	CV,PM	PR
Narayanan et al. (2020)	Multiple architectures considered for two-stage classification of pediatric pneumonia	SE	PM	PP
Rajaraman et al. (2019b)	Compares visualization methods for pneumonia localization	SE	PM	PP
Blumenfeld et al. (2018)	Classifies patches and uses the positive area size to classify the image	SE	PT	PR
Rajaraman et al. (2018b)	Feature extraction from CNN models and ensembling methods	SE	TB	MO,PR,S
Subramanian et al. (2019)	Detection of central venous catheters using segmentation shape analysis	SE	TU	C
Mansoor et al. (2016)	Detection of air-trapping in pediatric CXRs using Stacked Autoencoders	SE	Z	PR
Wang et al. (2020e)	Pneumoconiosis detection using Inception-v3 and evaluation against two radiologists	SE	Z	PR
Irvin et al. (2019)	Introduces CheXpert dataset and model performance on radiologist labeled test set	RP,LC	X	X
Oh et al. (2019)	Curates data for interval change detection, proposes method comparing local features	RP,IC	Z	PR
Daniels and Metaxas (2019)	Parses reports to define a topic model and predicts those using CXRs	RP	CM,PE,Z	O
Chauhan et al. (2020)	Trains model using image and reports to improve image only performance	RP	E	M
Karargyris et al. (2019b)	Extracts ambiguity of labels from reports, proposes model that uses this information	RP	E,PT,Z	M
Syeda-Mahmood et al. (2019)	Creates and parses reports for ChestX-ray14 AP data to obtain 73 labels for training	RP	MN	C
Laserson et al. (2018)	Obtains findings by tagging common report sentences to train models	RP	MN	PR
Annarumma et al. (2019)	An ensemble of two CNNs to predict priority level for CXR queue management	RP	T	PR
Baltruschat et al. (2019b)	Evaluates bone suppression and lung segmentation, detection of 8 abnormalities	PR,SE	CM,PE,PT,Z	O
Ferreira et al. (2020)	Classification of pediatric pneumonia types using adaped VGG-16 architecture	PR,SE	PM	PP
M. S. et al. (2019)	Evaluates various image preprocessing algorithms on the performance of DenseNet-121	PR	T	C,MO,PR,S
Baltruschat et al. (2020)	Detects 8 findings and analyzes how these can improve workflow prioritization	OT	CM,PE,PT,Z	C,O
Hermoza et al. (2020)	Proposes a model for weakly supervised classification and localization	LC,WS	C	C
Saednia et al. (2020)	Proposes a recurrent attention mechanism to improve model performance	LC	C	C
Baltruschat et al. (2019a)	Evaluates the use of various model configurations for classification	LC	C	C
Cai et al. (2018)	Attention mining and knowledge preservation for classification with localization	LC	C	C
Wang et al. (2020b)	Attention based model compared with well-known architectures	LC	C	C
Ma et al. (2019)	Minimizes the encoding differences of a CXR from multiple models	LC	C	C,X
Cohen et al. (2020a)	DenseNet used to predict COVID-19 severity as scored by radiologists	LC	CV	CC
Wang et al. (2021b)	Uses a ResNet-50 backed segmentation model to detect healthy, pneumonia, COVID-19	LC	CV,PM	CC,RP
Schwab et al. (2020)	Uses multi instance learning for classification with localization	LC	E,PM,PT	M,PR,RP
Yoo et al. (2020)	Lung cancer and nodule prediction using ResNet-34	LC	LC,ND	PR
Kashyap et al. (2020)	GradCam based attention mining loss, compared with labels extracted from reports	LC	LO	M
Majkowska et al. (2019)	Trains Xception using ¿750k CXRs, compares results with radiologist labels	LC	LO,ND,PT,Z	C,PR

continued on the next page

Table 2: Image-Level Prediction Studies (Section 4.1).
Tasks: AA=Adversarial Attack, DA=Domain Adaptation, IC=Interval Change, IG=Image Generation, IR=Image Retrieval, LC=Localization, OT=Other, PR=Preprocessing, RP=Report Parsing, SE=Segmentation, WS=Weak Supervision. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: C=ChestX-Ray14, CM=Cardiomegaly, CV=COVID, E=Edema, GA=Gender/Age, L=Lung, LC=Lung Cancer, LO=Lesion or Opacity, M=MIMIC-CXR, MN=Many, ND=Nodule, OR=Orientation, P=PadChest, PE=Effusion, PL=PLCO, PM=Pneumonia, PT=Pneumothorax, Q=Image Quality, T=Triage/Abnormal, TB=Tuberculosis, TU=Catheter or Tube, X=CheXpert, Z=Other.
Datasets: BL=Belarus, C=ChestX-ray14, CC=COVID-CXR, CG=COVIDGR, J=JSRT+SCR, M=MIMIC-CXR, MO=Montgomery, O=Open-i, P=PadChest, PL=PLCO, PP=Ped-pneumonia, PR=Private, RP=RSNA-Pneumonia, S=Shenzen, SI=SIIM-ACR, SM=Simulated CXR from CT, X=CheXpert.

continued from the previous page

Citation	Method	Other Tasks	Labels	Datasets
Hosch et al. (2020)	ResNet and VGG used to distinguish AP from PA images	LC	OR	PR,RP
Rueckel et al. (2020)	DenseNet-121 trained on public data evaluated using CT-based labels	LC	PE,PM	C,PL
Ureta et al. (2020)	Evaluates the performance of various models trained on pediatric CXRs on adult CXRs	LC	PM	PP
Chen et al. (2020b)	Evaluates ensembling methods and visualization on pediatric CXRs	LC	PM,PT,Z	PR
Yi et al. (2020)	Compares a ResNet-152 against radiologists and shows the statistical significance	LC	PT	C
Crosby et al. (2020a)	Compares GradCAM with radiologist segmentations for evaluation of VGG-19	LC	PT	C,PR
Crosby et al. (2020c)	Apical regions and patches from them extracted to detect pneumothorax	LC	PT	PR
Tang et al. (2020)	Detection of abnormality, various networks compared with radiologist labeling	LC	T	C,O,PP,RP
Rajaraman et al. (2020a)	Evaluates pre-training on ImageNet and CheXpert on various models/settings	LC	T	RP
Nakao et al. (2021)	Proposes a GAN-based model trained only with healthy images for anomaly detection	LC	T	RP
Pasa et al. (2019)	Proposes a new model for faster classification of TB	LC	TB	BL,MO,S
Hwang et al. (2019c)	Evalutes the use of a ResNet based model on a large gold standard dataset	LC	TB	PR
Singh et al. (2019)	Evaluates multiple models for detection of feeding tube malpositioning	LC	TU	PR
Chakravarty et al. (2020)	Graph CNN solution with ensembling which models disease dependencies	LC	X	X
Matsumoto et al. (2020)	Curates a dataset of heart failure cases and evaluates VGG-16 on it	LC	Z	C
Su et al. (2021)	CNN for identifiying the presence of subphrenic free air from CXR	LC	Z	PR
Zou et al. (2020)	Evaluates several models to predict hypertension and artery systolic pressure	LC	Z	PR
Zucker et al. (2020)	Uses ResNet-18 to measure the Brassfield Score, predicts Cystic Fibrosis based on it	LC	Z	PR
Campo et al. (2018)	Simulates CXRs from CT scans and predicts emphysema scores	LC	Z	PR
Toba et al. (2020)	Inception network to predict pulmonary to systemic flow ratio from pediatric CXR	LC	Z	PR
Li et al. (2020a)	Predicts COVID-19 severity by comparing CXRs to previous ones	IC,LC	CV	PR
Luo et al. (2020)	Addresses domain and label discrepancies in multi-dataset training	DA	C,TB,X	C,PR,X
Xue et al. (2019)	Method to increase robustness of CNN classifiers to adversarial samples	AA,IG	PM	RP
Li and Zhu (2020)	Uses the features extracted from the training dataset to detect adversarial CXRs	AA	C	C
Anand et al. (2020)	Self-supervision and adversarial training improves on transfer learning	AA	PM	PP
Khatibi et al. (2021)	Claims 0.99 AUC for predicting TB, uses complex feature engineering and ensembling			
Schroeder et al. (2021)	ResNet model trained with frontal and lateral images to predict COPD with PFT results			PR
Zhang et al. (2021a)	One-class identification of viral pneumonia cases compared with binary classification			PR
Balachandar et al. (2020)	A distributed learning method that overcomes problems of multi-institutional settings		C	C
Burwinkel et al. (2019)	Geometric deep learning including metadata with graph structure. Application to CXR		C	C
Nugroho (2021)	Proposes a new weighting scheme to impove abnormality classification		C	C
DSouza et al. (2019)	ResNet-34 used with various training settings for multi-label classification		C	C
Sirazitdinov et al. (2019)	Investigates effect of data augmentations on classification with Inception-Resnet-v2		C	C
Mao et al. (2018)	Proposes a variational/generative architecture, demonstrates performance on CXRs		C	C
Rajpurkar et al. (2018)	Evaluates the performance of an ensemble against many radiologists		C	C
Kurmann et al. (2019)	Novel method for multi-label classification, application to CXR		C	C
Paul et al. (2020)	Defines a few-shot learning method by extracting features from autoencoders		C	C
Unnikrishnan et al. (2020)	Mean teacher inspired a probablistic graphical model with a novel loss		C	C
Michael and Yoon (2020)	Examines the effect of denoising on pathology classification using DenseNet-121		C	C
Wang et al. (2021a)	Proposes integrating three attention mechanisms that work at different levels		C	C
Paul et al. (2021b)	Step-wise trained CNN and saliency-based autoeencoder for few shot learning		C	C,O
Paul et al. (2021a)	Uses CT and CXR reports with CXR images during training to diagnose unseen diseases		C	C,PR
Bustos et al. (2020)	Proposes a new dataset PadChest with multi-label labels and radiology reports		C	P
Li et al. (2021a)	Lesion detection network used to improve image-level classification		C	PR
Ghesu et al. (2019)	Method to produce confidence measure alongside probability, uses DenseNet-121		C,PL	C,PL
Haghighi et al. (2020)	Uses self-supervised learning for pretraining, compares with ImageNet pretraining		C,PT	C,SI
Zhou et al. (2020a)	Proposes a new CXR pre-training method, compares with pre-training on ImageNet		C,X	C,RP,X
Chen et al. (2020a)	Proposes a graph convolutional network framework which models disease dependencies		C,X	C,X
Zhou et al. (2019)	Compares several models for the detection of cardiomegaly		CM	C
Bougias et al. (2020)	Tests four off-the-shelf networks for prediction of cardiomegaly		CM	PR
Brestel et al. (2018)	Inception v3 trained to detect 4 abnormalities and compared with expert observers		CM,E,LO,Z	PR
Cicero et al. (2017)	GoogLeNet to classify normal and 5 abnormalities on a large proprietary dataset		CM,E,PE,PT,Z	PR
Bar et al. (2015a)	Compares the performance of deep learning with traditional feature extraction methods		CM,PE	PR
Bar et al. (2015b)	ImageNet pre-training and feature extraction methods for pathology detection		CM,PE,Z	PR
Griner et al. (2021)	An ensemble of DenseNet-121 networks used for COVID-19 classification		CV	C,PR
Hu et al. (2021)	Investigates the value of soft tissue CXR for training DenseNet-121 for COVID-19		CV	C,PR,RP
Zhu et al. (2020)	Labels and predicts COVID-19 severity stage using CNN		CV	CC
Fricks et al. (2021)	Uses a model trained on COVID-19 cases to evaluate the effect of an imaging parameter		CV	PR
Wehbe et al. (2020)	COVID-19 detection based on RT-PCR labels, evaluates an ensamble against radiologists		CV	PR
Castiglioni et al. (2021)	Ensemble of ResNet models for COVID-19 detection		CV	PR
Zhang et al. (2021c)	Compares the performance of a DenseNet-121 ensemble to radiologists		CV,PM	PR
Wang et al. (2019)	Various models and use of semi-supervised labels for edema severity estimation		E	M
Karargyris et al. (2019a)	Age prediction on PA or AP images using DenseNet-169		GA	C
Xue et al. (2018b)	Gender prediction using features from deep-learning models in traditional classifiers		GA	J,MO,O,PR,S
Sabottke et al. (2020)	Age prediction on AP images using DenseNet-121 and ResNet-50		GA	X
Lu et al. (2020a)	Combines the CXR with age/sex/smoking history to predict the lung cancer risk		LC	PL
Thammarach et al. (2020)	Densenet-121 pre-trained with public data used to identify 6 classes		LC,T,TB,Z	PR
Kuo et al. (2021)	Evaluates deep learning on pictures of CXRs captured with mobile phones		M,X	M,X

continued on the next page

Table 2: Image-Level Prediction Studies (Section 4.1).
Tasks: AA=Adversarial Attack, DA=Domain Adaptation, IC=Interval Change, IG=Image Generation, IR=Image Retrieval, LC=Localization, OT=Other, PR=Preprocessing, RP=Report Parsing, SE=Segmentation, WS=Weak Supervision. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: C=ChestX-Ray14, CM=Cardiomegaly, CV=COVID, E=Edema, GA=Gender/Age, L=Lung, LC=Lung Cancer, LO=Lesion or Opacity, M=MIMIC-CXR, MN=Many, ND=Nodule, OR=Orientation, P=PadChest, PE=Effusion, PL=PLCO, PM=Pneumonia, PT=Pneumothorax, Q=Image Quality, T=Triage/Abnormal, TB=Tuberculosis, TU=Catheter or Tube, X=CheXpert, Z=Other.
Datasets: BL=Belarus, C=ChestX-ray14, CC=COVID-CXR, CG=COVIDGR, J=JSRT+SCR, M=MIMIC-CXR, MO=Montgomery, O=Open-i, P=PadChest, PL=PLCO, PP=Ped-pneumonia, PR=Private, RP=RSNA-Pneumonia, S=Shenzen, SI=SIIM-ACR, SM=Simulated CXR from CT, X=CheXpert.

continued from the previous page

Citation	Method	Other Tasks	Labels	Datasets
Cohen et al. (2020b)	Investigates the domain and label shift across publicly available CXR datasets		MN	C,M,O,P,RP,X
Hashir et al. (2020)	Explores the use of the lateral view CXR for classification of 64 different labels		MN,P	P
Rajkomar et al. (2017)	Classification of CXRs as Frontal or Lateral using GoogLeNet architecture		OR	PR
Crosby et al. (2019)	Assesses the effect of imprinted labels on AP/PA classification		OR	PR
Crosby et al. (2020b)	Distinguishes the CXR orientation, bone CXRs and soft tissue CXRs from dual energy		OR,Z	PR
Bertrand et al. (2019)	Compares PA and Lateral images for pathology detection with DenseNet		P	P
Chen et al. (2019)	Introduces a loss term that uses the label hierarchy to improve model performance		PL	PL
Shah et al. (2020)	Trains VGG-16 on Ped-pneumonia dataset		PM	PP
Qu et al. (2020)	Methods to mitigate imbalanced class sizes. Applied to CXR using ResNet-18		PM	PP
Yue et al. (2020)	Evaluation of MobileNet to detect pneumonia on pediatric CXRs		PM	PP
Elshennawy and Ibrahim (2020)	Compares multiple architectures for pneumonia detection		PM	PP
Mittal et al. (2020)	Evaluates various capsule network architectures for pediatric pneumonia detection		PM	PP
Longjiang et al. (2020)	Uses ResNet-50 to classify paediatric pneumonia		PM	PR
Ganesan et al. (2019)	Compares traditional and generative data augmentation techniques on CXRs		PM	RP
Ravishankar et al. (2019)	Addresses catastrophic forgetting, application to pneumothorax detection using VGG-13		PT	C
Taylor et al. (2018)	Construction of large dataset, multiple architectures and hyperparameters optimized		PT	PR
Kitamura and Deible (2020)	Model pre-trained with public data and fine-tuned for pneumothorax detection		PT	PR
Moradi et al. (2019a)	DenseNet-121 used to detect CXRs with acquisition-based defects		Q	C
Takaki et al. (2020)	GoogleNet combined with rule-based approach to determine the image quality		Q	PR
Pan et al. (2019a)	Detects abnormal CXRs using several models. Evaluates on independent private data		T	C
Moradi et al. (2019b)	Defines a model on top of features extracted from Inception-ResNet-v2 for triaging		T	C
Wong et al. (2020)	Collects features from pretrained models and adds a CNN on top for triaging		T	C,M
Jang et al. (2020)	Studies the effect of various label noise levels on classification with DenseNet-121		T	C,PR,X
Dunnmon et al. (2019)	Various models for detection of abnormal CXRs, effect of different training set sizes		T	PR
Nam et al. (2020)	Defines 10 abnormalities to define a triaging model and uses CT based test labels		T	PR
Dyer et al. (2021)	Ensembe of DenseNet and EffecientNet for identification of normal CXR		T	PR
Ogawa et al. (2019)	Examines the use of data augmentation in small data setting		T	PR
Ellis et al. (2020)	Evaluation of extra supervision in the form of localized region of interest		T	PR
Rajaraman et al. (2019a)	Evalutates various models and ensembling methods for the triage task		T	RP
Lakhani and Sundaram (2017)	Evaluates deep learning approaches for tuberculosis detection		TB	BL,MO,PR,S
Hwang et al. (2016)	Evaluates the use of transfer learning for tuberculosis detection		TB	MO,PR,S
Sivaramakrishnan et al. (2018)	Extracts feaures using off-the-shelf models and trains a model using those		TB	MO,PR,S
Ayaz et al. (2021)	Combines hand-crafted features and CNN for tuberculosis diagnosis		TB	MO,S
Ul Abideen et al. (2020)	Evaluates a Bayesian-based CNN for detection of TB		TB	MO,S
Rajpurkar et al. (2020)	Evaluates assisting clinicians with an AI based system to improve diagnosis of TB		TB	PR
Heo et al. (2019)	Various architectures, inclusion of patient demographics in model considered		TB	PR
Kim et al. (2018)	Addresses preservation of learned data, application to TB detection using ResNet-21		TB	PR
Gozes and Greenspan (2019)	Pre-training using CXR pathology and metadata labels, application to TB detection		TB	S
Rajaraman and Antani (2020)	Compares various models using various pretraining and ensembling strategies		TB	S
Lakhani (2017)	Evaluates models on detecting the position of feeding tube in abdominal and CXRs		TU	PR
Mitra et al. (2020)	Comparison of seven architectures and ensembling for detection of nine pathologies		X	X
Pham et al. (2020)	A method to incorporate label dependencies and uncertainty data during classification		X	X
Rajan et al. (2021)	Proposes self-training and student-teacher model for sample effeciency		X	X
Calli et al. (2019)	Analyses the effect of label noise in training and test datasets		Z	C
Deshpande et al. (2020)	Labels 6 different foreign object types and detects using various architectures		Z	M
Lu et al. (2019)	Evaluates the use of CXRs to predict long term mortality using Inception-v4		Z	PL
Zhang et al. (2021b)	Low-res segmentation is used to crop high-res lung areas and predict pneumoconiosis		Z	PR
Devnath et al. (2021)	Pneumoconiosis prediction with DenseNet-121 and SVMs applied to extracted features		Z	PR
Liu et al. (2017)	Detection of coronary artery calcification using various CNN architectures		Z	PR
Hirata et al. (2021)	ResNet-50 for detection of the presence of elevated pulmonary arterial wedge pressure		Z	PR
Kusunose et al. (2020)	A network is designed to identify subjects with elevated pulmonary artery pressure		Z	PR,RP

pert observers in most of the 14 labels provided by ChestX-ray14. Following this, pneumonia is second most studied subject (26 studies). Of the 26 studies that worked with pneumonia, 12 studied pediatric chest X-rays and 11 of those used the Ped-Pneumonia dataset for training and evaluation (Rajaraman et al., 2018a; Yue et al., 2020; Liang and Zheng, 2020; Behzadikhormouji et al., 2020; Elshennawy and Ibrahim, 2020; Ureta et al., 2020; Mittal et al., 2020; Shah et al., 2020; Qu et al., 2020; Ferreira et al., 2020; Anand et al., 2020). Classification to Normal/Abnormal (or Triage) is another commonly studied topic (20 studies). Here, studies aim to distinguish normal CXRs or prioritize urgent/critical cases with the goal of reducing the radiologist workload or improving the reporting time. For example, Annarumma et al. (2019) develops a triaging pipeline based on the urgency of exams. Similarly, Tang et al. (2020) compares the performance of various deep learning models applied to several public chest X-ray datasets for distinguishing abnormal cases. Pneumothorax is another commonly studied condition (18 studies). For example, Taylor et al. (2018) aims to detect potentially critical patients and proposes that such models can be used to alert clinicians. Another common topic is tuberculosis detection (18 studies). The first studies that use deep learn-

ing to detect this infectious disease are Hwang et al. (2016); Lakhani and Sundaram (2017). Performance of a deep learning model and how the assistance of this model improves the radiologist performance is studied by Rajpurkar et al. (2020). This study in particular evaluates the use of extra clinical information such as age, white blood cell count, patient temperature and oxygen saturation to assist the deep learning model. Diagnosis or evaluation of COVID-19 from CXR is another topic that has attracted a lot of interest from researchers (17 studies). For example, Cohen et al. (2020a) predicts the disease severity, similarly Li et al. (2020a) predicts the disease progression by comparing an exam with the previous exams of the patient, and Tartaglione et al. (2020) detects COVID-19 using a very limited amount of data. Other than these most common tasks, there are many studies using deep learning to make Image-level Predictions from CXRs. Other commonly utilized labels are illustrated in Figure 4 and listed in Table 2.

A large proportion of the studies use pre-trained standard architectures that can easily be found in deep learning libraries such as Tensorflow or Pytorch. These architectures are commonly Resnet (He et al., 2016), DenseNet (Huang et al., 2017), Inception (Szegedy et al., 2015), VGG (Simonyan and Zisserman, 2014), or AlexNet (Krizhevsky et al., 2012). The choice of model depth (such as ResNet-18, ResNet-50, DenseNet-121, DenseNet-161) also varies between studies as there is no standard in this design choice. Most of those studies do not introduce methodological novelty but report or compare the performances of multiple architectures on a given task. For example, Deshpande et al. (2020) compares various Resnet and Densenet models using both pretrained and randomly initialized weights on the performance of detecting the existence of foreign objects. Similarly many other studies compare the performance of different architectures with various depths on a given task, for example Singh et al. (2019); Tang et al. (2020); Rajaraman et al. (2020a); Chakravarty et al. (2020). Just like the model depth and architecture, there are many factors that affect the performance of a deep learning model. The effect of various data augmentation and input pre-processing methods are evaluated by Sirazitdinov et al. (2019); M. S. et al. (2019). The effect of increasing or decreasing the image size are evaluated in DSouza et al. (2019); Baltruschat et al. (2019a). Various pretraining schemes are evaluated by (Gozes and Greenspan, 2019; Baltruschat et al., 2019a). More sophisticated pre-processing steps to improve model performance include bone suppression (Baltruschat et al., 2019b; Zhou et al., 2020b) and lung cropping (Liu et al., 2019).

Some studies bring methodological novelty by making use of methods that are known to work well to improve model performance elsewhere. For example, it is known that an ensemble of many models improves performance compared to a single model (Dietterich, 2000). Some studies that make use of this method are Rajpurkar et al. (2018); Rajaraman et al. (2019a); Rajaraman and Antani (2020); Zhang et al. (2021c). Attention mining (or object-region mining, attention-based) models are also found in the literature (Wei et al., 2018). Those models aim to improve performance and add localization capabilities to an image-level prediction model. Some studies making use of attention mining models are Cai et al. (2018); Saednia et al. (2020). Multiple-instance learning (multi-instance learning or MIL) (Xu et al., 2014) is another method that is used to add localization capabilities to image-level prediction models. MIL breaks the input image into smaller parts (instances), makes individual predictions relating to those instances and combines this information to make a prediction for the whole image. Some studies that make use of MIL are (Crosby et al., 2020c; Schwab et al., 2020). Other topics within the literature include model uncertainty (Ul Abideen et al., 2020; Ghesu et al., 2019), quality of the CXR (Moradi et al., 2020; McManigle et al., 2020; Moradi et al., 2019a; Takaki et al., 2020; McManigle et al., 2020) and defence against adversarial attack (Li and Zhu, 2020; Anand et al., 2020; Xue et al., 2019).

The different properties of datasets are also utilized to improve model capabilities or performance. Many of the public datasets make use of labels that are not mutually exclusive. This has resulted in a number of papers addressing the dependencies among abnormality labels (Pham et al., 2020; Chen et al., 2020b; Chakravarty et al., 2020). Since many of the labels are common between datasets from different institutes there has been investigation of the issues related to domain and/or label shift in images from different sources (Luo et al., 2020; Cohen et al., 2020b). The effect of dataset sizes is evaluated by Dunnmon et al. (2019). Semi-supervised learning methods combine a small set of labeled and a large set of unlabeled data to train a model (Gyawali et al., 2019, 2020; Wang et al., 2019; Unnikrishnan et al., 2020).

Most of the studies working on image-level prediction tasks deal with frontal CXR images. The importance of lateral chest X-rays and models that can deal with multiple views are evaluated in Bertrand et al. (2019); Hashir et al. (2020); Bayat et al. (2020).

4.2. Segmentation

Segmentation is one of the most commonly studied subjects in CXR analysis (58 papers) and includes literature focused on the identification of anatomy, foreign objects or abnormalities. The segmentation literature reviewed for this work is detailed fully in Table 3. Anatomical segmentation of the heart, lungs, clavicles or ribs, on chest radiographs, is a core part of many computer aided detection (CAD) pipelines. It is typically used as an initial step of such pipelines to define the region of interest for subsequent image analysis tasks to improve performance and efficiency (Baltruschat et al., 2019b; Wang et al., 2020e; Rajaraman et al., 2019b; Heo et al., 2019; Liu et al., 2019; Mansoor et al., 2016). Further, the segmentation itself can be useful to quantify clinical parameters based on shape or area measurements. For example, cardiothoracic ratio, a clinically used measurement to assess heart enlargement (cardiomegaly), can be directly calculated from heart and lung segmentations (Sogancioglu et al., 2020; Li et al., 2019). Organ segmentation has, for these reasons, become one of the most commonly studied subjects among CXR segmentation tasks as seen in Figure 5.

Another application found in the CXR literature is foreign object segmentation, i.e. catheter, tubes, lines, for which high

Table 3: Segmentation Studies (Section 4.2).
Tasks: DA=Domain Adaptation, IG=Image Generation, IL=Image-level Predictions, LC=Localization, PR=Preprocessing, WS=Weak Supervision. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: C=ChestX-Ray14, CL=Clavicle, CM=Cardiomegaly, CV=COVID, E=Edema, H=Heart, L=Lung, LO=Lesion or Opacity, PE=Effusion, PM=Pneumonia, PT=Pneumothorax, R=Rib, TU=Catheter or Tube, Z=Other.
Datasets: BL=Belarus, C=ChestX-ray14, J=JSRT+SCR, M=MIMIC-CXR, MO=Montgomery, O=Open-i, PP=Ped-pneumonia, PR=Private, RP=RSNA-Pneumonia, S=Shenzen, SI=SIIM-ACR, SM=Simulated CXR from CT.

Citation	Method	Other Tasks	Labels	Datasets
Yu et al. (2020)	A model based on U-Net and Faster R-CNN to detect PICC catether and its tip	**LC**,PR	TU	PR
Zhang et al. (2019b)	Tailored Mask R-CNN for simultaneous detection and segmentation	**LC**	L	PR
Wessel et al. (2019)	Uses Mask R-CNN iteratively to segment and detect ribs.	**LC**	R	PR
Liu et al. (2019)	Combines lung cropped CXR model and a CXR model to improve model performance	**IL**,LC,PR	C,L	C,J
Sogancioglu et al. (2020)	Comparison of image-level prediction and segmentation models for cardiomegaly	**IL**	CM	C
Que et al. (2018)	A network with DenseNet and U-Net for classification of cardiomegaly	**IL**	CM	C
Li et al. (2019)	U-Net based model for heart and lung segmentation for cardiothoracic ratio	**IL**	CM	PR
Moradi et al. (2020)	Combines lung cropped CXR model and a CXR model using the segmentation quality	**IL**	E,LO,PE,PT,Z	M
E et al. (2019)	Pneumonia detection is improved by use of lung segmentation	**IL**	PM	J,MO,PP,PR
Hurt et al. (2020)	U-Net based model to segment pneumonia	**IL**	PM	RP
Wang et al. (2020c)	Multi-scale DenseNet based model for pneumothorax segmentation	**IL**	PT	PR
Liu et al. (2020b)	DenseNet based U-Net for segmentation of the left and right humerus of the infant	**IL**	Z	PR
Tang et al. (2019b)	Attention-based network and CXR synthesis process for data augmentation	**IG**,IG	L	J,MO,PR
Eslami et al. (2020)	Conditional GANs for multi-class segmentation of heart,clavicles and lungs	**IG**	CL,H,L	J
Onodera et al. (2020)	Processing method to produce scatter-corrected CXRs and segments masses with U-Net	**IG**	LO	SM
Oliveira et al. (2020b)	MUNIT based DA model for lung segmentation	**DA**	CL,H,L	J
Dong et al. (2018)	Adversarial training of lung and heart segmentation for DA	**DA**	CM	J,PR
Zhang et al. (2018)	CycleGAN guided by a segmentation module to convert CXR to CT projection images	**DA**	H,L,Z	PR
Chen et al. (2018a)	CycleGAN based DA model with semantic aware loss for lung segmentation	**DA**	L	MO
Oliveira et al. (2020a)	Conditional GANs based DA for bone segmentation	**DA**	R	SM
Shah et al. (2018)	FCN based novel model incorporating weak landmarks and bounding boxes annotations	WS	CL,H,L	J
Bortsova et al. (2019)	U-Net segmentation model integrating unlabeled data through consistency loss	WS	CL,H,L	J
Ouyang et al. (2019)	Attention masks derived from classification model to guide the segmentation model	IL	PT	PR
Frid-Adar et al. (2019)	U-Net based network for classification and segmentation with simulated data	IL	Z	C,J
Sullivan et al. (2020)	U-Net based model for segmentation and a classification for existance of lines	IL	Z	PR
Cardenas et al. (2021)	U-Net for bone suppression given lung-segmented CXR image with patches			PR
Zhang et al. (2020)	Proposes teacher-student based learning with noisy segmentations		CL,H,L	J
Kholiavchenko et al. (2020)	Various FCN based models explored for simultaneous pixel and contour segmentation		CL,H,L	J
Novikov et al. (2018)	Investigates various FCN type architecture including U-Net for organ segmentation		CL,H,L	J
Bonheur et al. (2019)	Capsule networks adapted for multi-class organ segmentation		CL,H,L	J
Arsalan et al. (2020)	U-Net based architecture with residual connections for organ segmentation		CL,H,L	J,MO,S
Wang et al. (2020d)	U-Net based architecture based on dense connections		CL,R	PR
Mortani Barbosa et al. (2021)	CNN trained with CT projection images for quantification of airspace disease		CV	PR
Larrazabal et al. (2020)	Denoising autoencoder as post-processing to improve segmentations		H,L	J
Holste et al. (2020)	Evaluates U-Net performance with various loss functions, and data augmentation		H,L	PR
Mansoor et al. (2020)	Stacked denoising autoencoder model for space and shape parameter estimation		L	BL,J,PR
Amiri et al. (2020)	Investigates the effect of fine-tuning different layers for U-Net based model		L	J
Lu et al. (2020b)	Proposes a human-in-the-loop one shot anatomy segmentor		L	J
Li et al. (2021b)	U-Net with conditional random field post processing for lung segmentation		L	J
Portela et al. (2020)	Investigates U-Net with different optimizer and dropout		L	J
Yahyatabar et al. (2020)	U-Net with dense connections for reducing network parameters for lung segmentation		L	J,MO
Kim and Lee (2021)	U-Net with self attention for lung segmentation		L	J,MO,S
Arbabshirani et al. (2017)	Multi-scale and patch-based CNN to segment lungs		L	J,PR
Rahman et al. (2021)	U-Net based model for lung segmentation trained with CXR patches		L	MO
Souza et al. (2019)	Two stage patch based CNN for refined lung field segmentation		L	MO
Milletari et al. (2018)	Encoder-decoder architecture with ConvLSTM and ResNet for segmentation		L	MO
Zhang et al. (2019c)	Encoder-decoder based CNN with novel edge guidance module for lung segmentation		L	MO
Mathai et al. (2019)	Proposes a convolutional LSTM model for ultrasound, uses CXR as a secondary modality		L	MO
Kitahara et al. (2019)	U-Net based segmentation model for dynamic CXRs		L	PR
Furutani et al. (2019)	U-Net for whole lung region segmentation including where heart overlaps		L	PR
Xue et al. (2020)	Cascaded U-net with sample selection with imperfect segmentations		L	S
Wang et al. (2020a)	ResNet-50 based architecture with segmentation and classification branches		PT	SI
Tolkachev et al. (2020)	Investigates U-Net based models with various backbone encoders for pneumothorax		PT	SI
Groza and Kuzin (2020)	Ensemble of three LinkNet based networks and with multi-step postprocessing		PT	SI
Xue et al. (2018d)	Cascaded network with Faster R-CNN and U-Net for aortic knuckle		Z	J
Yi et al. (2019a)	Multi-scale U-Net based model with recurrent module for foreign objects		Z	O
Lee et al. (2018)	Two FCN to segment peripherally inserted central catheter line and its tip		Z	PR
Pan et al. (2019b)	Two Mask R-CNN to segment the spine and vertebral bodies and calculate the Cobb angle		Z	PR

performance levels have been reported using deep learning (Lee et al., 2018; Frid-Adar et al., 2019; Sullivan et al., 2020).

Interestingly, only a small number of works addressed segmentation of abnormalities. Hurt et al. (2020) focused on segmentation of pneumonia, and Tolkachev et al. (2020) developed a method to segment pneumothorax. Both of these works used recently published challenge datasets (hosted by Kaggle), namely RSNA-Pneumonia and SIIM-ACR. In general, the determination of abnormal locations on CXR is dominated by methods which addressed this as a localization task (i.e. via bounding-box type annotations) rather than exact delineation of abnormalities through segmentation. This is likely to be attributable to the difficulty of precise annotation on a projection image and to the high annotation cost for precise segmentations.

A small number of works tackled the segmentation task using a patch-based CNN, which is trained to classify the center of pixel in the patch as foreground or background by means of sliding-window approach (Arbabshirani et al., 2017; Souza

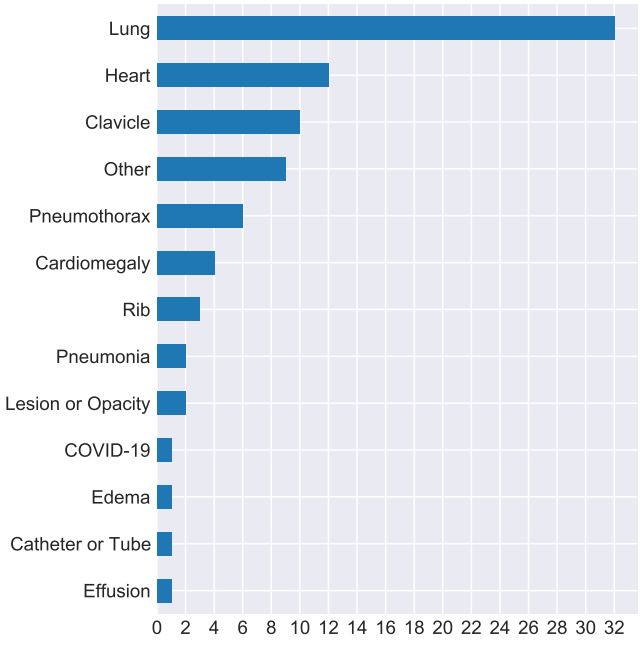

Figure 5: Number of studies for the Segmentation labels. 58 papers are included, each may study more than one label.

et al., 2019). However, this approach is generally considered inefficient for segmentation and most works use fully convolutional networks (FCN) (Shelhamer et al., 2017), which can take larger, arbitrary sized, images as input and produce a similar sized, per-pixel prediction, likelihood map in a single forward pass. In particular, the U-Net architecture (Ronneberger et al., 2015), a type of FCN, dominates the field with 50% of segmentation works in literature (29/58) employing it or some similar variant. Successful applications were built with this architecture to segment organs (Novikov et al., 2018; Furutani et al., 2019; Kitahara et al., 2019), pneumonia (Hurt et al., 2020) and foreign objects (Sullivan et al., 2020; Frid-Adar et al., 2019). For example, Novikov et al. (2018) compared three U-Net variant architectures for multi-class segmentation of the heart, clavicles and lungs on the JSRT dataset. Using regularization to prevent over-fitting and weighted cross entropy loss to balance the dataset, they outperformed the human observer at heart and lung segmentation. This result was in line with other works (Kholiavchenko et al., 2020; Bortsova et al., 2019; Arsalan et al., 2020) employing FCN-type architectures which also achieved very high performance levels on this dataset.

One commonly encountered challenge is that many algorithms produce noisy segmentation maps. In order to tackle this, several works employed post-processing techniques. Lee et al. (2018) used a probabilistic Hough line transform algorithm to remove false positives and produce a smoother segmentation of peripherally inserted central catheters (PICC). Groza and Kuzin (2020) used a heuristic approach to average cross-fold predictions with an optimized binarization threshold and a dilation technique for pneumothorax segmentation. Some authors proposed to learn post-processing by training an independent network, inputting segmentation predictions for refinement, rather than using conventional methods. For example, Larrazabal et al. (2020) used denoising autoencoders, trained to produce anatomically plausible segmentations from the initial predictions. Similarly, Souza et al. (2019) used a FCN to refine segmentation predictions. The final segmentation was achieved by combining the initial and reconstructed segmentation results.

A number of researchers used a multi-stage training strategy, where network predictions are refined in several steps during training (Wessel et al., 2019; Souza et al., 2019; Xue et al., 2018d, 2020). For example, Xue et al. (2018d) employed faster-RCNN to produce coarse segmentation results, which were then used to crop the images to a region of interest, which was provided to a U-Net trained to predict the final segmentation result. Similarly, Souza et al. (2019) employed two networks, where the second network received the predictions of the first to refine the segmentation results. Wessel et al. (2019) trained separate networks for segmentation of each rib in chest radiographs based on Mask R-CNN. The predicted segmentation results from the rib above was fed to each network as an additional input.

Although most of the works in the literature harnessed FCN architectures, a few authors employed recurrent neural networks (RNN) for segmentation tasks (Yi et al., 2019a; Milletari et al., 2018; Mathai et al., 2019) and report good performance. Milletari et al. (2018) proposed a novel architecture where the decoding component was long short term memory (LSTM) architecture to obtain multi-scale feature integration. The proposed approach achieved a Dice score of 0.97 for lung segmentation on Montgomery dataset. Similarly, Yi et al. (2019b) developed a scale RNN, a network based on encoder and decoder architecture with recurrent modules, for segmentation of catheter and tubes on pediatric chest X-rays.

The high cost of obtaining segmentation annotations motivates the development of segmentation systems which incorporate weak-labels or simulated datasets with the aim of reducing annotation costs (Frid-Adar et al., 2019; Ouyang et al., 2019; Lu et al., 2020b; Yi et al., 2019a). Several works addressed this using weakly supervised learning approaches (Lu et al., 2020b; Ouyang et al., 2019). Lu et al. (2020b) proposed a graph convolutional network based architecture which required only one labeled image and leveraged large amounts of unlabeled data (one-shot learning) through a newly introduced three contour-based loss function. Ouyang et al. (2019) proposed a pneumothorax segmentation framework which incorporated both images with pixel level annotations and weak image-level annotations. The authors trained an image classification network, ResNet-101, with weakly labeled data to derive attention maps. These attention maps were then used to train a segmentation model, Tiramisu, together with pixel level annotations.

4.3. Localization

Localization refers to the identification of a region of interest using a bounding box or point coordinates rather than a more specific pixel segmentation. In this section we discuss only the CXR localization literature which provides a quantitative evaluation of this task. It should be noted that there are many other works which train networks for an image-level prediction task

and provide some examples of heatmaps (e.g., saliency map or GradCAM) to suggest which region of the image determines the label. While this may be considered as a form of localization, these heatmaps are rarely quantitatively evaluated and such works are not included here. Table 4 details all the reviewed studies where localization was a primary focus of the work.

The majority of CXR analysis papers performing localization focus on identifying abnormalities rather than objects (e.g., catheter) or anatomy (e.g., ribs). Localization of nodules, tuberculosis and pneumonia are commonly studied applications in the literature, as illustrated in Figure 6.

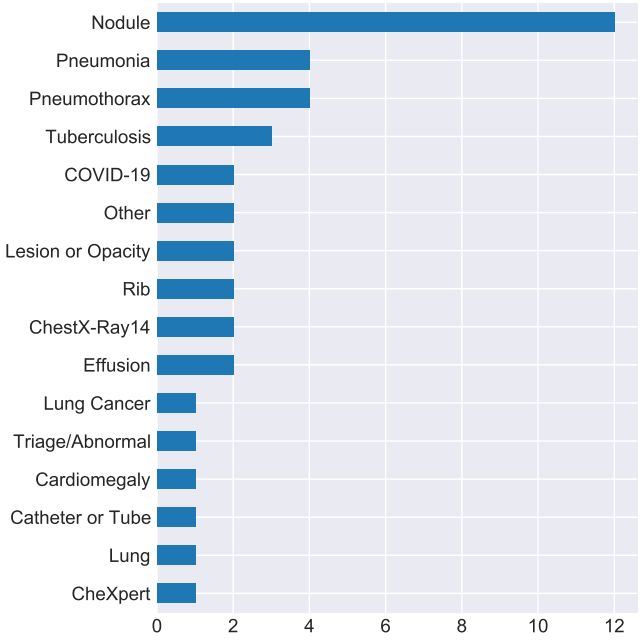

Figure 6: Number of studies for the Localization labels. 30 papers are included, each may study more than one label.

In recent years, a variety of specific architectures, i.e. YOLO, Mask R-CNN, Faster R-CNN, have been designed in computer vision research aiming at developing more accurate and faster algorithms for localization tasks (Zhao et al., 2019). Such state of the art architectures have been rapidly adapted for CXR analysis and shown to achieve high-level performance. For example, Park et al. (2019) demonstrated that the (original) YOLO architecture was successful at identifying the location of pneumothorax on chest radiographs. The model was evaluated on an external dataset with CXRs from 1,319 patients which were obtained after percutaneous transthoracic needle biopsy (PTNB) for pulmonary lesions; it achieved an AUC of 0.898 and 0.905 on 3-h and 1-day follow-up chest radiographs, respectively. Similarly, other studies (Kim et al., 2020; Schultheiss et al., 2020; Takemiya et al., 2019; Kim et al., 2019) harnessed architectures like RetinaNet, Mask R-CNN and RCNN for localization of nodules and masses. Kim et al. (2020) trained RetinaNet and Mask R-CNN for detection of nodule and mass and investigated the optimal input size. The authors showed that, using a square image with 896 pixels as the edge length, RetinaNet and Mask R-CNN achieved FROC of 0.906 and 0.869, respectively.

A number of papers adapted classification architectures (e.g., ResNet, DenseNet) to directly regress landmark locations for CXR localization tasks (Hwang et al., 2019b; Cha et al., 2019). One common way of tackling this is to adapt the networks to produce heatmap predictions and draw boxes around the areas that created the highest signals. For example, Hwang et al. (2019b) tailored a DenseNet-based classifier to produce heatmap predictions for each of four types of CXR abnormalities. The network was trained with pixel-wise cross entropy between the predictions and annotations. Similarly, Cha et al. (2019) adapted ResNet-50 and ResNet-101 architectures for localization of nodules and masses on CXR. Other studies (Xue et al., 2018c; Li et al., 2020c) tackled this problem using patch-based approaches, commonly referred as multiple instance learning, creating patches from chest X-rays and evaluating these for the presence of abnormalities.

One challenge in building robust deep learning localization systems is to collect large annotated datasets. Collecting such annotations is time-consuming and costly which has motivated researchers to build systems incorporating weaker labels during training. This research area is referred to as weakly supervised learning, and has been investigated by numerous works (Hwang and Kim, 2016; Hwang et al., 2019b; Nam et al., 2019; Pesce et al., 2019; Taghanaki et al., 2019b) for localization of a variety of abnormalities in CXR. Most of the works (Hwang et al., 2019b; Pesce et al., 2019; Nam et al., 2019; Hwang and Kim, 2016) leveraged weak image-level labels by adapting a CNN architecture to create two branches for localization (heatmap predictions) and classification. A hybrid loss function was used, combining localization and classification losses, which enabled training of the networks using images without localization annotations.

4.4. Image Generation

There are 35 studies identified in this work whose main focus is Image Generation, as detailed in Table 5. Image generation techniques have been harnessed for a wide variety of purposes including data augmentation (Salehinejad et al., 2019), visualization (Bigolin Lanfredi et al., 2019; Seah et al., 2019), abnormality detection through reconstruction (Tang et al., 2019c; Wolleb et al., 2020), domain adaptation (Zhang et al., 2018) or image enhancement techniques (Lee et al., 2019).

The generative adversarial network (GAN) (Goodfellow et al., 2014; Yi et al., 2019b) has became the method of choice for image generation in CXR and over 50% of the works reviewed here used GAN-based models.

A number of works focused on CXR generation to augment training datasets (Moradi et al., 2018b; Zhang et al., 2019a; Salehinejad et al., 2019) by using unconditional GANs which synthesize images from random noise. For example, Salehinejad et al. (2019) trained a DCGAN model, similar to Moradi et al. (2018b), independently for each class, to generate chest radiographs with five different abnormalities. The authors demonstrated that this augmentation process improved the abnormality classification performance of DCNN classifiers (ResNet, GoogleNet, AlexNet) by balancing the dataset classes.

Table 4: Localization Studies (Section 4.3).
Tasks: IC=Interval Change, IL=Image-level Predictions, PR=Preprocessing, RP=Report Parsing, SE=Segmentation, WS=Weak Supervision. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: C=ChestX-Ray14, CM=Cardiomegaly, CV=COVID, L=Lung, LC=Lung Cancer, LO=Lesion or Opacity, ND=Nodule, PE=Effusion, PM=Pneumonia, PT=Pneumothorax, R=Rib, T=Triage/Abnormal, TB=Tuberculosis, TU=Catheter or Tube, X=CheXpert, Z=Other.
Datasets: C=ChestX-ray14, CC=COVID-CXR, J=JSRT+SCR, M=MIMIC-CXR, O=Open-i, PP=Ped-pneumonia, PR=Private, RP=RSNA-Pneumonia, S=Shenzen, X=CheXpert.

Citation	Method	Other Tasks	Labels	Datasets
Yu et al. (2020)	A model based on U-Net and Faster R-CNN to detect PICC catether and its tip	**SE**,PR	TU	PR
Zhang et al. (2019b)	Tailored Mask R-CNN for simultaneous detection and segmentation	**SE**	L	PR
Wessel et al. (2019)	Uses Mask R-CNN iteratively to segment and detect ribs.	**SE**	R	PR
Ouyang et al. (2020)	Uses activation and gradient based attention for localization and classification	**IL**	C,X	C
Rajaraman et al. (2020b)	Detects and localizes COVID-19 using various networks and ensembling	**IL**	CV	C,CC,PP,RP,X
Samala et al. (2021)	GoogleNet trained with CXR patches, correlates with COVID-19 severity score	**IL**	CV,PM	C,PR
Park et al. (2020)	Proposes a segmentation and classification model compares with radiologist cohort	**IL**	LO,ND,PE,PT	PR
Nam et al. (2019)	Trains a semisupervised network on a large CXR dataset with CT-confirmed nodule cases	**IL**	ND	PR
Pesce et al. (2019)	Defines a loss that minimizes the saliency map errors to improve model performance	**IL**	ND	PR
Taghanaki et al. (2019b)	A weakly supervised localization with variational model, leverages attention maps	**IL**	PM	C
Li et al. (2019)	Attention guided CNN for pneumonia detection with bounding boxes	**IL**	PM	RP
Hwang et al. (2019b)	A CNN for identification of abnormal CXRs and localization of abnormalities	**IL**	T	PR
Lenis et al. (2020)	Introduces a visualization method to identify regions of interest from classification	**IL**	TB	PR
Hwang and Kim (2016)	Weakly supervised framework jointly trained with localization and classification	**IL**	TB	PR
Chen et al. (2020c)	Extract nodule candidates using traditional methods and trains GoogleNet	SE,PR	ND	J
Schultheiss et al. (2020)	RetinaNet for detecting nodules incorporating lung segmentation	SE	ND	J,PR
Tam et al. (2020)	Combines reports and CXRs for weakly supervised localization and classification	RP,WS	PM,PT	C,M
Moradi et al. (2018a)	Proposes a model using LSTM and CNN, combining reports and images as inputs	IL,RP	CM,ND	C,O
Khakzar et al. (2019)	Adversarially trained weakly supervised localization framework for interpretability	IL	C	C
Kim et al. (2020)	Evaluates the effect of image size for nodule detection with Mask R-CNN and RetinaNet	IL	ND	PR
Cho et al. (2020)	Evaluates the reproducibility of YOLO for disease localization in follow up exams	IC	LO,ND,PE,PT,Z	PR
Kim et al. (2019)	Evaluates the reproducability of various detection architectures in follow up exams.	IC	ND	PR
Cha et al. (2019)	ResNet model using CT and surgery based annotations for lung cancer prediction		LC	PR
Takemiya et al. (2019)	R-CNN for localization of lung nodules		ND	J
Wang et al. (2017a)	Fuses AlexNet and hand-crafted features to improve random forest performance		ND	J
Li et al. (2020c)	Patch-based nodule detectin, combines features from different resolutions		ND	J,PR
Park et al. (2019)	Evaluates the detection of pneumothorax before, 3h and 1d after biopsy		PT	PR
Mader et al. (2018)	Proposes a U-Net based model for localizing and labeling individual ribs		R	O
Xue et al. (2018c)	AlexNet for localizing tuberculosis with patch-based approach		TB	S
von Berg et al. (2020)	Localizes anatomical features for image quality check		Z	PR

Another work (Zhang et al., 2019a) proposed a novel GAN architecture to improve the quality of generated CXR by forcing the generator to learn different image representations. The authors proposed SkrGAN, where a sketch prior constraint is introduced by decomposing the generator into two modules for generating a sketched structural representation and the CXR image, respectively.

Abnormality detection is another task which has been addressed through a combination of image generation and one-class learning methods (Tang et al., 2019c; Mao et al., 2020). The underlying idea of these methods is that a generative model trained to reconstruct healthy images will have a high reconstruction error if abnormal images are input at test time, allowing them to be identified. Tang et al. (2019c) harnessed GANs and employed a U-Net type autoencoder to reconstruct images (as the generator), and a CNN-based discriminator and encoder. The discriminator received both reconstructed images and real images to provide supervisory signal for realistic reconstruction through adversarial training.

Similarly, Mao et al. (2020) proposed an autoencoder for abnormality detection which was trained only with healthy images. In this case the autoencoder was tailored to not only reconstruct healthy images but also produce uncertainty predictions. By leveraging uncertainty, the authors proposed a normalized reconstruction error to distinguish abnormal CXR images from normal ones.

The most widely studied subject in the image generation literature is image enhancement. Several researchers investigated bone suppression (Liu et al., 2020a; Matsubara et al., 2020; Zarshenas et al., 2019; Gozes and Greenspan, 2020; Lin et al., 2020; Zhou et al., 2020b) and lung enhancement (Li et al., 2019; Gozes and Greenspan, 2020) techniques to improve image interpretability. A number of works (Liu et al., 2020a; Zhou et al., 2020b) employed GANs to generate bone-suppressed images. For example, Liu et al. (2020a) employed GANs and leveraged additional input to the generator to guide the dual-energy subtraction (DES) soft-tissue image generation process. In this study, bones, edges and clavicles were first segmented by a CNN model, and the resulting edge maps were fed to the generator with the original CXR image as prior knowledge. For building a deep learning model for bone suppressed CXR generation, the paired dual energy (DE) imaging is needed, which is not always available in abundance. Several other studies (Li et al., 2019; Gozes and Greenspan, 2020) addressed this by leveraging digitally reconstructed radiographs for enhancing the lungs and bones in CXR. For instance, Li et al. (2019) trained an autoencoder for generating CXR with bone suppression and lung enhancement, and the knowledge obtained from DRR images were integrated through the encoder.

4.5. Domain Adaptation

Most of the papers surveyed in this work train and test their method on data from the same domain. This finding is inline with the previously reported studies (Kim et al., 2019;

Table 5: Image Generation Studies (Section 4.4).
Tasks: DA=Domain Adaptation, IC=Interval Change, IG=Image Generation, IL=Image-level Predictions, LC=Localization, PR=Preprocessing, RE=Registration, SE=Segmentation, SR=Super Resolution. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: BS=Bone Suppression, C=ChestX-Ray14, CL=Clavicle, CM=Cardiomegaly, CV=COVID, E=Edema, H=Heart, L=Lung, LO=Lesion or Opacity, PE=Effusion, PT=Pneumothorax, T=Triage/Abnormal, TB=Tuberculosis, Z=Other.
Datasets: C=ChestX-ray14, CC=COVID-CXR, J=JSRT+SCR, MO=Montgomery, O=Open-i, PL=PLCO, PP=Ped-pneumonia, PR=Private, RP=RSNA-Pneumonia, S=Shenzen, SM=Simulated CXR from CT, X=CheXpert.

Citation	Method	Other Tasks	Labels	Datasets
Tang et al. (2019b)	Attention-based network and CXR synthesis process for data augmentation	**SE**,IG	L	J,MO,PR
Eslami et al. (2020)	Conditional GANs for multi-class segmentation of heart,clavicles and lungs	**SE**	CL,H,L	J
Onodera et al. (2020)	Processing method to produce scatter-corrected CXRs and segments masses with U-Net	**SE**	LO	SM
Wang et al. (2018)	Combines classification loss and autoencoder reconstruction loss	**IL**,SE	T	J,MO,O,S
Seah et al. (2019)	Wasserstein GAN to permute diseased radiographs to appear healthy	**IL**,LC	Z	PR
Wolleb et al. (2020)	Novel GAN model trained with healthy and abnormal CXR to predict difference map	**IL**	PE	SM,X
Tang et al. (2019c)	GANs with U-Net autoencoder and CNN discriminator and encoder for one-class learning	**IL**	T	C
Mao et al. (2020)	Autoencoder uses uncertainty for reconstruction error in one-class learning setting	**IL**	T	PP,RP
Mahapatra and Ge (2019)	Conditional GAN based DA for image registration using segmentation guidance	**DA**,RE,SE	L	C
Madani et al. (2018)	Adversarial based method adapting new domains for abnormality classification	**DA**,IL	CM	PL
Umehara et al. (2017)	Proposes a patch-based CNN super resolution method	SR	Z	J
Uzunova et al. (2019)	Generates high resolution CXRs using multi-scale, patch based GANs	SR	Z	O
Zhang et al. (2019a)	Novel GAN model with sketch guidance module for high resolution CXR generation	SR	Z	PP
Lin et al. (2020)	AutoEncoder for bone suppression and segmentation with statistical similarity losses	SE,PR	BS	J
Dong et al. (2019)	Uses neural architecture search to find a discriminator network for GANs	SE	H,L	J,PR
Taghanaki et al. (2019a)	Proposes an iterative gradient based input preprocessing for improved performance	SE	L	S
Fang et al. (2020)	Learns transformations to register two CXRs, uses the difference for interval change	RE,IC	Z	PR
Yang et al. (2017)	Generates bone and soft tissue (dual energy) images from CXRs	PR	BS	PR
Zarshenas et al. (2019)	Proposes an CNN with multi-resolution decomposition for bone suppression images	PR	BS	PR
Gozes and Greenspan (2020)	U-Net for bone generation with CT projection images, used for CXR enhancement	PR	BS	SM
Lee et al. (2019)	U-Net based network to generate dual energy CXR	PR	Z	PR
Liu et al. (2020a)	GAN integrates edges of ribs and clavicles to guide DES-like images generation	PR	Z	PR
Xing et al. (2019)	Generates diseased CXRs, evaluates their realness with radiologists and trains models	LC	C	C
Li et al. (2019)	Novel CycleGAN model to decompose CXR images incorporating CT projection images	IL	C	C,PR,SM
Salehinejad et al. (2019)	Uses DCGAN model to generate CXR with abnormalities for data augmentation	IL	CM,E,PE,PT	PR
Albarqouni et al. (2017)	U-Net based architecture to decompose CXR structures, application to TB detection	IL	TB	PR
Moradi et al. (2018b)	Two DCGAN trained with normal and abnormal images for data augmentation	IL	Z	PL
Bigolin Lanfredi et al. (2019)	Novel conditional GAN using lung function test results to visualize COPD progression	IL	Z	PR
Zarei et al. (2021)	Conditional GAN and two variational autoencoders designed for CXR generation			PR
Gomi et al. (2020)	Novel reconstruction algorithm for CXR enhancement			PR
Zhou et al. (2020b)	Bone shadow suppression using conditional GANS with dilated U-Net variant		BS	J
Matsubara et al. (2020)	Generates CXRs from CT to train CNN for bone suppression		BS	PR
Zunair and Hamza (2021)	Generates COVID-19 CXR images to improve network training and performance		CV	CC,RP
Bayat et al. (2020)	2D-to-3D encoder-decoder network for generating 3D spine models from CXR studies		Z	PR
Bigolin Lanfredi et al. (2020)	Generates normal from abnormal CXRs, uses the deformations as disease evidence		Z	PR

Prevedello et al., 2019) and highlights an important concern: most of the performance levels reported in the literature might not generalize well to data from other domains (Zech et al., 2018). Several studies (Yao et al., 2019; Zech et al., 2018; Cohen et al., 2020b) demonstrated that there was a significant drop in performance when deep learning systems were tested on datasets outside their training domain for a variety of CXR applications. For example, Yao et al. (2019) investigated the performance of a DenseNet model for abnormality classification on CXR images using 10 diverse datasets varied by their location and patient distributions. The authors empirically demonstrated that there was a substantial drop in performance when a model was trained on a single dataset and tested on the other domains. Zech et al. (2018) observed a similar finding for pneumonia detection on chest radiographs.

Domain adaptation (DA) methods investigate how to improve the performance of a model on a dataset from a different domain than the training set. In CXR analysis, DA methods have been investigated in three main settings; adaptation of CXR images acquired from different hardware, adaptation of pediatric to adult CXR and adaptation of digitally reconstructed radiographs (generated by average intensity projections from CT) to real CXR images. All domain adaptation studies, and studies on generalization reviewed in this work are detailed in Table 6.

Most of the research on DA for CXR analysis harnessed adversarial-based DA methods, which either use generative models (e.g., CycleGANs) or non-generative models to adapt to new domains using a variety of different approaches. For example, Dong et al. (2018) investigated an unsupervised domain adaptation based on adversarial training for lung and heart segmentation. In this approach, a discriminator network, ResNet, learned to discriminate between segmentation predictions (heart and lung) from the target domain and reference standard segmentations from the source domain. This approach forced the FCN-based segmentation network to learn domain invariant features and produce realistic segmentation maps. A number of works (Chen et al., 2018a; Zhang et al., 2018; Oliveira and dos Santos, 2018) addressed unsupervised DA using CycleGAN-based models to transform source images to resemble those from the target domain. For example, Zhang et al. (2018) used a CycleGAN-based architecture to adapt CXR images to digitally reconstructed radiographs (DRR) (generated from CT scans), for anatomy segmentation in CXR. A CycleGAN-based model was employed to convert the CXR image appearance and a U-Net variant architecture to simultane-

Table 6: Domain Adaptation Studies (Section 4.5).
Tasks: IG=Image Generation, IL=Image-level Predictions, RE=Registration, SE=Segmentation. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: C=ChestX-Ray14, CL=Clavicle, CM=Cardiomegaly, H=Heart, L=Lung, M=MIMIC-CXR, PM=Pneumonia, R=Rib, TB=Tuberculosis, Z=Other.
Datasets: C=ChestX-ray14, J=JSRT+SCR, M=MIMIC-CXR, MO=Montgomery, O=Open-i, PL=PLCO, PP=Ped-pneumonia, PR=Private, RP=RSNA-Pneumonia, S=Shenzen, SM=Simulated CXR from CT.

Citation	Method	Other Tasks	Labels	Datasets
Oliveira et al. (2020b)	MUNIT based DA model for lung segmentation	**SE**	CL,H,L	J
Dong et al. (2018)	Adversarial training of lung and heart segmentation for DA	**SE**	CM	J,PR
Zhang et al. (2018)	CycleGAN guided by a segmentation module to convert CXR to CT projection images	**SE**	H,L,Z	PR
Chen et al. (2018a)	CycleGAN based DA model with semantic aware loss for lung segmentation	**SE**	L	MO
Oliveira et al. (2020a)	Conditional GANs based DA for bone segmentation	**SE**	R	SM
Lenga et al. (2020)	Continual learning methods to classify data from new domains	**IL**	C,M	C,M
Tang et al. (2019a)	CycleGAN model to adapt adult to pediatric CXR for pneumonia classification	**IL**	PM	PP,RP
Mahapatra and Ge (2019)	Conditional GAN based DA for image registration using segmentation guidance	**IG**,RE,SE	L	C
Madani et al. (2018)	Adversarial based method adapting new domains for abnormality classification	**IG**,IL	CM	PL
Zech et al. (2018)	Assessment of generalization to data from different institutes	IL	PM	C,O
Sathitratanacheewin et al. (2020)	Demonstrates the effect of training and test on data from different domains	IL	TB	S

ously segment organs of interest. Similarly, CycleGAN-based models were adapted to transfer DRR images to resemble CXR images for bone segmentation (Oliveira et al., 2020a) and to transform adult CXR to pediatric CXR for pneumonia classification (Tang et al., 2019c).

Unlike most of the studies which utilized DA methods in unsupervised setting, a few studies considered supervised and semi-supervised approaches to adapt to the target domain. Oliveira et al. (2020b) employed a MUNIT-based architecture (Huang et al., 2018) to map target images to resemble source images, subsequently feeding the transformed images to the segmentation model. The authors investigated both unsupervised and semi-supervised approaches in this work, where some labels from the target domain were available. Another work by Lenga et al. (2020) studied several recently proposed continual learning approaches, namely joint training, elastic weight consolidation and learning without forgetting, to improve the performance on a target domain and to mitigate effectively catastrophic forgetting for the source domain. The authors evaluated these methods for 2 publicly available datasets, ChestX-ray14 and MIMIC-CXR, for a multi-class abnormality classification task and demonstrated that joint training achieved the best performance.

4.6. Other Applications

In this section we review articles with a primary application that does not fit into any of the categories detailed in Sections 4.1 to 4.5 (14 studies). These works are detailed fully in Table 7.

Image retrieval is a task investigated by a number of authors (Anavi et al., 2015, 2016; Conjeti et al., 2017; Chen et al., 2018c; Silva et al., 2020; Owais et al., 2020; Haq et al., 2021). The aim of image retrieval tools is to search an image archive to find cases similar to a particular index image. Such algorithms are envisaged as a tool for radiologists in their daily workflow. Chen et al. (2018c) proposed a ranked feature extraction and hashing model, while Silva et al. (2020) proposed to use saliency maps as a similarity measure.

Another task that did not belong to previously defined categories is out-of-distribution detection. Studies working on this (Márquez-Neila and Sznitman, 2019; Çallı et al., 2019; Bozorgtabar et al., 2020) aim to verify whether a test sample belongs to the distribution of the training dataset as model performance is otherwise expected to be sub-optimal. Çallı et al. (2019) propose using the training dataset statistics on different layers of a deep learning model and applying Mahalanobis distance to see the distance of a sample from the training dataset. Bozorgtabar et al. (2020) approach the problem differently and train an unsupervised autoencoder. Later they use the feature encodings extracted from CXRs to define a database of known encodings and compare new samples to this database.

Report generation is another task which has attracted interest in deep learning for CXR (Li et al., 2020b; Yuan et al., 2019; Syeda-Mahmood et al., 2020; Xue et al., 2018a). These studies aim to partially automate the radiology workflow by evaluating the chest X-ray and producing a text radiology report. For example, Syeda-Mahmood et al. (2020) first determines the findings to be reported and then makes use of a large dataset of existing reports to find a similar case. This case report is then customized to produce the final output.

One other task of interest is image registration (Mansilla et al., 2020). This task aims to find the geometric transformation to convert a CXR so that it anatomically aligns with another CXR image or a statistically defined shape. The clinical goal of this task is typically to illustrate interval change between two images. Detecting new findings, tracking the course of a disease, or evaluating the efficacy of a treatment are among the many uses of image registration (Viergever et al., 2016). To that end, Mansilla et al. (2020) aims to create an anatomically plausible registration by using the heart and lung segmentations to guide the registration process.

5. Commercial Products

Computer-aided analysis of CXR images has been researched for many years, and in fact CXR was one of the first modalities for which a commercial product for automatic analyis became available in 2008. In spite of this promising start, and of the advances in the field achieved by deep learning, translation to clinical practice, even as an assistant to the reader, is relatively slow. There are a variety of legal and ethical considerations which may partly account for this (Recht et al., 2020; Strohm et al., 2020), however there is growing acceptance that

Table 7: Other Studies (Section 4.6).
Tasks: IL=Image-level Predictions, IR=Image Retrieval, OD=Out-of-Distribution, RE=Registration, RG=Report Generation, RP=Report Parsing. **Bold font** in tasks implies that this additional task is central to the work and the study also appears in another table in this paper.
Labels: C=ChestX-Ray14, H=Heart, L=Lung, Q=Image Quality, T=Triage/Abnormal, TB=Tuberculosis, X=CheXpert, Z=Other.
Datasets: C=ChestX-ray14, J=JSRT+SCR, M=MIMIC-CXR, MO=Montgomery, O=Open-i, PR=Private, S=Shenzen, X=CheXpert.

Citation	Method	Tasks	Labels	Datasets
Owais et al. (2020)	Uses a database of the intermediate ResNet-50 features to find similar studies	**IL**,IR	TB	MO,S
Syeda-Mahmood et al. (2020)	Generate reports by classifying CXRs, and finding and modifying similar reports	RG,RP	Z	C,M
Li et al. (2020b)	Extracts features from Chest X-rays and uses another network to write reports.	RG,IL	C	C,O
Yuan et al. (2019)	Generates radiology reports by training on classification labels and report text	RG,IL	Z	O,X
Xue et al. (2018a)	A novel recurrent generation network with attention mechanism	RG	Z	O
Mansilla et al. (2020)	Anatomical priors to improve deep learning based image registration	RE	H,L	J,MO,S
Márquez-Neila and Sznitman (2019)	Proposes a method to reject out-of-distribution images during test time	OD,IL	Z	C
Bozorgtabar et al. (2020)	Proposes to detect anomalies based on a dataset of autoencoder features	OD	Q,T	C
Çallı et al. (2019)	Mahalanobis distance on network layers to detect out-of-distribution samples	OD	Z	C
Anavi et al. (2015)	Compares the extracted feature and classification similarities for ranking	IR		PR
Haq et al. (2021)	Uses extracted features to cluster similarly labeled CXRs across datasets	IR	C,X	C,X
Chen et al. (2018c)	Proposes a learnable hash to retrieve CXRs with similar pathologies	IR	Z	C
Conjeti et al. (2017)	Residual network to retrieve images with similar abnormalities	IR	Z	O
Anavi et al. (2016)	Combines features extracted from CXRs and metadata for image retrieval	IR	Z	PR
Silva et al. (2020)	Proposes to use the saliency maps as a similarity measure for image retrieval	IR	Z	X

artificial intelligence (AI) products have a place in the radiological workflow and attempts are underway to understand and address the issues to be overcome (Chokshi et al., 2019). In this section we examine the currently available commercial products for CXR analysis.

An up to date list of commercial products for medical image analysis (Grand-challenge, 2021; van Leeuwen et al., 2021) was searched for products applicable to chest X-ray. One product was excluded as it is not specifically a CXR diagnostic tool, but a texture analysis product for many modalities. The 21 remaining products are listed in Table 8. A number of these products have already been evaluated in peer-reviewed publications, as shown in Table 8 and it is beyond the scope of this work to make an assessment of their performance. All of the listed products are CE marked (Europe) and/or FDA cleared (United States) and are thus available for clinical use (Grand-challenge, 2021; van Leeuwen et al., 2021).

The commercial products include applications for a wide range of abnormalities, with 6 of them reporting results for more than 5 (and up to 30) different labels. The most commonly addressed task is pneumothorax identification (8 products), followed by pleural effusion (7), nodules (6) and tuberculosis (4). In contrast with the literature, which is dominated by image-level prediction algorithms, 17 of 21 products in Table 8 claim to provide localization of one or more abnormalities which they are designed to detect, usually visualized with heatmaps or contouring of abnormalities. Two further products are designed for generation of bone suppression images, one for interval change visualization and one for identification and reporting of healthy images. Products contribute differently to the workflow of the radiologist. Five products focus on detecting acute cases to prioritize the worklist and speed up time to diagnosis. Draft reports are produced by five other products, for either the normal (healthy) cases only or for all cases. The production of draft reports, like workflow prioritization, is aimed at optimizing the speed and efficiency of the radiologist.

6. Discussion

In this work we have detailed datasets, literature and commercial products relevant to deep learning in CXR analysis. It is clear that this area of research has thrived on the release of multiple large, public, labeled datasets in recent years, with 209 of 295 publications reviewed here using one or more public datasets in their research. The number of publications in the field has grown consistently as more public data becomes available, as demonstrated in Figure 2. However, although these datasets are extremely valuable, there are multiple caveats to be considered in relation to their use, as described in Section 3. In particular, the caution required in the use of NLP-extracted labels is often overlooked by researchers, especially for the evaluation and comparison of models. For accurate assessment of model performance, the use of 'gold-standard' test data labels is recommended. These labels can be acquired through expert radiological interpretation of CXRs (preferably with multiple readers) or via associated CT scans, laboratory test results, or other appropriate measurements.

Other important factors to be considered when using public data include the image quality (if it has been reduced prior to release, is this a limiting factor for the application?) and the potential overlap between labels. Although a few publications address label dependencies, this is most often overlooked, frequently resulting in the loss of valuable diagnostic information.

While the increased interest in CXR analysis following the release of public datasets is a positive development in the field, a secondary consequence of this readily available labeled data is the appearance of many publications from researchers with limited experience or understanding of deep learning or CXR analysis. The literature reviewed during the preparation for this paper was very variable in quality. A substantial number of the papers included offer limited novel contributions although they are technically sound. Many of these studies report experiments predicting the labels on public datasets using off-the-shelf architectures and without regard to the label inaccuracies and overlap, or the clinical utility of such generic image-level algorithms. A large number of works were excluded for rea-

Table 8: Commercial Products for CXR analysis. (Section 5)
Labels: T=Triage/Abnormal, PM=Pneumonia, CV=COVID, TB=Tuberculosis, LO=Lesion or Opacity, CM=Cardiomegaly, ND=Nodule, PE=Effusion, PT=Pneumothorax, TU=Catheter or Tube, LC=Lung Cancer, BS=Bone Suppression, E=Edema, Z=Other
Output: LOC=Localization, PRI=Prioritization, REP=Report, SCOR=Scoring

Company	Product	Literature (4 most recent)	Labels (Total number)	Output
Siemens Healthineers	AI-Rad Companion Chest X-Ray	Fischer et al. (2020)	LO PE PT Z (5)	LOC, SCOR, REP
Samsung Healthcare	Auto Lung Nodule Detection	Sim et al. (2019)	ND (1)	LOC
Thirona	CAD4COVID-XRay	Murphy et al. (2020b)	CV (1)	LOC, SCOR
Thirona	CAD4TB	Murphy et al. (2020a); Habib et al. (2020); Qin et al. (2019); Santos et al. (2020)	TB (1)	LOC, SCOR
Oxipit	ChestEye CAD		T (1)	REP (healthy)
Arterys	Chest | MSK AI		LO, ND, PE, PT (4)	LOC, SCOR, PRI
Quibim	Chest X-Ray Classifier	Liang et al. (2020)	PM CM ND PE PT E Z (16)	LOC, SCOR, REP
GE	Critical Care Suite		PT (1)	LOC, SCOR
InferVision	InferRead DR Chest		TB PE PT LC Z (9)	LOC, SCOR
JLK	JLD-O2K		LC Z (16)	LOC, SCOR
Lunit	Lunit INSIGHT CXR	Hwang et al. (2019b,c,a); Qin et al. (2019)	TB CM ND PE PT Z (11)	LOC, SCOR, PRI, REP
qure.ai	qXR	Singh et al. (2018); Nash et al. (2020); Engle et al. (2020); Qin et al. (2019)	T CV TB Z (30)	LOC, SCOR, PRI, REP
Digitec	TIRESYA		BS (1)	Bone Suppressed Image
VUNO	VUNO Med-Chest X-Ray	Kim et al. (2017)	LO ND PE PT Z (5)	LOC, SCOR
Riverain Technologies	ClearRead Xray - Bone Suppress	Homayounieh et al. (2020); Dellios et al. (2017); Schalekamp et al. (2016, 2014b)	BS(1)	Bone Suppressed Image
Riverain Technologies	ClearRead Xray - Compare		LC(1)	Subtraction Image
Riverain Technologies	ClearRead Xray - Confirm		TU(1)	LOC
Riverain Technologies	ClearRead Xray - Detect	Dellios et al. (2017); Schalekamp et al. (2014a); Szucs-Farkas et al. (2013)	ND LC (2)	LOC
behold.ai	Red Dot		T PT (2)	LOC
Zebra Medical Vision	Triage Pleural Effusion		PE	LOC, PRI
Zebra Medical Vision	Triage Pneumothorax		PT	LOC, PRI

sons of poor scientific quality (142). In 112 of these the construction of the dataset gave cause for concern, the most common example being that the training dataset was constructed such that images with certain labels came from different data sources, meaning that the images could be easily differentiated by factors other than the label of interest. In particular, a large number of papers (61) combined adult COVID-19 subjects with pediatric (healthy and other-pneumonia) subjects in an attempt to classify COVID-19. Other reasons for exclusion included the presentation of results optimized on a validation set (without a held-out test set), or the inclusion of the same images multiple times in the dataset prior to splitting train and test sets. This latter issue has been exacerbated by the publication of several COVID-19 related datasets which combine data from multiple public sources in one location, and are then themselves combined by authors building deep-learning systems. Such concerns about dataset construction for COVID-19 studies have been discussed in several other works (López-Cabrera et al., 2021; DeGrave et al., 2020; Cruz et al., 2021; Maguolo and Nanni, 2020; Tartaglione et al., 2020).

Although a broad range of off-the-shelf architectures are employed in the literature surveyed for this review, there is little evidence to suggest that one architecture outperforms another for any specific task. Many papers evaluate multiple different architectures for their task but differences between the various architecture results are typically small, proper hyperparameter optimization is not usually performed and statistical significance or data-selection influence are rarely considered. Many such evaluations use inaccurate NLP-extracted labels for evaluation which serves to muddy the waters even further.

While it is not possible to suggest an optimal architecture for a specific task, it is observed that ensembles of networks typically perform better than individual models (Dietterich, 2000). At the time of writing, most of the top-10 submissions from the public challenges (CheXpert (Irvin et al., 2019), SIIM-ACR (ACR, 2019), and RSNA-Pneumonia (RSNA, 2018)) consist of network ensembles. There is also promise in the development of self-adapting frameworks such as the nnU-Net (Isensee et al., 2021) which has achieved an excellent performance in many medical image segmentation challenges. This framework adapts specifically to the task at hand by selecting the optimal choice for a number of steps such as preprocessing, hyperparameter optimization, architecture etc., and it is likely that a similar optimization framework would perform well for classification or localization tasks, including those for CXR images.

In spite of the pervasiveness of CXR in clinics worldwide, translation of AI systems for clinical use has been relatively slow. Apart from legal and ethical considerations regarding the use of AI in medical decision making (Recht et al., 2020; Strohm et al., 2020), a discussion which is outside the scope of this work, there are still a number of technical hurdles where progress can be made towards the goal of clinical translation. Firstly, the generalizability of AI algorithms is an important issue which needs further work. A large majority of papers in this review draw training, validation and test samples from the same dataset. However, it is well known that such models tend to have a weaker performance on datasets from external domains. If access to reliable data from multiple domains remains problematic then domain adaptation or active learning methods could be considered to address the generalization issue. An alternative method to utilize data from multiple hospitals without breaching regulatory and privacy codes is federated learning, whereby an algorithm can be trained using data from multiple remote locations (Sheller et al., 2019). Further research is required to determine how this type of system will work in clinical practice.

A final issue for deep learning researchers to consider is frequently referred to as 'explainable AI'. Systems which produce classification labels without any indication of reasoning raise concerns of trustworthiness for radiologists. It is also significantly faster for experts to accept or reject the findings of an AI system if there is some indication of how the finding was reached (e.g., identification of nodule location with a bounding box, identification of cardiac and thoracic diameters for cardiomegaly detection). Every commercial product for detection of abnormality in CXR provides a localization feature to indicate the abnormal location, however the literature is heavily focused on image-level predictions with relatively few publications where localization is evaluated.

Beyond the resolution of technical issues, researchers aiming to produce clinically useful systems need to consider the workflow and requirements of the end-user, the radiologist or clinician, more carefully. At present, in the industrialized world, it is expected that an AI system will act, at least initially, as an assistant to (not a replacement for) a radiologist. As a 2D image, the CXR is already relatively quickly interpreted by a radiologist, and so the challenge for AI researchers is to produce systems that will save the radiologist time, prioritize urgent cases or improve the sensitivity/specificity of their findings. Image-level classification for a long list of (somewhat arbitrarily defined) labels is unlikely to be clinically useful. Reviewing such a list of labels and associated probabilities for every CXR would require substantial time and effort, without a proportional improvement in diagnostic accuracy. A simple system with bounding boxes indicating abnormal regions is likely to be more helpful in directing the attention of the radiologist and has the potential to increase sensitivity to subtle findings or in difficult regions with many projected structures. Similarly, a system to quickly identify normal cases has the potential to speed up the workflow as identified by multiple vendors and in the literature (Dyer et al., 2021; Dunnmon et al., 2019; Baltruschat et al., 2020).

To further understand how AI could assist with CXR interpretation, we first must consider the current typical workflow of the radiologist, which notably involves a number of additional inputs beyond the CXR image, that are rarely considered in the research literature. In most scenarios (excluding bedside/AP imaging) both a frontal and lateral CXR are acquired as part of standard imaging protocol, to reduce the interpretation difficulties associated with projected anatomy. Very few studies included in this review made use of the lateral image, although there are indications that it can improve classification accuracy (Hashir et al., 2020). Furthermore, the reviewing radiologist has access to the clinical question being asked, the patient history and symptoms and in many cases other supporting data from blood tests or other investigations. All of this information assists the radiologist to not only identify the visible abnormalities on CXR (e.g., consolidation), but to infer likely causes of these abnormalities (e.g., pneumonia). Incorporation of data from multiple sources along with the CXR image information will almost certainly improve sensitivity and specificity and avoid an algorithm erroneously suggesting labels which are not compatible with data from external sources. Another extremely important and time-consuming element in the radiological review of CXR is comparison with previous images from the same patient, to assess changes over time. Interval change is a topic studied by very few authors and addressed by only a single commercial vendor (by provision of a subtraction image). Innovative AI systems for the visualization and quantification of interval change with one or more previous images could substantially improve the efficiency of the radiologist. Finally, the radiologist is required to produce a report as a result of the CXR review, which is another time-consuming process addressed by very few researchers and just a handful of commercial vendors. A system which can convert radiological findings to a preliminary report has the potential to save time and cost for the care provider.

In many areas of the world, medical facilities that do perform CXR imaging do not have access to radiological expertise. This presents a further opportunity for AI to play a role in diagnostic pathways, as an assistant to the clinician who is not trained in the interpretation of CXR. Researchers and commercial vendors have already identified the need for AI systems to detect signs of tuberculosis (TB), a condition which is endemic in many parts of the world, and frequently in low-resource settings where radiologists are not available. While such regions of the world could potentially benefit from AI systems to detect other conditions, it is important to identify in advance what conditions could be feasibly both detected and treated in these areas where resources are severely limited.

The findings of this work suggest that while the deep learning community has benefited from large numbers of publicly available CXR images, the direction of the research has been largely determined by the available data and labels, rather than the needs of the clinician or radiologist. Future work, in data provision and labelling, and in deep learning, should have a more direct focus on the clinical needs for AI in CXR interpretation. More accurate comparison and benchmarking of algorithms would be enabled by additional public challenges using appropriately annotated data for clinically relevant tasks.

7. Acknowledgements

This work was supported by the Dutch Technology Foundation STW, which formed the NWO Domain Applied and Engineering Sciences and partly funded by the Ministry of Economic Affairs (Perspectief programme P15-26 'DLMedIA: Deep Learning for Medical Image Analysis'.

References

ACR, 2019. SIIM-ACR Pneumothorax Segmentation. URL: https://kaggle.com/c/siim-acr-pneumothorax-segmentation.

Albarqouni, S., Fotouhi, J., Navab, N., 2017. X-Ray In-Depth Decomposition: Revealing the Latent Structures, in: Medical Image Computing and Computer Assisted Intervention - MICCAI 2017. Springer. volume 10435, pp. 444–452. doi:10.1007/978-3-319-66179-7_51.

Amiri, M., Brooks, R., Rivaz, H., 2020. Fine-Tuning U-Net for Ultrasound Image Segmentation: Different Layers, Different Outcomes. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 67, 2510–2518. doi:10.1109/TUFFC.2020.3015081.

Anand, D., Tank, D., Tibrewal, H., Sethi, A., 2020. Self-Supervision vs. Transfer Learning: Robust Biomedical Image Analysis Against Adversarial Attacks, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1159–1163. doi:10.1109/ISBI45749.2020.9098369.

Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H., 2015. A comparative study for chest radiograph image retrieval using binary texture and deep learning classification. International Conference of the IEEE Engineering in Medicine and Biology Society 2015, 2940–2943. doi:10.1109/EMBC.2015.7319008.

Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H., 2016. Visualizing and enhancing a deep learning framework using patients age and gender for chest x-ray image retrieval, in: Medical Imaging 2016: Computer-Aided Diagnosis, SPIE. p. 978510. doi:10.1117/12.2217587.

Anis, S., Lai, K.W., Chuah, J.H., Shoaib, M.A., Mohafez, H., Hadizadeh, M., Ding, Y., Ong, Z.C., 2020. An overview of deep learning approaches in chest radiograph. IEEE Access doi:10.1109/access.2020.3028390.

Annarumma, M., Withey, S.J., Bakewell, R.J., Pesce, E., Goh, V., Montana, G., 2019. Automated Triaging of Adult Chest Radiographs with Deep Artificial Neural Networks. Radiology 291, 272–272. doi:10.1148/radiol.2019194005.

Arbabshirani, M.R., Dallal, A.H., Agarwal, C., Patel, A., Moore, G., 2017. Accurate segmentation of lung fields on chest radiographs using deep convolutional networks, in: Medical Imaging 2017: Image Processing, SPIE. p. 1013305. doi:10.1117/12.2254526.

Arjovsky, M., Chintala, S., Bottou, L., 2017. Wasserstein generative adversarial networks, in: Proceedings of the 34th International Conference on Machine Learning, PMLR. pp. 214–223.

Arsalan, M., Owais, M., Mahmood, T., Choi, J., Park, K.R., 2020. Artificial Intelligence-Based Diagnosis of Cardiac and Related Diseases. Journal of Clinical Medicine 9, 871. doi:10.3390/jcm9030871.

Ayaz, M., Shaukat, F., Raja, G., 2021. Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors. Physical and Engineering Sciences in Medicine doi:10.1007/s13246-020-00966-0.

Balabanova, Y., Coker, R., Fedorin, I., Zakharova, S., Plavinskij, S., Krukov, N., Atun, R., Drobniewski, F., 2005. Variability in interpretation of chest radiographs among russian clinicians and implications for screening programmes: observational study. BMJ 331, 379–382. doi:10.1136/bmj.331.7513.379.

Balachandar, N., Chang, K., Kalpathy-Cramer, J., Rubin, D.L., 2020. Accounting for data variability in multi-institutional distributed deep learning for medical imaging. Journal of the American Medical Informatics Association 27, 700–708. doi:10.1093/jamia/ocaa017.

Baltruschat, I., Steinmeister, L., Nickisch, H., Saalbach, A., Grass, M., Adam, G., Knopp, T., Ittrich, H., 2020. Smart chest X-ray worklist prioritization using artificial intelligence: a clinical workflow simulation. Eur. Radiol. doi:10.1007/s00330-020-07480-7.

Baltruschat, I.M., Nickisch, H., Grass, M., Knopp, T., Saalbach, A., 2019a. Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. Scientific Reports 9, 6381. doi:10.1038/s41598-019-42294-8.

Baltruschat, I.M., Steinmeister, L., Ittrich, H., Adam, G., Nickisch, H., Saalbach, A., von Berg, J., Grass, M., Knopp, T., 2019b. When Does Bone Suppression And Lung Field Segmentation Improve Chest X-Ray Disease Classification?, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 1362–1366. doi:10.1109/ISBI.2019.8759510.

Bar, Y., Diamant, I., Wolf, L., Greenspan, H., 2015a. Deep learning with non-medical training used for chest pathology identification, in: Medical Imaging 2015: Computer-Aided Diagnosis, SPIE. p. 94140V. doi:10.1117/12.2083124.

Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., Greenspan, H., 2015b. Chest pathology detection using deep learning with non-medical training, in: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 294–297. doi:10.1109/ISBI.2015.7163871. iSSN: 1945-8452.

Bayat, A., Sekuboyina, A., Paetzold, J.C., Payer, C., Stern, D., Urschler, M., Kirschke, J.S., Menze, B.H., 2020. Inferring the 3D Standing Spine Posture from 2D Radiographs, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12266, pp. 775–784. doi:10.1007/978-3-030-59725-2_75.

Becker, H.C., Nettleton, W.J., Meyers, P.H., Sweeney, J.W., Nice, C.M., 1964. Digital computer determination of a medical diagnostic index directly from chest x-ray images. IEEE Transactions on Biomedical Engineering BME-11, 67–72. doi:10.1109/tbme.1964.4502309.

Behzadi-khormouji, H., Rostami, H., Salehi, S., Derakhshande-Rishehri, T., Masoumi, M., Salemi, S., Keshavarz, A., Gholamrezanezhad, A., Assadi, M., Batouli, A., 2020. Deep learning, reusable and problem-based architectures for detection of consolidation on chest X-ray images. Computer Methods and Programs in Biomedicine 185, 105162. doi:10.1016/j.cmpb.2019.105162.

von Berg, J., Krönke, S., Gooßen, A., Bystrov, D., Brück, M., Harder, T., Wieberneit, N., Young, S., 2020. Robust chest x-ray quality assessment using convolutional neural networks and atlas regularization, in: Medical Imaging 2020: Image Processing, SPIE. p. 56. doi:10.1117/12.2549541.

Bertrand, H., Hashir, M., Cohen, J.P., 2019. Do lateral views help automated chest x-ray predictions?, in: International Conference on Medical Imaging with Deep Learning – Extended Abstract Track, p. 1.

Bigolin Lanfredi, R., Schroeder, J.D., Vachet, C., Tasdizen, T., 2019. Adversarial Regression Training for Visualizing the Progression of Chronic Obstructive Pulmonary Disease with Chest X-Rays, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 685–693. doi:10.1007/978-3-030-32226-7_76.

Bigolin Lanfredi, R., Schroeder, J.D., Vachet, C., Tasdizen, T., 2020. Interpretation of Disease Evidence for Medical Images Using Adversarial Deformation Fields, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12262, pp. 738–748. doi:10.1007/978-3-030-59713-9_71.

Blain, M., T Kassin, M., Varble, N., Wang, X., Xu, Z., Xu, D., Carrafiello, G., Vespro, V., Stellato, E., Ierardi, A.M., Di Meglio, L., D Suh, R., A Walker, S., Xu, S., H Sanford, T., B Turkbey, E., Harmon, S., Turkbey, B., J Wood, B., 2020. Determination of disease severity in COVID-19 patients using deep learning in chest X-ray images. Diagnostic and Interventional Radiology (Ankara, Turkey) doi:10.5152/dir.2020.20205.

Blumenfeld, A., Greenspan, H., Konen, E., 2018. Pneumothorax detection in chest radiographs using convolutional neural networks, in: Medical Imaging 2018: Computer-Aided Diagnosis, SPIE. p. 3. doi:10.1117/12.2292540.

Bodenreider, O., 2004. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Research 32, D267–D270. doi:10.1093/nar/gkh061.

Bonheur, S., Štern, D., Payer, C., Pienn, M., Olschewski, H., Urschler, M., 2019. Matwo-CapsNet: A Multi-label Semantic Segmentation Capsules Network, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11768, pp. 664–672. doi:10.1007/978-3-030-32254-0_74.

Bortsova, G., Dubost, F., Hogeweg, L., Katramados, I., de Bruijne, M., 2019. Semi-supervised Medical Image Segmentation via Learning Consistency Under Transformations, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 810–818. doi:10.1007/978-3-030-32226-7_90.

Bougias, H., Georgiadou, E., Malamateniou, C., Stogiannos, N., 2020. Identifying cardiomegaly in chest X-rays: a cross-sectional study of evaluation and comparison between different transfer learning methods. Acta Radiol. , 028418512097363doi:10.1177/0284185120973630.

Bozorgtabar, B., Mahapatra, D., Vray, G., Thiran, J.P., 2020. SALAD: Self-supervised Aggregation Learning for Anomaly Detection on X-Rays, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 468–478. doi:10.1007/978-3-030-59710-8_46.

Brestel, C., Shadmi, R., Tamir, I., Cohen-Sfaty, M., Elnekave, E., 2018. RadBot-CXR: Classification of Four Clinical Finding Categories in Chest X-Ray Using Deep Learning, in: International Conference on Medical Imaging with Deep Learning, pp. 1–8.

Burwinkel, H., Kazi, A., Vivar, G., Albarqouni, S., Zahnd, G., Navab, N., Ahmadi, S.A., 2019. Adaptive Image-Feature Learning for Disease Classification Using Inductive Graph Networks, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 640–648. doi:10.1007/978-3-030-32226-7_71.

Bustos, A., Pertusa, A., Salinas, J.M., de la Iglesia-Vayá, M., 2020. PadChest: A large chest x-ray image dataset with multi-label annotated reports. Medical Image Analysis 66, 101797. doi:10.1016/j.media.2020.101797.

Cai, J., Lu, L., Harrison, A.P., Shi, X., Chen, P., Yang, L., 2018. Iterative Attention Mining for Weakly Supervised Thoracic Disease Pattern Localization in Chest X-Rays, in: Medical Image Computing and Computer As-

sisted Intervention – MICCAI 2018. Springer. volume 11071, pp. 589–598. doi:10.1007/978-3-030-00934-2_66.

Çallı, E., Murphy, K., Sogancioglu, E., van Ginneken, B., 2019. {FRODO}: Free rejection of out-of-distribution samples: application to chest x-ray analysis, in: International Conference on Medical Imaging with Deep Learning – Extended Abstract Track, pp. 1–4.

Calli, E., Scholten, E.T., Murphy, K., van Ginneken, B., Sogancioglu, E., 2019. Handling label noise through model confidence and uncertainty: application to chest radiograph classification, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. p. 41. doi:10.1117/12.2514290.

Campo, M.I., Pascau, J., Estepar, R.S.J., 2018. Emphysema quantification on simulated X-rays through deep learning techniques, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE. pp. 273–276. doi:10.1109/ISBI.2018.8363572.

Cardenas, D.A.C., Jr, J.R.F., Moreno, R.A., Rebelo, M.d.F.d.S., Krieger, J.E., Gutierrez, M.A., 2021. Automated radiographic bone suppression with deep convolutional neural networks, in: Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging, International Society for Optics and Photonics. p. 116001D. doi:10.1117/12.2582210.

Castiglioni, I., Ippolito, D., Interlenghi, M., Monti, C.B., Salvatore, C., Schiaffino, S., Polidori, A., Gandola, D., Messa, C., Sardanelli, F., 2021. Machine learning applied on chest x-ray can aid in the diagnosis of COVID-19: a first experience from Lombardy, Italy. European Radiology Experimental 5, 7. doi:10.1186/s41747-020-00203-z.

Cha, M.J., Chung, M.J., Lee, J.H., Lee, K.S., 2019. Performance of deep learning model in detecting operable lung cancer with chest radiographs. Journal of Thoracic Imaging 34, 86–91. doi:10.1097/rti.0000000000000388.

Chakravarty, A., Sarkar, T., Ghosh, N., Sethuraman, R., Sheet, D., 2020. Learning Decision Ensemble using a Graph Neural Network for Comorbidity Aware Chest Radiograph Screening, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE. pp. 1234–1237. doi:10.1109/EMBC44109.2020.9176693.

Chauhan, G., Liao, R., Wells, W., Andreas, J., Wang, X., Berkowitz, S., Horng, S., Szolovits, P., Golland, P., 2020. Joint Modeling of Chest Radiographs and Radiology Reports for Pulmonary Edema Assessment, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12262, pp. 529–539. doi:10.1007/978-3-030-59713-9_51.

Chen, B., Li, J., Lu, G., Yu, H., Zhang, D., 2020a. Label Co-Occurrence Learning With Graph Convolutional Networks for Multi-Label Chest X-Ray Image Classification. IEEE Journal of Biomedical and Health Informatics 24, 2292–2302. doi:10.1109/JBHI.2020.2967084.

Chen, C., Dou, Q., Chen, H., Heng, P.A., 2018a. Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest x-ray segmentation, in: Machine Learning in Medical Imaging. Springer, pp. 143–151. doi:10.1007/978-3-030-00919-9_17.

Chen, H., Miao, S., Xu, D., Hager, G.D., Harrison, A.P., 2019. Deep Hierarchical Multi-label Classification of Chest X-ray Images, in: International Conference on Medical Imaging with Deep Learning, PMLR. pp. 109–120.

Chen, K.C., Yu, H.R., Chen, W.S., Lin, W.C., Lee, Y.C., Chen, H.H., Jiang, J.H., Su, T.Y., Tsai, C.K., Tsai, T.A., Tsai, C.M., Lu, H.H.S., 2020b. Diagnosis of common pulmonary diseases in children by X-ray images and deep learning. Scientific Reports 10, 17374. doi:10.1038/s41598-020-73831-5.

Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018b. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence 40, 834–848. doi:10.1109/tpami.2017.2699184.

Chen, S., Han, Y., Lin, J., Zhao, X., Kong, P., 2020c. Pulmonary nodule detection on chest radiographs using balanced convolutional neural network and classic candidate detection. Artificial Intelligence in Medicine 107, 101881. doi:10.1016/j.artmed.2020.101881.

Chen, X., Duan, Y., Houthooft, R., Schulman, J., Sutskever, I., Abbeel, P., 2016. Infogan: Interpretable representation learning by information maximizing generative adversarial nets, in: Advances in Neural Information Processing Systems, pp. 2172–2180.

Chen, Z., Cai, R., Lu, J., Feng, J., Zhou, J., 2018c. Order-Sensitive Deep Hashing for Multimorbidity Medical Image Retrieval, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11070, pp. 620–628. doi:10.1007/978-3-030-00928-1_70.

Cho, Y., Kim, Y.G., Lee, S.M., Seo, J.B., Kim, N., 2020. Reproducibility of abnormality detection on chest radiographs using convolutional neural network in paired radiographs obtained within a short-term interval. Scientific Reports 10, 17417. doi:10.1038/s41598-020-74626-4.

Chokshi, F.H., Flanders, A.E., Prevedello, L.M., Langlotz, C.P., 2019. Fostering a healthy AI ecosystem for radiology: Conclusions of the 2018 RSNA summit on AI in radiology. Radiology: Artificial Intelligence 1, 190021. doi:10.1148/ryai.2019190021.

Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 1251–1258. doi:10.1109/cvpr.2017.195.

Cicero, M., Bilbily, A., Colak, E., Dowdell, T., Gray, B., Perampaladas, K., Barfett, J., 2017. Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs:. Investigative Radiology 52, 281–287. doi:10.1097/RLI.0000000000000341.

Cohen, J.P., Dao, L., Roth, K., Morrison, P., Bengio, Y., Abbasi, A.F., Shen, B., Mahsa, H.K., Ghassemi, M., Li, H., Duong, T., 2020a. Predicting COVID-19 Pneumonia Severity on Chest X-ray With Deep Learning. Cureus doi:10.7759/cureus.9448.

Cohen, J.P., Hashir, M., Brooks, R., Bertrand, H., 2020b. On the limits of cross-domain generalization in automated x-ray prediction. Proceedings of the Third Conference on Medical Imaging with Deep Learning, PMLR , 121:136–155arXiv:2002.02497.

Cohen, J.P., Morrison, P., Dao, L., 2020c. Covid-19 image data collection: Prospective predictions are the future. arXiv preprint arXiv:2006.11988 .

Conjeti, S., Roy, A.G., Katouzian, A., Navab, N., 2017. Hashing with Residual Networks for Image Retrieval, in: Medical Image Computing and Computer Assisted Intervention - MICCAI 2017. Springer. volume 10435, pp. 541–549. doi:10.1007/978-3-319-66179-7_62.

Crosby, J., Chen, S., Li, F., MacMahon, H., Giger, M., 2020a. Network output visualization to uncover limitations of deep learning detection of pneumothorax, in: Medical Imaging 2020: Image Perception, Observer Performance, and Technology Assessment, SPIE. p. 22. doi:10.1117/12.2550066.

Crosby, J., Rhines, T., Duan, C., Li, F., MacMahon, H., Giger, M., 2019. Impact of imprinted labels on deep learning classification of AP and PA thoracic radiographs, in: Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications, SPIE. p. 13. doi:10.1117/12.2513026.

Crosby, J., Rhines, T., Li, F., MacMahon, H., Giger, M., 2020b. Deep convolutional neural networks in the classification of dual-energy thoracic radiographic views for efficient workflow: analysis on over 6500 clinical radiographs. Journal of Medical Imaging 7, 1. doi:10.1117/1.JMI.7.1.016501.

Crosby, J., Rhines, T., Li, F., MacMahon, H., Giger, M., 2020c. Deep learning for pneumothorax detection and localization using networks fine-tuned with multiple institutional datasets, in: Medical Imaging 2020: Computer-Aided Diagnosis, SPIE. p. 11. doi:10.1117/12.2549709.

Cruz, B.G.S., Bossa, M.N., Sölter, J., Husch, A.D., 2021. Public covid-19 x-ray datasets and their impact on model bias - a systematic review of a significant problem. medRxiv doi:10.1101/2021.02.15.21251775.

Daniels, Z.A., Metaxas, D.N., 2019. Exploiting Visual and Report-Based Information for Chest X-RAY Analysis by Jointly Learning Visual Classifiers and Topic Models, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 1270–1274. doi:10.1109/ISBI.2019.8759548.

DeGrave, A.J., Janizek, J.D., Lee, S.I., 2020. AI for radiographic COVID-19 detection selects shortcuts over signal. medRxiv doi:10.1101/2020.09.13.20193565.

Dellios, N., Teichgraeber, U., Chelaru, R., Malich, A., Papageorgiou, I.E., 2017. Computer-aided Detection Fidelity of Pulmonary Nodules in Chest Radiograph. Journal of Clinical Imaging Science 7. doi:10.4103/jcis.JCIS_75_16.

Demner-Fushman, D., Antani, S., Simpson, M., Thoma, G.R., 2012. Design and Development of a Multimodal Biomedical Information Retrieval System. Journal of Computing Science and Engineering 6, 168–177. doi:10.5626/JCSE.2012.6.2.168.

Deng, J., Dong, W., Socher, R., Li, L., Kai Li, Li Fei-Fei, 2009. ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. doi:10.1109/CVPR.2009.5206848. iSSN: 1063-6919.

Deshpande, H., Harder, T., Saalbach, A., Sawarkar, A., Buelow, T., 2020. Detection Of Foreign Objects In Chest Radiographs Using Deep Learning, in: 2020 IEEE 17th International Symposium on Biomedical

Imaging Workshops (ISBI Workshops), IEEE. pp. 1–4. doi:10.1109/ISBIWorkshops50223.2020.9153350.

Devnath, L., Luo, S., Summons, P., Wang, D., 2021. Automated detection of pneumoconiosis with multilevel deep features learned from chest X-Ray radiographs. Comput. Biol. Med. 129, 104125. doi:10.1016/j.compbiomed.2020.104125.

Dietterich, T.G., 2000. Ensemble Methods in Machine Learning, in: Multiple Classifier Systems, Springer. pp. 1–15. doi:10.1007/3-540-45014-9_1.

Dong, N., Kampffmeyer, M., Liang, X., Wang, Z., Dai, W., Xing, E., 2018. Unsupervised Domain Adaptation for Automatic Estimation of Cardiothoracic Ratio, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11071, pp. 544–552. doi:10.1007/978-3-030-00934-2_61.

Dong, N., Xu, M., Liang, X., Jiang, Y., Dai, W., Xing, E., 2019. Neural Architecture Search for Adversarial Medical Image Segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 828–836. doi:10.1007/978-3-030-32226-7_92.

DSouza, A.M., Abidin, A.Z., Wismüller, A., 2019. Automated identification of thoracic pathology from chest radiographs with enhanced training pipeline, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. p. 123. doi:10.1117/12.2512600.

Dunnmon, J.A., Yi, D., Langlotz, C.P., Ré, C., Rubin, D.L., Lungren, M.P., 2019. Assessment of Convolutional Neural Networks for Automated Classification of Chest Radiographs. Radiology 290, 537–544. doi:10.1148/radiol.2018181422.

Dyer, T., Dillard, L., Harrison, M., Morgan, T.N., Tappouni, R., Malik, Q., Rasalingham, S., 2021. Diagnosis of normal chest radiographs using an autonomous deep-learning algorithm. Clin. Radiol. , S0009926021000763doi:10.1016/j.crad.2021.01.015.

E, L., Zhao, B., Guo, Y., Zheng, C., Zhang, M., Lin, J., Luo, Y., Cai, Y., Song, X., Liang, H., 2019. Using deep-learning techniques for pulmonary-thoracic segmentations and improvement of pneumonia diagnosis in pediatric chest radiographs. Pediatric Pulmonology 54, 1617–1626. doi:10.1002/ppul.24431.

Ellis, R., Ellestad, E., Elicker, B., Hope, M.D., Tosun, D., 2020. Impact of hybrid supervision approaches on the performance of artificial intelligence for the classification of chest radiographs. Computers in Biology and Medicine 120, 103699. doi:10.1016/j.compbiomed.2020.103699.

Elshennawy, N.M., Ibrahim, D.M., 2020. Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images. Diagnostics 10, 649. doi:10.3390/diagnostics10090649.

Engle, E., Gabrielian, A., Long, A., Hurt, D.E., Rosenthal, A., 2020. Performance of Qure.ai automatic classifiers against a large annotated database of patients with diverse forms of tuberculosis. PLOS ONE 15, e0224445. doi:10.1371/journal.pone.0224445.

Eslami, M., Tabarestani, S., Albarqouni, S., Adeli, E., Navab, N., Adjouadi, M., 2020. Image-to-images translation for multi-task organ segmentation and bone suppression in chest x-ray radiography. IEEE Transactions on Medical Imaging 39, 1–1. doi:10.1109/tmi.2020.2974159.

Fang, Q., Yan, J., Gu, X., Zhao, J., Li, Q., 2020. Unsupervised learning-based deformable registration of temporal chest radiographs to detect interval change, in: Medical Imaging 2020: Image Processing, SPIE. p. 104. doi:10.1117/12.2549211.

Feng, Y., Teh, H.S., Cai, Y., 2019. Deep learning for chest radiology: A review. Current Radiology Reports 7. doi:10.1007/s40134-019-0333-9.

Ferreira, J.R., Armando Cardona Cardenas, D., Moreno, R.A., de Fatima de Sa Rebelo, M., Krieger, J.E., Antonio Gutierrez, M., 2020. Multi-View Ensemble Convolutional Neural Network to Improve Classification of Pneumonia in Low Contrast Chest X-Ray Images, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE. pp. 1238–1241. doi:10.1109/EMBC44109.2020.9176517.

Fischer, A.M., Varga-Szemes, A., Martin, S.S., Sperl, J.I., Sahbaee, P., Neumann, D., Gawlitza, J., Henzler, T., Johnson, C.M., Nance, J.W., Schoenberg, S.O., Schoepf, U.J., 2020. Artificial Intelligence-based Fully Automated Per Lobe Segmentation and Emphysema-quantification Based on Chest Computed Tomography Compared With Global Initiative for Chronic Obstructive Lung Disease Severity of Smokers. Journal of Thoracic Imaging 35 Suppl 1, S28–S34. doi:10.1097/RTI.0000000000000500.

Fricks, R.B., Abadi, E., Ria, F., Samei, E., 2021. Classification of COVID-19 in chest radiographs: assessing the impact of imaging parameters using clinical and simulated images, in: Medical Imaging 2021: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 115970A. doi:10.1117/12.2582223.

Frid-Adar, M., Amer, R., Greenspan, H., 2019. Endotracheal Tube Detection and Segmentation in Chest Radiographs Using Synthetic Data, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 784–792. doi:10.1007/978-3-030-32226-7_87.

Fukushima, K., Miyake, S., 1982. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition, in: Competition and Cooperation in Neural Nets, Springer. pp. 267–285. doi:10.1007/978-3-642-46466-9_18.

Furutani, K., Hirano, Y., Kido, S., 2019. Segmentation of lung region from chest x-ray images using U-net, in: International Forum on Medical Imaging in Asia 2019, SPIE. p. 48. doi:10.1117/12.2521594.

Ganesan, P., Rajaraman, S., Long, R., Ghoraani, B., Antani, S., 2019. Assessment of Data Augmentation Strategies Toward Performance Improvement of Abnormality Classification in Chest Radiographs, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 841–844. doi:10.1109/EMBC.2019.8857516.

Ghesu, F.C., Georgescu, B., Gibson, E., Guendel, S., Kalra, M.K., Singh, R., Digumarthy, S.R., Grbic, S., Comaniciu, D., 2019. Quantifying and Leveraging Classification Uncertainty for Chest Radiograph Assessment, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 676–684. doi:10.1007/978-3-030-32226-7_75.

van Ginneken, B., 2017. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. Radiological Physics and Technology 10, 23–32. doi:10.1007/s12194-017-0394-5.

van Ginneken, B., Stegmann, M.B., Loog, M., 2006. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. Medical Image Analysis 10, 19–40. doi:10.1016/j.media.2005.02.002.

Girshick, R., 2015. Fast r-CNN, in: 2015 IEEE International Conference on Computer Vision (ICCV), IEEE. pp. 1440–1448. doi:10.1109/iccv.2015.169.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition, IEEE. pp. 580–587. doi:10.1109/cvpr.2014.81.

Gomi, T., Hara, H., Watanabe, Y., Mizukami, S., 2020. Improved digital chest tomosynthesis image quality by use of a projection-based dual-energy virtual monochromatic convolutional neural network with super resolution. PLoS One 15, e0244745. doi:10.1371/journal.pone.0244745.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets, in: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, MIT Press. pp. 2672–2680.

Gozes, O., Greenspan, H., 2019. Deep Feature Learning from a Hospital-Scale Chest X-ray Dataset with Application to TB Detection on a Small-Scale Dataset, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 4076–4079. doi:10.1109/EMBC.2019.8856729.

Gozes, O., Greenspan, H., 2020. Bone Structures Extraction and Enhancement in Chest Radiographs via CNN Trained on Synthetic Data, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 858–861. doi:10.1109/ISBI45749.2020.9098738.

Grand-challenge, 2021. Grand challenge: Ai for radiology. https://grand-challenge.org/aiforradiology/.

Griner, D., Zhang, R., Tie, X., Zhang, C., Garrett, J.W., Li, K., Chen, G.H., 2021. COVID-19 pneumonia diagnosis using chest x-ray radiograph and deep learning, in: Medical Imaging 2021: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 1159706. doi:10.1117/12.2581972.

Groza, V., Kuzin, A., 2020. Pneumothorax Segmentation with Effective Conditioned Post-Processing in Chest X-Ray, in: 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops), IEEE. pp. 1–4. doi:10.1109/ISBIWorkshops50223.2020.9153444.

Gyawali, P.K., Ghimire, S., Bajracharya, P., Li, Z., Wang, L., 2020. Semi-supervised Medical Image Classification with Global Latent Mixing, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 604–613. doi:10.1007/

978-3-030-59710-8_59.

Gyawali, P.K., Li, Z., Ghimire, S., Wang, L., 2019. Semi-supervised Learning by Disentangling and Self-ensembling over Stochastic Latent Space, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 766–774. doi:10.1007/978-3-030-32226-7_85.

Habib, S.S., Rafiq, S., Zaidi, S.M.A., Ferrand, R.A., Creswell, J., Van Ginneken, B., Jamal, W.Z., Azeemi, K.S., Khowaja, S., Khan, A., 2020. Evaluation of computer aided detection of tuberculosis on chest radiography among people with diabetes in Karachi Pakistan. Scientific Reports 10, 6276. doi:10.1038/s41598-020-63084-7.

Haghighi, F., Hosseinzadeh Taher, M.R., Zhou, Z., Gotway, M.B., Liang, J., 2020. Learning Semantics-Enriched Representation via Self-discovery, Self-classification, and Self-restoration, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 137–147. doi:10.1007/978-3-030-59710-8_14.

Haq, N.F., Moradi, M., Wang, Z.J., 2021. A deep community based approach for large scale content based X-ray image retrieval. Med. Image Anal. 68, 101847. doi:10.1016/j.media.2020.101847.

Hashir, M., Bertrand, H., Cohen, J.P., 2020. Quantifying the value of lateral views in deep learning for chest x-rays, in: Proceedings of the Third Conference on Medical Imaging with Deep Learning, PMLR. pp. 288–303.

He, K., Gkioxari, G., Dollar, P., Girshick, R., 2017. Mask r-CNN, in: 2017 IEEE International Conference on Computer Vision (ICCV), IEEE. pp. 2961–2969. doi:10.1109/iccv.2017.322.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. doi:10.1109/CVPR.2016.90.

Heo, S.J., Kim, Y., Yun, S., Lim, S.S., Kim, J., Nam, C.M., Park, E.C., Jung, I., Yoon, J.H., 2019. Deep Learning Algorithms with Demographic Information Help to Detect Tuberculosis in Chest Radiographs in Annual Workers' Health Examination Data. International Journal of Environmental Research and Public Health 16, 250. doi:10.3390/ijerph16020250.

Hermoza, R., Maicas, G., Nascimento, J.C., Carneiro, G., 2020. Region Proposals for Saliency Map Refinement for Weakly-Supervised Disease Localisation and Classification, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12266, pp. 539–549. doi:10.1007/978-3-030-59725-2_52.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S., 2017. Gans trained by a two time-scale update rule converge to a local nash equilibrium, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, p. 6629–6640.

Hirata, Y., Kusunose, K., Tsuji, T., Fujimori, K., Kotoku, J., Sata, M., 2021. Deep Learning for Detection of Elevated Pulmonary Artery Wedge Pressure using Standard Chest X-Ray. Can. J. Cardiol. , S0828282X21001094doi:10.1016/j.cjca.2021.02.007.

HMHospitales, 2020. Hmhospitales.

Holste, G., Sullivan, R.P., Bindschadler, M., Nagy, N., Alessio, A., 2020. Multiclass semantic segmentation of pediatric chest radiographs, in: Medical Imaging 2020: Image Processing, SPIE. p. 49. doi:10.1117/12.2544426.

Homayounieh, F., Digumarthy, S.R., Febbo, J.A., Garrana, S., Nitiwarangkul, C., Singh, R., Khera, R.D., Gilman, M., Kalra, M.K., 2020. Comparison of Baseline, Bone-Subtracted, and Enhanced Chest Radiographs for Detection of Pneumothorax. Canadian Association of Radiologists Journal = Journal l'Association Canadienne Des Radiologistes , 846537120908852doi:10.1177/0846537120908852.

Hosch, R., Kroll, L., Nensa, F., Koitka, S., 2020. Differentiation Between Anteroposterior and Posteroanterior Chest X-Ray View Position With Convolutional Neural Networks. RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren , a–1183–5227doi:10.1055/a-1183-5227.

Hu, Q., Drukker, K., Giger, M.L., 2021. Role of standard and soft tissue chest radiography images in COVID-19 diagnosis using deep learning, in: Medical Imaging 2021: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 1159704. doi:10.1117/12.2581977.

Huang, G., Liu, Z., v. d. Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269. doi:10.1109/CVPR.2017.243.

Huang, X., Liu, M.Y., Belongie, S., Kautz, J., 2018. Multimodal unsupervised image-to-image translation, in: Computer Vision – ECCV 2018. Springer, pp. 179–196. doi:10.1007/978-3-030-01219-9_11.

Hurt, B., Yen, A., Kligerman, S., Hsiao, A., 2020. Augmenting Interpretation of Chest Radiographs With Deep Learning Probability Maps. Journal of Thoracic Imaging 35, 285–293. doi:10.1097/RTI.0000000000000505.

Hwang, E.J., Nam, J.G., Lim, W.H., Park, S.J., Jeong, Y.S., Kang, J.H., Hong, E.K., Kim, T.M., Goo, J.M., Park, S., Kim, K.H., Park, C.M., 2019a. Deep learning for chest radiograph diagnosis in the emergency department. Radiology 293, 573–580. doi:10.1148/radiol.2019191225.

Hwang, E.J., Park, S., Jin, K.N., Kim, J.I., Choi, S.Y., Lee, J.H., Goo, J.M., Aum, J., Yim, J.J., Cohen, J.G., Ferretti, G.R., and, C.M.P., 2019b. Development and validation of a deep learning–based automated detection algorithm for major thoracic diseases on chest radiographs. JAMA Network Open 2, e191095. doi:10.1001/jamanetworkopen.2019.1095.

Hwang, E.J., Park, S., Jin, K.N., Kim, J.I., Choi, S.Y., Lee, J.H., Goo, J.M., Aum, J., Yim, J.J., Park, C.M., Deep Learning-Based Automatic Detection Algorithm Development and Evaluation Group, Kim, D.H., Woo, W., Choi, C., Hwang, I.P., Song, Y.S., Lim, L., Kim, K., Wi, J.Y., Oh, S.S., Kang, M.J., 2019c. Development and Validation of a Deep Learning–based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. Clinical Infectious Diseases 69, 739–747. doi:10.1093/cid/ciy967.

Hwang, S., Kim, H.E., 2016. Self-Transfer Learning for Weakly Supervised Lesion Localization, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Springer. volume 9901, pp. 239–246. doi:10.1007/978-3-319-46723-8_28.

Hwang, S., Kim, H.E., Jeong, J., Kim, H.J., 2016. A novel approach for tuberculosis screening based on deep convolutional neural networks, in: Medical Imaging 2016: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 97852W. doi:10.1117/12.2216198.

Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R.L., Shpanskaya, K.S., Seekins, J., Mong, D.A., Halabi, S.S., Sandberg, J.K., Jones, R., Larson, D.B., Langlotz, C.P., Patel, B.N., Lungren, M.P., Ng, A.Y., 2019. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison, in: AAAI Conference on Artificial Intelligence, pp. 590–597.

Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H., 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods 18, 203–211. doi:10.1038/s41592-020-01008-z.

Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 1125–1134. doi:10.1109/cvpr.2017.632.

Jaeger, S., Candemir, S., Antani, S., Wáng, Y.X.J., Lu, P.X., Thoma, G., 2014. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quantitative Imaging in Medicine and Surgery 4, 475–477. doi:10.3978/j.issn.2223-4292.2014.11.20.

Jang, R., Kim, N., Jang, M., Lee, K.H., Lee, S.M., Lee, K.H., Noh, H.N., Seo, J.B., 2020. Assessment of the Robustness of Convolutional Neural Networks in Labeling Noise by Using Chest X-Ray Images From Multiple Centers. JMIR Medical Informatics 8, e18089. doi:10.2196/18089.

Johnson, A.E.W., Pollard, T.J., Berkowitz, S.J., Greenbaum, N.R., Lungren, M.P., ying Deng, C., Mark, R.G., Horng, S., 2019. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. Scientific Data 6. doi:10.1038/s41597-019-0322-0.

Junior, J.R.F., Cardenas, D.A.C., Moreno, R.A., Rebelo, M.d.F.d.S., Krieger, J.E., Gutierrez, M.A., 2021. A general fully automated deep-learning method to detect cardiomegaly in chest x-rays, in: Medical Imaging 2021: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 115972B. doi:10.1117/12.2581980.

Kallianos, K., Mongan, J., Antani, S., Henry, T., Taylor, A., Abuya, J., Kohli, M., 2019. How far have we come? artificial intelligence for chest radiograph interpretation. Clinical Radiology 74, 338–345. doi:10.1016/j.crad.2018.12.015.

Karargyris, A., Kashyap, S., Wu, J.T., Sharma, A., Moradi, M., Syeda-Mahmood, T., 2019a. Age prediction using a large chest x-ray dataset, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. p. 66. doi:10.1117/12.2512922.

Karargyris, A., Wong, K.C.L., Wu, J.T., Moradi, M., Syeda-Mahmood, T., 2019b. Boosting the Rule-Out Accuracy of Deep Disease Detection Using Class Weight Modifiers, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 877–881. doi:10.1109/ISBI.2019.8759532.

Karras, T., Aila, T., Laine, S., Lehtinen, J., 2018. Progressive growing of GANs

for improved quality, stability, and variation, in: International Conference on Learning Representations, pp. 1–8.

Kashyap, S., Karargyris, A., Wu, J., Gur, Y., Sharma, A., Wong, K.C.L., Moradi, M., Syeda-Mahmood, T., 2020. Looking in the Right Place for Anomalies: Explainable Ai Through Automatic Location Learning, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1125–1129. doi:10.1109/ISBI45749.2020.9098370.

Kermany, D., 2018. Large dataset of labeled optical coherence tomography (oct) and chest x-ray images. doi:10.17632/RSCBJBR9SJ.3.

Khakzar, A., Albarqouni, S., Navab, N., 2019. Learning Interpretable Features via Adversarially Robust Optimization, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 793–800. doi:10.1007/978-3-030-32226-7_88.

Khatibi, T., Shahsavari, A., Farahani, A., 2021. Proposing a novel multi-instance learning model for tuberculosis recognition from chest X-ray images based on CNNs, complex networks and stacked ensemble. Physical and Engineering Sciences in Medicine doi:10.1007/s13246-021-00980-w.

Kholiavchenko, M., Sirazitdinov, I., Kubrak, K., Badrutdinova, R., Kuleev, R., Yuan, Y., Vrtovec, T., Ibragimov, B., 2020. Contour-aware multi-label chest X-ray organ segmentation. International Journal of Computer Assisted Radiology and Surgery 15, 425–436. doi:10.1007/s11548-019-02115-9.

Kim, H.E., Kim, S., Lee, J., 2018. Keep and Learn: Continual Learning by Constraining the Latent Space for Knowledge Preservation in Neural Networks, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11070, pp. 520–528. doi:10.1007/978-3-030-00928-1_59.

Kim, J.R., Shim, W.H., Yoon, H.M., Hong, S.H., Lee, J.S., Cho, Y.A., Kim, S., 2017. Computerized Bone Age Estimation Using Deep Learning Based Program: Evaluation of the Accuracy and Efficiency. AJR. American journal of roentgenology 209, 1374–1380. doi:10.2214/AJR.17.18224.

Kim, M., Lee, B.D., 2021. Automatic Lung Segmentation on Chest X-rays Using Self-Attention Deep Neural Network. Sensors 21, 369. doi:10.3390/s21020369.

Kim, Y.G., Cho, Y., Wu, C.J., Park, S., Jung, K.H., Seo, J.B., Lee, H.J., Hwang, H.J., Lee, S.M., Kim, N., 2019. Short-term Reproducibility of Pulmonary Nodule and Mass Detection in Chest Radiographs: Comparison among Radiologists and Four Different Computer-Aided Detections with Convolutional Neural Net. Scientific Reports 9, 18738. doi:10.1038/s41598-019-55373-7.

Kim, Y.G., Lee, S.M., Lee, K.H., Jang, R., Seo, J.B., Kim, N., 2020. Optimal matrix size of chest radiographs for computer-aided detection on lung nodule or mass with deep learning. European Radiology 30, 4943–4951. doi:10.1007/s00330-020-06892-9.

Kitahara, Y., Tanaka, R., Roth, H., Oda, H., Mori, K., Kasahara, K., Matsumoto, I., 2019. Lung segmentation based on a deep learning approach for dynamic chest radiography, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. p. 130. doi:10.1117/12.2512711.

Kitamura, G., Deible, C., 2020. Retraining an open-source pneumothorax detecting machine learning algorithm for improved performance to medical images. Clinical Imaging 61, 15–19. doi:10.1016/j.clinimag.2020.01.008.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks.

Kruger, R.P., Townes, J.R., Hall, D.L., Dwyer, S.J., Lodwick, G.S., 1972. Automated radiographic diagnosis via feature extraction and classification of cardiac size and shape descriptors. IEEE Transactions on Biomedical Engineering BME-19, 174–186. doi:10.1109/tbme.1972.324115.

Kuo, P.C., Tsai, C.C., López, D.M., Karargyris, A., Pollard, T.J., Johnson, A.E.W., Celi, L.A., 2021. Recalibration of deep learning models for abnormality detection in smartphone-captured chest radiograph. npj Digital Medicine 4, 25. doi:10.1038/s41746-021-00393-9.

Kurmann, T., Márquez-Neila, P., Wolf, S., Sznitman, R., 2019. Deep Multi-label Classification in Affine Subspaces, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11764, pp. 165–173. doi:10.1007/978-3-030-32239-7_19.

Kusakunniran, W., Karnjanapreechakorn, S., Siriapisith, T., Borwarnginn, P., Sutassananon, K., Tongdee, T., Saiviroonporn, P., 2021. COVID-19 detection and heatmap generation in chest x-ray images. Journal of Medical Imaging 8, 014001. doi:10.1117/1.JMI.8.S1.014001.

Kusunose, K., Hirata, Y., Tsuji, T., Kotoku, J., Sata, M., 2020. Deep learning to predict elevated pulmonary artery pressure in patients with suspected pulmonary hypertension using standard chest X ray. Sci. Rep. 10, 19311. doi:10.1038/s41598-020-76359-w.

Lakhani, P., 2017. Deep Convolutional Neural Networks for Endotracheal Tube Position and X-ray Image Classification: Challenges and Opportunities. Journal of Digital Imaging 30, 460–468. doi:10.1007/s10278-017-9980-7.

Lakhani, P., Sundaram, B., 2017. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. Radiology 284, 574–582. doi:10.1148/radiol.2017162326.

Larrazabal, A.J., Martinez, C., Glocker, B., Ferrante, E., 2020. Post-DAE: Anatomically Plausible Segmentation via Post-Processing With Denoising Autoencoders. IEEE Transactions on Medical Imaging 39, 3813–3820. doi:10.1109/TMI.2020.3005297.

Laserson, J., Lantsman, C.D., Cohen-Sfady, M., Tamir, I., Goz, E., Brestel, C., Bar, S., Atar, M., Elnekave, E., 2018. TextRay: Mining Clinical Reports to Gain a Broad Understanding of Chest X-Rays, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11071, pp. 553–561. doi:10.1007/978-3-030-00934-2_62.

LeCun, Y., Bengio, Y., 1998. Convolutional networks for images, speech, and time series, in: The handbook of brain theory and neural networks. MIT Press, pp. 255–258.

Lee, D., Kim, H., Choi, B., Kim, H.J., 2019. Development of a deep neural network for generating synthetic dual-energy chest x-ray images with single x-ray exposure. Physics in Medicine & Biology 64, 115017. doi:10.1088/1361-6560/ab1cee.

Lee, H., Mansouri, M., Tajmir, S., Lev, M.H., Do, S., 2018. A Deep-Learning System for Fully-Automated Peripherally Inserted Central Catheter (PICC) Tip Detection. Journal of Digital Imaging 31, 393–402. doi:10.1007/s10278-017-0025-z.

van Leeuwen, K.G., Schalekamp, S., Rutten, M.J., van Ginneken, B., de Rooij, M., 2021. Artificial intelligence in radiology; 100 commercially available products and their scientific evidence, 2021. European Radiology (in press) .

Lenga, M., Schulz, H., Saalbach, A., 2020. Continual Learning for Domain Adaptation in Chest X-ray Classification. Proceedings of the Third Conference on Medical Imaging with Deep Learning, PMLR , 121:413–423arXiv:2001.05922.

Lenis, D., Major, D., Wimmer, M., Berg, A., Sluiter, G., Bühler, K., 2020. Domain Aware Medical Image Classifier Interpretation by Counterfactual Impact Analysis, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 315–325. doi:10.1007/978-3-030-59710-8_31.

Li, B., Kang, G., Cheng, K., Zhang, N., 2019. Attention-Guided Convolutional Neural Network for Detecting Pneumonia on Chest X-Rays, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 4851–4854. doi:10.1109/EMBC.2019.8857277.

Li, F., Shi, J.X., Yan, L., Wang, Y.G., Zhang, X.D., Jiang, M.S., Wu, Z.Z., Zhou, K.Q., 2021a. Lesion-aware convolutional neural network for chest radiograph classification. Clin. Radiol. 76, 155.e1–155.e14. doi:10.1016/j.crad.2020.08.027.

Li, M.D., Arun, N.T., Gidwani, M., Chang, K., Deng, F., Little, B.P., Mendoza, D.P., Lang, M., Lee, S.I., O'Shea, A., Parakh, A., Singh, P., Kalpathy-Cramer, J., 2020a. Automated assessment of COVID-19 pulmonary disease severity on chest radiographs using convolutional siamese neural networks. medRxiv doi:10.1101/2020.05.20.20108159.

Li, X., Cao, R., Zhu, D., 2020b. Vispi: Automatic visual perception and interpretation of chest x-rays, in: Medical Imaging with Deep Learning, pp. 1–8.

Li, X., Shen, L., Xie, X., Huang, S., Xie, Z., Hong, X., Yu, J., 2020c. Multi-resolution convolutional networks for chest X-ray radiograph based lung nodule detection. Artificial Intelligence in Medicine 103, 101744. doi:10.1016/j.artmed.2019.101744.

Li, X., Zhu, D., 2020. Robust Detection of Adversarial Attacks on Medical Images, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1154–1158. doi:10.1109/ISBI45749.2020.9098628.

Li, Y., Dong, X., Shi, W., Miao, Y., Yang, H., Jiang, Z., 2021b. Lung fields segmentation in chest radiographs using Dense-U-Net and fully connected CRF, in: Twelfth International Conference on Graphics and Image Processing (ICGIP 2020), International Society for Optics and Photonics. p. 1172011. doi:10.1117/12.2589384.

Li, Z., Hou, Z., Chen, C., Hao, Z., An, Y., Liang, S., Lu, B., 2019. Automatic

cardiothoracic ratio calculation with deep learning. IEEE Access 7, 37749–37756. doi:10.1109/ACCESS.2019.2900053.

Li, Z., Li, H., Han, H., Shi, G., Wang, J., Zhou, S.K., 2019. Encoding CT Anatomy Knowledge for Unpaired Chest X-ray Image Decomposition, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 275–283. doi:10.1007/978-3-030-32226-7_31.

Liang, C.H., Liu, Y.C., Wu, M.T., Garcia-Castro, F., Alberich-Bayarri, A., Wu, F.Z., 2020. Identifying pulmonary nodules or masses on chest radiography using deep learning: external validation and strategies to improve clinical practice. Clinical Radiology 75, 38–45. doi:10.1016/j.crad.2019.08.005.

Liang, G., Zheng, L., 2020. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Computer Methods and Programs in Biomedicine 187, 104964. doi:10.1016/j.cmpb.2019.06.023.

Lin, C., Tang, R., Lin, D.D., Liu, L., Lu, J., Chen, Y., Gao, D., Zhou, J., 2020. Deep Feature Disentanglement Learning for Bone Suppression in Chest Radiographs, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 795–798. doi:10.1109/ISBI45749.2020.9098399.

Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P., 2017. Focal loss for dense object detection, in: 2017 IEEE International Conference on Computer Vision (ICCV), IEEE. pp. 2980–2988. doi:10.1109/iccv.2017.324.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical Image Analysis 42, 60–88. doi:10.1016/j.media.2017.07.005.

Liu, H., Wang, L., Nan, Y., Jin, F., Wang, Q., Pu, J., 2019. SDFN: Segmentation-based deep fusion network for thoracic disease classification in chest X-ray images. Computerized Medical Imaging and Graphics 75, 66–73. doi:10.1016/j.compmedimag.2019.05.005.

Liu, X., Wang, S., Deng, Y., Chen, K., 2017. Coronary artery calcification (CAC) classification with deep convolutional neural networks, in: Medical Imaging 2017: Computer-Aided Diagnosis, SPIE. p. 101340M. doi:10.1117/12.2253974.

Liu, Y., Liu, M., Xi, Y., Qin, G., Shen, D., Yang, W., 2020a. Generating Dual-Energy Subtraction Soft-Tissue Images from Chest Radiographs via Bone Edge-Guided GAN, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12262, pp. 678–687. doi:10.1007/978-3-030-59713-9_65.

Liu, Y.C., Lin, Y.C., Tsai, P.Y., Iwata, O., Chuang, C.C., Huang, Y.H., Tsai, Y.S., Sun, Y.N., 2020b. Convolutional Neural Network-Based Humerus Segmentation and Application to Bone Mineral Density Estimation from Chest X-ray Images of Critical Infants. Diagnostics 10, 1028. doi:10.3390/diagnostics10121028.

Lodwick, G.S., Keats, T.E., Dorst, J.P., 1963. The coding of roentgen images for computer analysis as applied to lung cancer. Radiology 81, 185–200. doi:10.1148/81.2.185.

Longjiang, E., Baisong Zhao, Liu, H., Zheng, C., Song, X., Cai, Y., Liang, H., 2020. Image-based Deep Learning in Diagnosing the Etiology of Pneumonia on Pediatric Chest X-rays. Pediatr. Pulmonol. , ppul.25229doi:10.1002/ppul.25229.

Lu, M.T., Ivanov, A., Mayrhofer, T., Hosny, A., Aerts, H.J.W.L., Hoffmann, U., 2019. Deep Learning to Assess Long-term Mortality From Chest Radiographs. JAMA Network Open 2, e197416. doi:10.1001/jamanetworkopen.2019.7416.

Lu, M.T., Raghu, V.K., Mayrhofer, T., Aerts, H.J., Hoffmann, U., 2020a. Deep Learning Using Chest Radiographs to Identify High-Risk Smokers for Lung Cancer Screening Computed Tomography: Development and Validation of a Prediction Model. Annals of Internal Medicine 173, 704–713. doi:10.7326/M20-1868.

Lu, Y., Li, W., Zheng, K., Wang, Y., Harrison, A.P., Lin, C., Wang, S., Xiao, J., Lu, L., Kuo, C.F., Miao, S., 2020b. Learning to Segment Anatomical Structures Accurately from One Exemplar, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 678–688. doi:10.1007/978-3-030-59710-8_66.

Luo, L., Yu, L., Chen, H., Liu, Q., Wang, X., Xu, J., Heng, P.A., 2020. Deep Mining External Imperfect Data for Chest X-Ray Disease Screening. IEEE Transactions on Medical Imaging 39, 3583–3594. doi:10.1109/TMI.2020.3000949.

López-Cabrera, J.D., Orozco-Morales, R., Portal-Diaz, J.A., Lovelle-Enríquez, O., Pérez-Díaz, M., 2021. Current limitations to identify COVID-19 using artificial intelligence with chest X-ray imaging. Health and Technology 11, 411–424. doi:10.1007/s12553-021-00520-2.

M. S., V., V., M.K., Gonnabhaktula, A., Kundeti, S.R., J., V., 2019. Local and global transformations to improve learning of medical images applied to chest radiographs, in: Medical Imaging 2019: Image Processing, SPIE. p. 114. doi:10.1117/12.2512717.

Ma, C., Wang, H., Hoi, S.C.H., 2019. Multi-label Thoracic Disease Image Classification with Cross-Attention Networks, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 730–738. doi:10.1007/978-3-030-32226-7_81.

Madani, A., Moradi, M., Karargyris, A., Syeda-Mahmood, T., 2018. Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE. pp. 1038–1042. doi:10.1109/ISBI.2018.8363749.

Mader, A.O., von Berg, J., Fabritz, A., Lorenz, C., Meyer, C., 2018. Localization and Labeling of Posterior Ribs in Chest Radiographs Using a CRF-regularized FCN with Local Refinement, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11071, pp. 562–570. doi:10.1007/978-3-030-00934-2_63.

Maguolo, G., Nanni, L., 2020. A critic evaluation of methods for covid-19 automatic detection from x-ray images. arXiv preprint arXiv:2004.12823 .

Mahapatra, D., Ge, Z., 2019. Training Data Independent Image Registration with Gans Using Transfer Learning and Segmentation Information, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 709–713. doi:10.1109/ISBI.2019.8759247.

Majkowska, A., Mittal, S., Steiner, D.F., Reicher, J.J., McKinney, S.M., Duggan, G.E., Eswaran, K., Cameron Chen, P.H., Liu, Y., Kalidindi, S.R., Ding, A., Corrado, G.S., Tse, D., Shetty, S., 2019. Chest Radiograph Interpretation with Deep Learning Models: Assessment with Radiologist-adjudicated Reference Standards and Population-adjusted Evaluation. Radiology 294, 421–431. doi:10.1148/radiol.2019191293. publisher: Radiological Society of North America.

Mansilla, L., Milone, D.H., Ferrante, E., 2020. Learning deformable registration of medical images with anatomical constraints. Neural Networks 124, 269–279. doi:10.1016/j.neunet.2020.01.023.

Mansoor, A., Cerrolaza, J.J., Perez, G., Biggs, E., Okada, K., Nino, G., Linguraru, M.G., 2020. A Generic Approach to Lung Field Segmentation From Chest Radiographs Using Deep Space and Shape Learning. IEEE Transactions on Biomedical Engineering 67, 1206–1220. doi:10.1109/TBME.2019.2933508.

Mansoor, A., Perez, G., Nino, G., Linguraru, M.G., 2016. Automatic tissue characterization of air trapping in chest radiographs using deep neural networks, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 97–100. doi:10.1109/EMBC.2016.7590649.

Mao, C., Yao, L., Pan, Y., Luo, Y., Zeng, Z., 2018. Deep Generative Classifiers for Thoracic Disease Diagnosis with Chest X-ray Images, in: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE. pp. 1209–1214. doi:10.1109/BIBM.2018.8621107.

Mao, Y., Xue, F.F., Wang, R., Zhang, J., Zheng, W.S., Liu, H., 2020. Abnormality Detection in Chest X-Ray Images Using Uncertainty Prediction Autoencoders, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12266, pp. 529–538. doi:10.1007/978-3-030-59725-2_51.

Mathai, T.S., Gorantla, V., Galeotti, J., 2019. Segmentation of Vessels in Ultra High Frequency Ultrasound Sequences Using Contextual Memory, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11765, pp. 173–181. doi:10.1007/978-3-030-32245-8_20.

Matsubara, N., Teramoto, A., Saito, K., Fujita, H., 2020. Bone suppression for chest X-ray image using a convolutional neural filter. Physical and Engineering Sciences in Medicine 43, 97–108. doi:10.1007/s13246-019-00822-w.

Matsumoto, T., Kodera, S., Shinohara, H., Ieki, H., Yamaguchi, T., Higashikuni, Y., Kiyosue, A., Ito, K., Ando, J., Takimoto, E., Akazawa, H., Morita, H., Komuro, I., 2020. Diagnosing Heart Failure from Chest X-Ray Images Using Deep Learning. International Heart Journal 61, 781–786. doi:10.1536/ihj.19-714.

McDonald, C.J., Overhage, J.M., Barnes, M., Schadow, G., Blevins, L., Dexter, P.R., Mamlin, B., 2005. The Indiana Network For Patient Care: A Working Local Health Information Infrastructure. Health Affairs 24, 1214–1220.

doi:10.1377/hlthaff.24.5.1214. publisher: Health Affairs.

McManigle, J.E., Bartz, R.R., Carin, L., 2020. Y-Net for Chest X-Ray Preprocessing: Simultaneous Classification of Geometry and Segmentation of Annotations, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE. pp. 1266–1269. doi:10.1109/EMBC44109.2020.9176334.

Mettler, F.A., Bhargavan, M., Faulkner, K., Gilley, D.B., Gray, J.E., Ibbott, G.S., Lipoti, J.A., Mahesh, M., McCrohan, J.L., Stabin, M.G., Thomadsen, B.R., Yoshizumi, T.T., 2009. Radiologic and nuclear medicine studies in the united states and worldwide: Frequency, radiation dose, and comparison with other radiation sources—1950–2007. Radiology 253, 520–531. doi:10.1148/radiol.2532082010.

Meyers, P.H., Nice, C.M., Becker, H.C., Nettleton, W.J., Sweeney, J.W., Meckstroth, G.R., 1964. Automated computer analysis of radiographic images. Radiology 83, 1029–1034. doi:10.1148/83.6.1029.

Michael, P., Yoon, H.J., 2020. Survey of image denoising methods for medical image classification, in: Medical Imaging 2020: Computer-Aided Diagnosis, SPIE. p. 132. doi:10.1117/12.2549695.

Milletari, F., Rieke, N., Baust, M., Esposito, M., Navab, N., 2018. CFCM: Segmentation via Coarse to Fine Context Memory, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11073, pp. 667–674. doi:10.1007/978-3-030-00937-3_76.

Mitra, A., Chakravarty, A., Ghosh, N., Sarkar, T., Sethuraman, R., Sheet, D., 2020. A Systematic Search over Deep Convolutional Neural Network Architectures for Screening Chest Radiographs, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE. pp. 1225–1228. doi:10.1109/EMBC44109.2020.9175246.

Mittal, A., Kumar, D., Mittal, M., Saba, T., Abunadi, I., Rehman, A., Roy, S., 2020. Detecting Pneumonia Using Convolutions and Dynamic Capsule Routing for Chest X-ray Images. Sensors 20, 1068. doi:10.3390/s20041068.

Moradi, M., Madani, A., Gur, Y., Guo, Y., Syeda-Mahmood, T., 2018a. Bimodal Network Architectures for Automatic Generation of Image Annotation from Text, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11070, pp. 449–456. doi:10.1007/978-3-030-00928-1_51.

Moradi, M., Madani, A., Karargyris, A., Syeda-Mahmood, T.F., 2018b. Chest x-ray generation and data augmentation for cardiovascular abnormality classification, in: Medical Imaging 2018: Image Processing, SPIE. p. 57. doi:10.1117/12.2293971.

Moradi, M., Siegel, E., Kashyap, S., Karargyris, A., Wu, J.T., Saboury, B., Morris, M., Syeda-Mahmood, T., 2019a. Artificial intelligence for point of care radiograph quality assessment, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. p. 128. doi:10.1117/12.2513092.

Moradi, M., Wong, K.C.L., Syeda-Mahmood, T., Wu, J.T., 2019b. Identifying disease-free chest x-ray images with deep transfer learning, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. p. 24. doi:10.1117/12.2513164.

Moradi, M., Wong, K.L., Karargyris, A., Syeda-Mahmood, T., 2020. Quality controlled segmentation to aid disease detection, in: Medical Imaging 2020: Computer-Aided Diagnosis, SPIE. p. 138. doi:10.1117/12.2549426.

Mortani Barbosa, E.J., Gefter, W.B., Ghesu, F.C., Liu, S., Mailhe, B., Mansoor, A., Grbic, S., Vogt, S., 2021. Automated Detection and Quantification of COVID-19 Airspace Disease on Chest Radiographs: A Novel Approach Achieving Expert Radiologist-Level Performance Using a Deep Convolutional Neural Network Trained on Digital Reconstructed Radiographs From Computed Tomography–Derived Ground Truth. Invest. Radiol. Publish Ahead of Print. doi:10.1097/RLI.0000000000000763.

Murphy, K., Habib, S.S., Zaidi, S.M.A., Khowaja, S., Khan, A., Melendez, J., Scholten, E.T., Amad, F., Schalekamp, S., Verhagen, M., Philipsen, R.H.H.M., Meijers, A., van Ginneken, B., 2020a. Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system. Scientific Reports 10, 5492. doi:10.1038/s41598-020-62148-y.

Murphy, K., Smits, H., Knoops, A.J.G., Korst, M.B.J.M., Samson, T., Scholten, E.T., Schalekamp, S., Schaefer-Prokop, C.M., Phi lipsen, R.H.H.M., Meijers, A., Melendez, J., van Ginneken, B., Rutten, M., 2020b. COVID-19 on Chest Radiographs: A Multireader Evaluation of an Artificial Intelligence System. Radiology 296, E166–E172. doi:10.1148/radiol.2020201874.

Márquez-Neila, P., Sznitman, R., 2019. Image Data Validation for Medical Systems, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11767, pp. 329–337. doi:10.1007/978-3-030-32251-9_36.

Nakao, T., Hanaoka, S., Nomura, Y., Murata, M., Takenaga, T., Miki, S., Watadani, T., Yoshikawa, T., Hayashi, N., Abe, O., 2021. Unsupervised Deep Anomaly Detection in Chest Radiographs. J. Digit. Imaging doi:10.1007/s10278-020-00413-2.

Nam, J.G., Kim, M., Park, J., Hwang, E.J., Lee, J.H., Hong, J.H., Goo, J.M., Park, C.M., 2020. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. Eur. Respir. J. , 2003061doi:10.1183/13993003.03061-2020.

Nam, J.G., Park, S., Hwang, E.J., Lee, J.H., Jin, K.N., Lim, K.Y., Vu, T.H., Sohn, J.H., Hwang, S., Goo, J.M., Park, C.M., 2019. Development and Validation of Deep Learning–based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. Radiology 290, 218–228. doi:10.1148/radiol.2018180237.

Narayanan, B.N., Davuluru, V.S.P., Hardie, R.C., 2020. Two-stage deep learning architecture for pneumonia detection and its diagnosis in chest radiographs, in: Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications, SPIE. p. 15. doi:10.1117/12.2547635.

Nash, M., Kadavigere, R., Andrade, J., Sukumar, C.A., Chawla, K., Shenoy, V.P., Pande, T., Huddart, S., Pai, M., Saravu, K., 2020. Deep learning, computer-aided radiography reading for tuberculosis: a diagnostic accuracy study from a tertiary hospital in India. Scientific Reports 10, 210. doi:10.1038/s41598-019-56589-3.

National Lung Screening Trial Research Team, N., Aberle, D.R., Adams, A.M., Berg, C.D., Black, W.C., Clapp, J.D., Fagerstrom, R.M., Gareen, I.F., Gatsonis, C., Marcus, P.M., Sicks, J.D., 2011. Reduced lung-cancer mortality with low-dose computed tomographic screening. The New England Journal of Medicine 365, 395–409. doi:10.1056/NEJMoa1102873.

Novikov, A.A., Lenis, D., Major, D., Hladuvka, J., Wimmer, M., Buhler, K., 2018. Fully convolutional architectures for multiclass segmentation in chest radiographs. IEEE Transactions on Medical Imaging 37, 1865–1876. doi:10.1109/tmi.2018.2806086.

Nugroho, B.A., 2021. An aggregate method for thorax diseases classification. Sci. Rep. 11, 3242. doi:10.1038/s41598-021-81765-9.

Oakden-Rayner, L., 2019. Half a million x-rays! first impressions of the stanford and mit chest x-ray datasets. https://lukeoakdenrayner.wordpress.com/2019/02/25/half-a-million-x-rays-first-impressions-of-the-stanford-and-mit-ch

Oakden-Rayner, L., 2020. Exploring large-scale public medical image datasets. Academic Radiology 27, 106–112. doi:10.1016/j.acra.2019.10.006.

Odena, A., Olah, C., Shlens, J., 2017. Conditional image synthesis with auxiliary classifier GANs, in: Proceedings of the 34th International Conference on Machine Learning, pp. 2642–2651.

Ogawa, R., Kido, T., Kido, T., Mochizuki, T., 2019. Effect of augmented datasets on deep convolutional neural networks applied to chest radiographs. Clinical Radiology 74, 697–701. doi:10.1016/j.crad.2019.04.025.

Oh, D.Y., Kim, J., Lee, K.J., 2019. Longitudinal Change Detection on Chest X-rays Using Geometric Correlation Maps, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 748–756. doi:10.1007/978-3-030-32226-7_83.

Oh, Y., Park, S., Ye, J.C., 2020. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. IEEE Transactions on Medical Imaging 39, 2688–2700. doi:10.1109/TMI.2020.2993291.

Olatunji, T., Yao, L., Covington, B., Upton, A., 2019. Caveats in generating medical imaging labels from radiology reports with natural language processing, in: International Conference on Medical Imaging with Deep Learning – Extended Abstract Track, pp. 1–4.

Oliveira, H., Mota, V., Machado, A.M., dos Santos, J.A., 2020a. From 3d to 2d: Transferring knowledge for rib segmentation in chest x-rays. Pattern Recognition Letters 140, 10–17. doi:10.1016/j.patrec.2020.09.021.

Oliveira, H., dos Santos, J., 2018. Deep transfer learning for segmentation of anatomical structures in chest radiographs, in: 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), IEEE. pp. 204–211. doi:10.1109/sibgrapi.2018.00033.

Oliveira, H.N., Ferreira, E., Santos, J.A.D., 2020b. Truly generalizable radiograph segmentation with conditional domain adaptation. IEEE Access 8, 84037–84062. doi:10.1109/access.2020.2991688.

Onodera, S., Lee, Y., Tanaka, Y., 2020. Evaluation of dose reduction potential in scatter-corrected bedside chest radiography using U-net. Radiological Physics and Technology 13, 336–347. doi:10.1007/s12194-020-00586-z.

Ouyang, X., Karanam, S., Wu, Z., Chen, T., Huo, J., Zhou, X.S., Wang, Q.,

Cheng, J.Z., 2020. Learning Hierarchical Attention for Weakly-supervised Chest X-Ray Abnormality Localization and Diagnosis. IEEE Transactions on Medical Imaging , 1–1doi:10.1109/TMI.2020.3042773.

Ouyang, X., Xue, Z., Zhan, Y., Zhou, X.S., Wang, Q., Zhou, Y., Wang, Q., Cheng, J.Z., 2019. Weakly Supervised Segmentation Framework with Uncertainty: A Study on Pneumothorax Segmentation in Chest X-ray, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 613–621. doi:10.1007/978-3-030-32226-7_68.

Owais, M., Arsalan, M., Mahmood, T., Kim, Y.H., Park, K.R., 2020. Comprehensive Computer-Aided Decision Support Framework to Diagnose Tuberculosis From Chest X-Ray Images: Data Mining Study. JMIR Medical Informatics 8, e21790. doi:10.2196/21790.

Pan, I., Agarwal, S., Merck, D., 2019a. Generalizable Inter-Institutional Classification of Abnormal Chest Radiographs Using Efficient Convolutional Neural Networks. Journal of Digital Imaging 32, 888–896. doi:10.1007/s10278-019-00180-9.

Pan, Y., Chen, Q., Chen, T., Wang, H., Zhu, X., Fang, Z., Lu, Y., 2019b. Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays. European Spine Journal 28, 3035–3043. doi:10.1007/s00586-019-06115-w.

Park, S., Lee, S.M., Kim, N., Choe, J., Cho, Y., Do, K.H., Seo, J.B., 2019. Application of deep learning–based computer-aided detection system: detecting pneumothorax on chest radiograph after biopsy. European Radiology 29, 5341–5348. doi:10.1007/s00330-019-06130-x.

Park, S., Lee, S.M., Lee, K.H., Jung, K.H., Bae, W., Choe, J., Seo, J.B., 2020. Deep learning-based detection system for multiclass lesions on chest radiographs: comparison with observer readings. European Radiology 30, 1359–1368. doi:10.1007/s00330-019-06532-x.

Pasa, F., Golkov, V., Pfeiffer, F., Cremers, D., Pfeiffer, D., 2019. Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization. Scientific Reports 9, 6268. doi:10.1038/s41598-019-42557-4.

Paul, A., Shen, T.C., Lee, S., Balachandar, N., Peng, Y., Lu, Z., Summers, R.M., 2021a. Generalized Zero-shot Chest X-ray Diagnosis through Trait-Guided Multi-view Semantic Embedding with Self-training. IEEE Transactions on Medical Imaging , 1–1doi:10.1109/TMI.2021.3054817.

Paul, A., Tang, Y.X., Shen, T.C., Summers, R.M., 2021b. Discriminative ensemble learning for few-shot chest x-ray diagnosis. Med. Image Anal. 68, 101911. doi:10.1016/j.media.2020.101911.

Paul, A., Tang, Y.X., Summers, R.M., 2020. Fast few-shot transfer learning for disease identification from chest x-ray images using autoencoder ensemble, in: Medical Imaging 2020: Computer-Aided Diagnosis, SPIE. p. 6. doi:10.1117/12.2549060.

Pesce, E., Joseph Withey, S., Ypsilantis, P.P., Bakewell, R., Goh, V., Montana, G., 2019. Learning to detect chest radiographs containing pulmonary lesions using visual attention networks. Medical Image Analysis 53, 26–38. doi:10.1016/j.media.2018.12.007.

Pham, H.H., Le, T.T., Ngo, D.T., Tran, D.Q., Nguyen, H.Q., 2020. Interpreting chest x-rays via {cnn}s that exploit hierarchical disease dependencies and uncertainty labels, in: Medical Imaging with Deep Learning, pp. 1–8.

Portela, R.D.S., Pereira, J.R.G., Costa, M.G.F., Filho, C.F.F.C., 2020. Lung Region Segmentation in Chest X-Ray Images using Deep Convolutional Neural Networks, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE. pp. 1246–1249. doi:10.1109/EMBC44109.2020.9175478.

Prevedello, L.M., Halabi, S.S., Shih, G., Wu, C.C., Kohli, M.D., Chokshi, F.H., Erickson, B.J., Kalpathy-Cramer, J., Andriole, K.P., Flanders, A.E., 2019. Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions. Radiology: Artificial Intelligence 1, e180031. doi:10.1148/ryai.2019180031.

Qin, C., Yao, D., Shi, Y., Song, Z., 2018. Computer-aided detection in chest radiography based on artificial intelligence: a survey. BioMedical Engineering OnLine 17. doi:10.1186/s12938-018-0544-y.

Qin, Z.Z., Sander, M.S., Rai, B., Titahong, C.N., Sudrungrot, S., Laah, S.N., Adhikari, L.M., Carter, E.J., Puri, L., Codlin, A.J., Creswell, J., 2019. Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. Scientific Reports 9, 15000. doi:10.1038/s41598-019-51503-3.

Qu, W., Balki, I., Mendez, M., Valen, J., Levman, J., Tyrrell, P.N., 2020. Assessing and mitigating the effects of class imbalance in machine learning with application to X-ray imaging. International Journal of Computer Assisted Radiology and Surgery 15, 2041–2048. doi:10.1007/s11548-020-02260-6.

Que, Q., Tang, Z., Wang, R., Zeng, Z., Wang, J., Chua, M., Gee, T.S., Yang, X., Veeravalli, B., 2018. CardioXNet: Automated detection for cardiomegaly based on deep learning, in: International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE. pp. 612–615. doi:10.1109/embc.2018.8512374.

Quekel, L.G., Kessels, A.G., Goei, R., van Engelshoven, J.M., 2001. Detection of lung cancer on the chest radiograph: a study on observer performance. European Journal of Radiology 39, 111–116. doi:10.1016/s0720-048x(01)00301-1.

Rahman, M.F., Tseng, T.L.B., Pokojovy, M., Qian, W., Totada, B., Xu, H., 2021. An automatic approach to lung region segmentation in chest x-ray images using adapted U-Net architecture, in: Medical Imaging 2021: Physics of Medical Imaging, International Society for Optics and Photonics. p. 115953I. doi:10.1117/12.2581882.

Rajan, D., Thiagarajan, J.J., Karargyris, A., Kashyap, S., 2021. Self-training with improved regularization for sample-efficient chest x-ray classification, in: Medical Imaging 2021: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 115971S. doi:10.1117/12.2582290.

Rajaraman, S., Antani, S.K., 2020. Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs. IEEE Access 8, 27318–27326. doi:10.1109/ACCESS.2020.2971257.

Rajaraman, S., Candemir, S., Kim, I., Thoma, G., Antani, S., 2018a. Visualization and Interpretation of Convolutional Neural Network Predictions in Detecting Pneumonia in Pediatric Chest Radiographs. Applied Sciences 8, 1715. doi:10.3390/app8101715.

Rajaraman, S., Candemir, S., Xue, Z., Alderson, P.O., Kohli, M., Abuya, J., Thoma, G.R., Antani, S., 2018b. A novel stacked generalization of models for improved TB detection in chest radiographs, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 718–721. doi:10.1109/EMBC.2018.8512337.

Rajaraman, S., Kim, I., Antani, S.K., 2020a. Detection and visualization of abnormality in chest radiographs using modality-specific convolutional neural network ensembles. PeerJ 8, e8693. doi:10.7717/peerj.8693.

Rajaraman, S., Sornapudi, S., Alderson, P.O., Folio, L.R., Antani, S.K., 2020b. Analyzing inter-reader variability affecting deep ensemble learning for COVID-19 detection in chest radiographs. PLoS One 15, e0242301. doi:10.1371/journal.pone.0242301.

Rajaraman, S., Sornapudi, S., Kohli, M., Antani, S., 2019a. Assessment of an ensemble of machine learning models toward abnormality detection in chest radiographs, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 3689–3692. doi:10.1109/EMBC.2019.8856715.

Rajaraman, S., Thoma, G., Antani, S.., Candemir, S., 2019b. Visualizing and explaining deep learning predictions for pneumonia detection in pediatric chest radiographs, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. p. 27. doi:10.1117/12.2512752.

Rajkomar, A., Lingam, S., Taylor, A.G., Blum, M., Mongan, J., 2017. High-Throughput Classification of Radiographs Using Deep Convolutional Neural Networks. Journal of Digital Imaging 30, 95–101. doi:10.1007/s10278-016-9914-9.

Rajpurkar, P., Irvin, J., Ball, R.L., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C.P., Patel, B.N., Yeom, K.W., Shpanskaya, K., Blankenberg, F.G., Seekins, J., Amrhein, T.J., Mong, D.A., Halabi, S.S., Zucker, E.J., Ng, A.Y., Lungren, M.P., 2018. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLOS Medicine 15, e1002686. doi:10.1371/journal.pmed.1002686.

Rajpurkar, P., O'Connell, C., Schechter, A., Asnani, N., Li, J., Kiani, A., Ball, R.L., Mendelson, M., Maartens, G., van Hoving, D.J., Griesel, R., Ng, A.Y., Boyles, T.H., Lungren, M.P., 2020. CheXaid: deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV. npj Digital Medicine 3, 115. doi:10.1038/s41746-020-00322-2.

Raoof, S., Feigin, D., Sung, A., Raoof, S., Irugulpati, L., Rosenow, E.C., 2012. Interpretation of plain chest roentgenogram. Chest 141, 545–558. doi:10.1378/chest.10-1302.

Ravishankar, H., Venkataramani, R., Anamandra, S., Sudhakar, P., Annangi, P., 2019. Feature Transformers: Privacy Preserving Lifelong Learners for Medical Imaging, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11767, pp. 347–355. doi:10.1007/978-3-030-32251-9_38.

Recht, M.P., Dewey, M., Dreyer, K., Langlotz, C., Niessen, W., Prainsack, B., Smith, J.J., 2020. Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations. European Radiology 30, 3576–3584. doi:10.1007/s00330-020-06672-5.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 779–788. doi:10.1109/cvpr.2016.91.

Redmon, J., Farhadi, A., 2017. YOLO9000: Better, faster, stronger, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 7263–7271. doi:10.1109/cvpr.2017.690.

Redmon, J., Farhadi, A., 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767 .

Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster r-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 1137–1149. doi:10.1109/tpami.2016.2577031.

Rolnick, D., Veit, A., Belongie, S., Shavit, N., 2018. Deep Learning is Robust to Massive Label Noise. arXiv:1705.10694 [cs] ArXiv: 1705.10694.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 234–241.

RSNA, 2018. RSNA Pneumonia Detection Challenge. Library Catalog: www.kaggle.com.

Rueckel, J., Kunz, W.G., Hoppe, B.F., Patzig, M., Notohamiprodjo, M., Meinel, F.G., Cyran, C.C., Ingrisch, M., Ricke, J., Sabel, B.O., 2020. Artificial Intelligence Algorithm Detecting Lung Infection in Supine Chest Radiographs of Critically Ill Patients With a Diagnostic Accuracy Similar to Board-Certified Radiologists. Critical Care Medicine Publish Ahead of Print. doi:10.1097/CCM.0000000000004397.

Sabottke, C.F., Breaux, M.A., Spieler, B.M., 2020. Estimation of age in unidentified patients via chest radiography using convolutional neural network regression. Emergency Radiology 27, 463–468. doi:10.1007/s10140-020-01782-5.

Saednia, K., Jalalifar, A., Ebrahimi, S., Sadeghi-Naini, A., 2020. An Attention-Guided Deep Neural Network for Annotating Abnormalities in Chest X-ray Images: Visualization of Network Decision Basis *, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE. pp. 1258–1261. doi:10.1109/EMBC44109.2020.9175378.

Sahiner, B., Pezeshk, A., Hadjiiski, L.M., Wang, X., Drukker, K., Cha, K.H., Summers, R.M., Giger, M.L., 2018. Deep learning in medical imaging and radiation therapy. Medical Physics 46, e1–e36. doi:10.1002/mp.13264.

Salehinejad, H., Colak, E., Dowdell, T., Barfett, J., Valaee, S., 2019. Synthesizing Chest X-Ray Pathology for Training Deep Convolutional Neural Networks. IEEE Transactions on Medical Imaging 38, 1197–1206. doi:10.1109/TMI.2018.2881415.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X., 2016. Improved techniques for training gans, in: Proceedings of the 30th International Conference on Neural Information Processing Systems, p. 2234–2242.

Samala, R.K., Hadjiiski, L., Chan, H.P., Zhou, C., Stojanovska, J., Agarwal, P., Fung, C., 2021. Severity assessment of COVID-19 using imaging descriptors: a deep-learning transfer learning approach from non-COVID-19 pneumonia, in: Medical Imaging 2021: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 115971T. doi:10.1117/12.2582115.

Santos, A.d.S., Oliveira, R.D.d., Lemos, E.F., Lima, F., Cohen, T., Cords, O., Martinez, L., Gonçalves, C., Ko, A., Andrews, J.R., Croda, J., 2020. Yield, Efficiency and Costs of Mass Screening Algorithms for Tuberculosis in Brazilian Prisons. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America doi:10.1093/cid/ciaa135.

Sathitratanacheewin, S., Sunanta, P., Pongpirul, K., 2020. Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. Heliyon 6, e04614. doi:10.1016/j.heliyon.2020.e04614.

Schalekamp, S., van Ginneken, B., Koedam, E., Snoeren, M.M., Tiehuis, A.M., Wittenberg, R., Karssemeijer, N., Schaefer-Prokop, C.M., 2014a. Computer-aided detection improves detection of pulmonary nodules in chest radiographs beyond the support by bone-suppressed images. Radiology 272, 252–261. doi:10.1148/radiol.14131315.

Schalekamp, S., Ginneken, B.v., Berk, I.A.H.v.d., Hartmann, I.J.C., Snoeren, M.M., Odink, A.E., Lankeren, W.v., Pegge, S.A.H., Schijf, L.J., Karssemeijer, N., Schaefer-Prokop, C.M., 2014b. Bone Suppression Increases the Visibility of Invasive Pulmonary Aspergillosis in Chest Radiographs. PLOS ONE 9, e108551. doi:10.1371/journal.pone.0108551.

Schalekamp, S., Karssemeijer, N., Cats, A.M., De Hoop, B., Geurts, B.H.J., Berger-Hartog, O., van Ginneken, B., Schaefer-Prokop, C.M., 2016. The Effect of Supplementary Bone-Suppressed Chest Radiographs on the Assessment of a Variety of Common Pulmonary Abnormalities: Results of an Observer Study. Journal of Thoracic Imaging 31, 119–125. doi:10.1097/RTI.0000000000000195.

Schroeder, J.D., Bigolin Lanfredi, R., Li, T., Chan, J., Vachet, C., Paine, R., Srikumar, V., Tasdizen, T., 2021. Prediction of Obstructive Lung Disease from Chest Radiographs via Deep Learning Trained on Pulmonary Function Data. International Journal of Chronic Obstructive Pulmonary Disease Volume 15, 3455–3466. doi:10.2147/COPD.S279850.

Schultheiss, M., Schober, S.A., Lodde, M., Bodden, J., Aichele, J., Müller-Leisse, C., Renger, B., Pfeiffer, F., Pfeiffer, D., 2020. A robust convolutional neural network for lung nodule detection in the presence of foreign bodies. Scientific Reports 10, 12987. doi:10.1038/s41598-020-69789-z.

Schwab, E., Goossen, A., Deshpande, H., Saalbach, A., 2020. Localization of Critical Findings in Chest X-Ray Without Local Annotations Using Multi-Instance Learning, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1879–1882. doi:10.1109/ISBI45749.2020.9098551.

Seah, J.C.Y., Tang, J.S.N., Kitchen, A., Gaillard, F., Dixon, A.F., 2019. Chest radiographs in congestive heart failure: Visualizing neural network learning. Radiology 290, 514–522. doi:10.1148/radiol.2018180887.

Shah, M.P., Merchant, S.N., Awate, S.P., 2018. MS-Net: Mixed-Supervision Fully-Convolutional Networks for Full-Resolution Segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11073, pp. 379–387. doi:10.1007/978-3-030-00937-3_44.

Shah, U., Abd-Alrazeq, A., Alam, T., Househ, M., Shah, Z., 2020. An Efficient Method to Predict Pneumonia from Chest X-Rays Using Deep Learning Approach. Studies in Health Technology and Informatics 272, 457–460. doi:10.3233/SHTI200594.

Shelhamer, E., Long, J., Darrell, T., 2017. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 640–651. doi:10.1109/tpami.2016.2572683.

Sheller, M.J., Reina, G.A., Edwards, B., Martin, J., Bakas, S., 2019. Multi-institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation, in: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Springer. pp. 92–104. doi:10.1007/978-3-030-11723-8_9.

Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K.i., Matsui, M., Fujita, H., Kodera, Y., Doi, K., 2000. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. American Journal of Roentgenology 174, 71–74. doi:10.2214/ajr.174.1.1740071.

Silva, W., Poellinger, A., Cardoso, J.S., Reyes, M., 2020. Interpretability-Guided Content-Based Medical Image Retrieval, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 305–314. doi:10.1007/978-3-030-59710-8_30.

Sim, Y., Chung, M.J., Kotter, E., Yune, S., Kim, M., Do, S., Han, K., Kim, H., Yang, S., Lee, D.J., Choi, B.W., 2019. Deep Convolutional Neural Network–based Software Improves Radiologist Detection of Malignant Lung Nodules on Chest Radiographs. Radiology 294, 199–209. doi:10.1148/radiol.2019182465.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 .

Singh, R., Kalra, M.K., Nitiwarangkul, C., Patti, J.A., Homayounieh, F., Padole, A., Rao, P., Putha, P., Muse, V.V., Sharma, A., Digumarthy, S.R., 2018. Deep learning in chest radiography: Detection of findings and presence of change. PloS One 13, e0204155. doi:10.1371/journal.pone.0204155.

Singh, V., Danda, V., Gorniak, R., Flanders, A., Lakhani, P., 2019. Assessment of Critical Feeding Tube Malpositions on Radiographs Using Deep Learning. Journal of Digital Imaging 32, 651–655. doi:10.1007/s10278-019-00229-9.

Sirazitdinov, I., Kholiavchenko, M., Kuleev, R., Ibragimov, B., 2019. Data Augmentation for Chest Pathologies Classification, in: 2019 IEEE 16th

International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 1216–1219. doi:10.1109/ISBI.2019.8759573.

Sivaramakrishnan, R., Antani, S., Candemir, S., Xue, Z., Thoma, G., Alderson, P., Abuya, J., Kohli, M., 2018. Comparing deep learning models for population screening using chest radiography, in: Medical Imaging 2018: Computer-Aided Diagnosis, SPIE. p. 49. doi:10.1117/12.2293140.

Soganciogl u, E., Murphy, K., Calli, E., Scholten, E.T., Schalekamp, S., Ginneken, B.V., 2020. Cardiomegaly detection on chest radiographs: Segmentation versus classification. IEEE Access 8, 94631–94642. doi:10.1109/access.2020.2995567.

Souza, J.C., Bandeira Diniz, J.O., Ferreira, J.L., França da Silva, G.L., Corrêa Silva, A., de Paiva, A.C., 2019. An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks. Computer Methods and Programs in Biomedicine 177, 285–296. doi:10.1016/j.cmpb.2019.06.005.

Strohm, L., Hehakaya, C., Ranschaert, E.R., Boon, W.P.C., Moors, E.H.M., 2020. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. European Radiology 30, 5525–5532. doi:10.1007/s00330-020-06946-y.

Su, C.Y., Tsai, T.Y., Tseng, C.Y., Liu, K.H., Lee, C.W., 2021. A Deep Learning Method for Alerting Emergency Physicians about the Presence of Subphrenic Free Air on Chest Radiographs. Journal of Clinical Medicine 10, 254. doi:10.3390/jcm10020254.

Subramanian, V., Wang, H., Wu, J.T., Wong, K.C.L., Sharma, A., Syeda-Mahmood, T., 2019. Automated Detection and Type Classification of Central Venous Catheters in Chest X-Rays, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 522–530. doi:10.1007/978-3-030-32226-7_58.

Sullivan, R.P., Holste, G., Burkow, J., Alessio, A., 2020. Deep learning methods for segmentation of lines in pediatric chest radiographs, in: Medical Imaging 2020: Computer-Aided Diagnosis, SPIE. p. 87. doi:10.1117/12.2550686.

Syeda-Mahmood, T., Ahmad, H., Ansari, N., Gur, Y., Kashyap, S., Karargyris, A., Moradi, M., Pillai, A., Seshadhri, K., Wang, W., Wong, K.C.L., Wu, J., 2019. Building a Benchmark Dataset and Classifiers for Sentence-Level Findings in AP Chest X-Rays, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 863–867. doi:10.1109/ISBI.2019.8759162.

Syeda-Mahmood, T., Wong, K.C.L., Gur, Y., Wu, J.T., Jadhav, A., Kashyap, S., Karargyris, A., Pillai, A., Sharma, A., Syed, A.B., Boyko, O., Moradi, M., 2020. Chest X-Ray Report Generation Through Fine-Grained Label Learning, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12262, pp. 561–571. doi:10.1007/978-3-030-59713-9_54.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI Press. p. 4278–4284.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions, in: IEEE conference on computer vision and pattern recognition, pp. 1–9.

Szucs-Farkas, Z., Schick, A., Cullmann, J.L., Ebner, L., Megyeri, B., Vock, P., Christe, A., 2013. Comparison of dual-energy subtraction and electronic bone suppression combined with computer-aided detection on chest radiographs: effect on human observers' performance in nodule detection. AJR. American journal of roentgenology 200, 1006–1013. doi:10.2214/AJR.12.8877.

Tabik, S., Gomez-Rios, A., Martin-Rodriguez, J.L., Sevillano-Garcia, I., Rey-Area, M., Charte, D., Guirado, E., Suarez, J.L., Luengo, J., Valero-Gonzalez, M.A., Garcia-Villanova, P., Olmedo-Sanchez, E., Herrera, F., 2020. COVIDGR Dataset and COVID-SDNet Methodology for Predicting COVID-19 Based on Chest X-Ray Images. IEEE Journal of Biomedical and Health Informatics 24, 3595–3605. doi:10.1109/JBHI.2020.3037127.

Taghanaki, S.A., Abhishek, K., Hamarneh, G., 2019a. Improved Inference via Deep Input Transfer, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 819–827. doi:10.1007/978-3-030-32226-7_91.

Taghanaki, S.A., Havaei, M., Berthier, T., Dutil, F., Di Jorio, L., Hamarneh, G., Bengio, Y., 2019b. InfoMask: Masked Variational Latent Representation to Localize Chest Disease, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 739–747. doi:10.1007/978-3-030-32226-7_82.

Takaki, T., Murakami, S., Watanabe, R., Aoki, T., Fujibuchi, T., 2020. Calculating the target exposure index using a deep convolutional neural network and a rule base. Physica Medica 71, 108–114. doi:10.1016/j.ejmp.2020.02.012.

Takemiya, R., Kido, S., Hirano, Y., Mabu, S., 2019. Detection of pulmonary nodules on chest x-ray images using R-CNN, in: International Forum on Medical Imaging in Asia 2019, SPIE. p. 58. doi:10.1117/12.2521652.

Tam, L.K., Wang, X., Turkbey, E., Lu, K., Wen, Y., Xu, D., 2020. Weakly Supervised One-Stage Vision and Language Disease Detection Using Large Scale Pneumonia and Pneumothorax Studies, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12264, pp. 45–55. doi:10.1007/978-3-030-59719-1_5.

Tang, Y., Tang, Y., Sandfort, V., Xiao, J., Summers, R.M., 2019a. TUNA-Net: Task-Oriented UNsupervised Adversarial Network for Disease Recognition in Cross-domain Chest X-rays, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 431–440. doi:10.1007/978-3-030-32226-7_48.

Tang, Y.B., Tang, Y.X., Xiao, J., Summers, R.M., 2019b. Xlsor: A robust and accurate lung segmentor on chest x-rays using criss-cross attention and customized radiorealistic abnormalities generation, in: International Conference on Medical Imaging with Deep Learning, PMLR. pp. 457–467.

Tang, Y.X., Tang, Y.B., Han, M., Xiao, J., Summers, R.M., 2019c. Abnormal Chest X-Ray Identification With Generative Adversarial One-Class Classifier, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE. pp. 1358–1361. doi:10.1109/ISBI.2019.8759442.

Tang, Y.X., Tang, Y.B., Peng, Y., Yan, K., Bagheri, M., Redd, B.A., Brandon, C.J., Lu, Z., Han, M., Xiao, J., Summers, R.M., 2020. Automated abnormality classification of chest radiographs using deep convolutional neural networks. npj Digital Medicine 3, 70. doi:10.1038/s41746-020-0273-z.

Tartaglione, E., Barbano, C.A., Berzovini, C., Calandri, M., Grangetto, M., 2020. Unveiling COVID-19 from CHEST X-Ray with Deep Learning: A Hurdles Race with Small Data. International Journal of Environmental Research and Public Health 17, 6933. doi:10.3390/ijerph17186933.

Taylor, A.G., Mielke, C., Mongan, J., 2018. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. PLOS Medicine 15, e1002697. doi:10.1371/journal.pmed.1002697.

Thammarach, P., Khaengthanyakan, S., Vongsurakrai, S., Phienphanich, P., Pooprasert, P., Yaemsuk, A., Vanichvarodom, P., Munpolsri, N., Khwayotha, S., Lertkowit, M., Tungsagunwattana, S., Vijitsanguan, C., Lertrojanapunya, S., Noisiri, W., Chiawiriyabunya, I., Aphikulvanich, N., Tantibundhit, C., 2020. AI Chest 4 All, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), pp. 1229–1233. doi:10.1109/EMBC44109.2020.9175862. iSSN: 2694-0604.

Toba, S., Mitani, Y., Yodoya, N., Ohashi, H., Sawada, H., Hayakawa, H., Hirayama, M., Futsuki, A., Yamamoto, N., Ito, H., Konuma, T., Shimpo, H., Takao, M., 2020. Prediction of Pulmonary to Systemic Flow Ratio in Patients With Congenital Heart Disease Using Deep Learning–Based Analysis of Chest Radiographs. JAMA Cardiology 5, 449. doi:10.1001/jamacardio.2019.5620.

Tolkachev, A., Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Ibragimov, B., 2020. Deep Learning for Diagnosis and Segmentation of Pneumothorax: The Results on The Kaggle Competition and Validation Against Radiologists. IEEE Journal of Biomedical and Health Informatics , 1–1doi:10.1109/JBHI.2020.3023476.

Toriwaki, J.I., Suenaga, Y., Negoro, T., Fukumura, T., 1973. Pattern recognition of chest x-ray images. Computer Graphics and Image Processing 2, 252–271. doi:10.1016/0146-664x(73)90005-1.

Ul Abideen, Z., Ghafoor, M., Munir, K., Saqib, M., Ullah, A., Zia, T., Tariq, S.A., Ahmed, G., Zahra, A., 2020. Uncertainty Assisted Robust Tuberculosis Identification With Bayesian Convolutional Neural Networks. IEEE Access 8, 22812–22825. doi:10.1109/ACCESS.2020.2970023.

Umehara, K., Ota, J., Ishimaru, N., Ohno, S., Okamoto, K., Suzuki, T., Shirai, N., Ishida, T., 2017. Super-resolution convolutional neural network for the improvement of the image quality of magnified images in chest radiographs, in: Medical Imaging 2017: Image Processing, International Society for Optics and Photonics. p. 101331P. doi:10.1117/12.2249969.

United Nations, 2008. United nations scientific committee on the effects of atomic radiation (UNSCEAR), 2008 report on sources and effects of ionizing radiation. http://www.unscear.org/docs/publications/2008/UNSCEAR_2008_Annex-A-CORR.pdf.

Unnikrishnan, B., Nguyen, C.M., Balaram, S., Foo, C.S., Krishnaswamy, P., 2020. Semi-supervised Classification of Diagnostic Radiographs with

NoTeacher: A Teacher that is Not Mean, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 624–634. doi:10.1007/978-3-030-59710-8_61.

Ureta, J., Aran, O., Rivera, J.P., 2020. Detecting pneumonia in chest radiographs using convolutional neural networks, in: Twelfth International Conference on Machine Vision (ICMV 2019), SPIE. p. 116. doi:10.1117/12.2559527.

Uzunova, H., Ehrhardt, J., Jacob, F., Frydrychowicz, A., Handels, H., 2019. Multi-scale GANs for Memory-efficient Generation of High Resolution Medical Images, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 112–120. doi:10.1007/978-3-030-32226-7_13.

Vayá, M.d.l.I., Saborit, J.M., Montell, J.A., Pertusa, A., Bustos, A., Cazorla, M., Galant, J., Barber, X., Orozco-Beltrán, D., García-García, F., Caparrós, M., González, G., Salinas, J.M., 2020. BIMCV COVID-19+: a large annotated dataset of RX and CT images from COVID-19 patients. arXiv:2006.01174 [cs, eess] ArXiv: 2006.01174.

Viergever, M.A., Maintz, J.A., Klein, S., Murphy, K., Staring, M., Pluim, J.P., 2016. A survey of medical image registration – under review. Medical Image Analysis 33, 140–144. doi:10.1016/j.media.2016.06.030.

Wang, C., Elazab, A., Jia, F., Wu, J., Hu, Q., 2018. Automated chest screening based on a hybrid model of transfer learning and convolutional sparse denoising autoencoder. BioMedical Engineering OnLine 17, 63. doi:10.1186/s12938-018-0496-2.

Wang, C., Elazab, A., Wu, J., Hu, Q., 2017a. Lung nodule classification using deep feature fusion in chest radiography. Computerized Medical Imaging and Graphics 57, 10–18. doi:10.1016/j.compmedimag.2016.11.004.

Wang, H., Gu, H., Qin, P., Wang, J., 2020a. CheXLocNet: Automatic localization of pneumothorax in chest radiographs using deep convolutional neural networks. PLoS One 15, e0242013. doi:10.1371/journal.pone.0242013.

Wang, H., Jia, H., Lu, L., Xia, Y., 2020b. Thorax-Net: An Attention Regularized Deep Neural Network for Classification of Thoracic Diseases on Chest Radiography. IEEE Journal of Biomedical and Health Informatics 24, 475–485. doi:10.1109/JBHI.2019.2928369.

Wang, H., Wang, S., Qin, Z., Zhang, Y., Li, R., Xia, Y., 2021a. Triple attention learning for classification of 14 thoracic diseases using chest radiography. Med. Image Anal. 67, 101846. doi:10.1016/j.media.2020.101846.

Wang, M., Deng, W., 2018. Deep visual domain adaptation: A survey. Neurocomputing 312, 135–153. doi:10.1016/j.neucom.2018.05.083.

Wang, Q., Liu, Q., Luo, G., Liu, Z., Huang, J., Zhou, Y., Zhou, Y., Xu, W., Cheng, J.Z., 2020c. Automated segmentation and diagnosis of pneumothorax on chest X-rays with fully convolutional multi-scale ScSE-DenseNet: a retrospective study. BMC Med. Inform. Decis. Mak. 20, 317. doi:10.1186/s12911-020-01325-5.

Wang, W., Feng, H., Bu, Q., Cui, L., Xie, Y., Zhang, A., Feng, J., Zhu, Z., Chen, Z., 2020d. MDU-Net: A Convolutional Network for Clavicle and Rib Segmentation from a Chest Radiograph. Journal of Healthcare Engineering 2020, 1–9. doi:10.1155/2020/2785464.

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M., 2017b. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2097–2106. doi:10.1109/cvpr.2017.369.

Wang, X., Schwab, E., Rubin, J., Klassen, P., Liao, R., Berkowitz, S., Golland, P., Horng, S., Dalal, S., 2019. Pulmonary edema severity estimation in chest radiographs using deep learning, in: International Conference on Medical Imaging with Deep Learning–Extended Abstract Track, pp. 1,4.

Wang, X., Yu, J., Zhu, Q., Li, S., Zhao, Z., Yang, B., Pu, J., 2020e. Potential of deep learning in assessing pneumoconiosis depicted on digital chest radiography. Occupational and Environmental Medicine 77, 597–602. doi:10.1136/oemed-2019-106386.

Wang, Z., Xiao, Y., Li, Y., Zhang, J., Lu, F., Hou, M., Liu, X., 2021b. Automatically discriminating and localizing COVID-19 from community-acquired pneumonia on chest X-rays. Pattern Recognition 110, 107613. doi:10.1016/j.patcog.2020.107613.

Wehbe, R.M., Sheng, J., Dutta, S., Chai, S., Dravid, A., Barutcu, S., Wu, Y., Cantrell, D.R., Xiao, N., Allen, B.D., MacNealy, G.A., Savas, H., Agrawal, R., Parekh, N., Katsaggelos, A.K., 2020. DeepCOVID-XR: An Artificial Intelligence Algorithm to Detect COVID-19 on Chest Radiographs Trained and Tested on a Large US Clinical Dataset. Radiology , 203511doi:10.1148/radiol.2020203511.

Wei, Y., Feng, J., Liang, X., Cheng, M.M., Zhao, Y., Yan, S., 2018. Object Region Mining with Adversarial Erasing: A Simple Classification to Semantic Segmentation Approach. arXiv:1703.08448 [cs] ArXiv: 1703.08448.

Wessel, J., Heinrich, M.P., von Berg, J., Franz, A., Saalbach, A., 2019. Sequential Rib Labeling and Segmentation in Chest X-Ray using Mask R-CNN. arXiv:1908.08329.

Wolleb, J., Sandkühler, R., Cattin, P.C., 2020. DeScarGAN: Disease-Specific Anomaly Detection with Weak Supervision, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12264, pp. 14–24. doi:10.1007/978-3-030-59719-1_2.

Wong, K.C.L., Moradi, M., Wu, J., Pillai, A., Sharma, A., Gur, Y., Ahmad, H., Chowdary, M.S., Chiranjeevi, J., Reddy Polaka, K.K., Wunnava, V., Reddy, D., Syeda-Mahmood, T., 2020. A Robust Network Architecture to Detect Normal Chest X-Ray Radiographs, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1851–1855. doi:10.1109/ISBI45749.2020.9098671.

Xing, Y., Ge, Z., Zeng, R., Mahapatra, D., Seah, J., Law, M., Drummond, T., 2019. Adversarial Pulmonary Pathology Translation for Pairwise Chest X-Ray Data Augmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 757–765. doi:10.1007/978-3-030-32226-7_84.

Xu, Y., Mo, T., Feng, Q., Zhong, P., Lai, M., Chang, E.I., 2014. Deep learning of feature representation with multiple instance learning for medical image analysis, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1626–1630. doi:10.1109/ICASSP.2014.6853873. iSSN: 2379-190X.

Xue, C., Deng, Q., Li, X., Dou, Q., Heng, P.A., 2020. Cascaded Robust Learning at Imperfect Labels for Chest X-ray Segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12266, pp. 579–588. doi:10.1007/978-3-030-59725-2_56.

Xue, F.F., Peng, J., Wang, R., Zhang, Q., Zheng, W.S., 2019. Improving Robustness of Medical Image Diagnosis with Denoising Convolutional Neural Networks, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 846–854. doi:10.1007/978-3-030-32226-7_94.

Xue, Y., Xu, T., Rodney Long, L., Xue, Z., Antani, S., Thoma, G.R., Huang, X., 2018a. Multimodal Recurrent Model with Attention for Automated Radiology Report Generation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11070, pp. 457–466. doi:10.1007/978-3-030-00928-1_52.

Xue, Z., Antani, S., Long, R., Thoma, G.R., 2018b. Using deep learning for detecting gender in adult chest radiographs, in: Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications, SPIE. p. 10. doi:10.1117/12.2293027.

Xue, Z., Jaeger, S., Antani, S., Long, R., Karagyris, A., Siegelman, J., Folio, L.R., Thoma, G.R., 2018c. Localizing tuberculosis in chest radiographs with deep learning, in: Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications, SPIE. p. 28. doi:10.1117/12.2293022.

Xue, Z., Long, R., Jaeger, S., Folio, L., George Thoma, R., Antani, a.S., 2018d. Extraction of Aortic Knuckle Contour in Chest Radiographs Using Deep Learning, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 5890–5893. doi:10.1109/EMBC.2018.8513560.

Yahyatabar, M., Jouvet, P., Cheriet, F., 2020. Dense-Unet: a light model for lung fields segmentation in Chest X-Ray images, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE. pp. 1242–1245. doi:10.1109/EMBC44109.2020.9176033.

Yang, W., Chen, Y., Liu, Y., Zhong, L., Qin, G., Lu, Z., Feng, Q., Chen, W., 2017. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. Medical Image Analysis 35, 421–433. doi:10.1016/j.media.2016.08.004.

Yao, L., Prosky, J., Covington, B., Lyman, K., 2019. A strong baseline for domain adaptation and generalization in medical imaging, in: International Conference on Medical Imaging with Deep Learning – Extended Abstract Track, pp. 1–4.

Yi, P.H., Kim, T.K., Yu, A.C., Bennett, B., Eng, J., Lin, C.T., 2020. Can AI outperform a junior resident? Comparison of deep neural network to first-year radiology residents for identification of pneumothorax. Emergency Radiology 27, 367–375. doi:10.1007/s10140-020-01767-4.

Yi, X., Adams, S., Babyn, P., Elnajmi, A., 2019a. Automatic Catheter and Tube

Detection in Pediatric X-ray Images Using a Scale-Recurrent Network and Synthetic Data. Journal of Digital Imaging 33, 181–190. doi:10.1007/s10278-019-00201-7.

Yi, X., Walia, E., Babyn, P., 2019b. Generative adversarial network in medical imaging: A review. Medical Image Analysis 58, 101552. doi:10.1016/j.media.2019.101552.

Yoo, H., Kim, K.H., Singh, R., Digumarthy, S.R., Kalra, M.K., 2020. Validation of a Deep Learning Algorithm for the Detection of Malignant Pulmonary Nodules in Chest Radiographs. JAMA Network Open 3, e2017135. doi:10.1001/jamanetworkopen.2020.17135.

Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks?, in: Advances in Neural Information Processing Systems, Curran Associates, Inc.. pp. 1–9.

Young, M., 1994. Interobserver variability in the interpretation of chest roentgenograms of patients with possible pneumonia. Archives of Internal Medicine 154, 2729. doi:10.1001/archinte.1994.00420230122014.

Yu, D., Zhang, K., Huang, L., Zhao, B., Zhang, X., Guo, X., Li, M., Gu, Z., Fu, G., Hu, M., Ping, Y., Sheng, Y., Liu, Z., Hu, X., Zhao, R., 2020. Detection of peripherally inserted central catheter (PICC) in chest X-ray images: A multi-task deep learning model. Computer Methods and Programs in Biomedicine 197, 105674. doi:10.1016/j.cmpb.2020.105674.

Yuan, J., Liao, H., Luo, R., Luo, J., 2019. Automatic Radiology Report Generation Based on Multi-view Image Fusion and Medical Concept Enrichment, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 721–729. doi:10.1007/978-3-030-32226-7_80.

Yue, Z., Ma, L., Zhang, R., 2020. Comparison and Validation of Deep Learning Models for the Diagnosis of Pneumonia. Computational Intelligence and Neuroscience 2020, 1–8. doi:10.1155/2020/8876798.

Zarei, M., Abadi, E., Fricks, R., Segars, W.P., Samei, E., 2021. A probabilistic conditional adversarial neural network to reduce imaging variation in radiography, in: Medical Imaging 2021: Physics of Medical Imaging, International Society for Optics and Photonics. p. 115953Y. doi:10.1117/12.2582336.

Zarshenas, A., Liu, J., Forti, P., Suzuki, K., 2019. Separation of bones from soft tissue in chest radiographs: Anatomy-specific orientation-frequency-specific deep neural network convolution. Medical Physics 46, 2232–2242. doi:10.1002/mp.13468.

Zech, J.R., Badgeley, M.A., Liu, M., Costa, A.B., Titano, J.J., Oermann, E.K., 2018. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. PLOS Medicine 15, e1002683. doi:10.1371/journal.pmed.1002683.

Zhang, J., Xie, Y., Pang, G., Liao, Z., Verjans, J., Li, W., Sun, Z., He, J., Li, Y., Shen, C., Xia, Y., 2021a. Viral Pneumonia Screening on Chest X-Rays Using Confidence-Aware Anomaly Detection. IEEE transactions on medical imaging 40, 879–890. doi:10.1109/TMI.2020.3040950.

Zhang, L., Rong, R., Li, Q., Yang, D.M., Yao, B., Luo, D., Zhang, X., Zhu, X., Luo, J., Liu, Y., Yang, X., Ji, X., Liu, Z., Xie, Y., Sha, Y., Li, Z., Xiao, G., 2021b. A deep learning-based model for screening and staging pneumoconiosis. Sci. Rep. 11, 2201. doi:10.1038/s41598-020-77924-z.

Zhang, M., Gao, J., Lyu, Z., Zhao, W., Wang, Q., Ding, W., Wang, S., Li, Z., Cui, S., 2020. Characterizing Label Errors: Confident Learning for Noisy-Labeled Image Segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 721–730. doi:10.1007/978-3-030-59710-8_70.

Zhang, R., Tie, X., Qi, Z., Bevins, N.B., Zhang, C., Griner, D., Song, T.K., Nadig, J.D., Schiebler, M.L., Garrett, J.W., Li, K., Reeder, S.B., Chen, G.H., 2021c. Diagnosis of Coronavirus Disease 2019 Pneumonia by Using Chest Radiography: Value of Artificial Intelligence. Radiology 298, E88–E97. doi:10.1148/radiol.2020202944.

Zhang, T., Fu, H., Zhao, Y., Cheng, J., Guo, M., Gu, Z., Yang, B., Xiao, Y., Gao, S., Liu, J., 2019a. SkrGAN: Sketching-Rendering Unconditional Generative Adversarial Networks for Medical Image Synthesis, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11767, pp. 777–785. doi:10.1007/978-3-030-32251-9_85.

Zhang, W., Li, G., Wang, F., E, L., Yu, Y., Lin, L., Liang, H., 2019b. Simultaneous Lung Field Detection and Segmentation for Pediatric Chest Radiographs, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11769, pp. 594–602. doi:10.1007/978-3-030-32226-7_66.

Zhang, Y., Miao, S., Mansi, T., Liao, R., 2018. Task Driven Generative Modeling for Unsupervised Domain Adaptation: Application to X-ray Image Segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer. volume 11071, pp. 599–607. doi:10.1007/978-3-030-00934-2_67.

Zhang, Z., Fu, H., Dai, H., Shen, J., Pang, Y., Shao, L., 2019c. ET-Net: A Generic Edge-aTtention Guidance Network for Medical Image Segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer. volume 11764, pp. 442–450. doi:10.1007/978-3-030-32239-7_49.

Zhao, Z.Q., Zheng, P., Xu, S.T., Wu, X., 2019. Object detection with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems 30, 3212–3232. doi:10.1109/tnnls.2018.2876865.

Zhou, H.Y., Yu, S., Bian, C., Hu, Y., Ma, K., Zheng, Y., 2020a. Comparing to Learn: Surpassing ImageNet Pretraining on Radiographs by Comparing Image Representations, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer. volume 12261, pp. 398–407. doi:10.1007/978-3-030-59710-8_39.

Zhou, S., Zhang, X., Zhang, R., 2019. Identifying Cardiomegaly in ChestX-ray8 Using Transfer Learning. Studies in Health Technology and Informatics 264, 482–486. doi:10.3233/SHTI190268.

Zhou, Z., Zhou, L., Shen, K., 2020b. Dilated conditional GAN for bone suppression in chest radiographs with enforced semantic features. Medical Physics , mp.14371doi:10.1002/mp.14371.

Zhu, C.S., Pinsky, P.F., Kramer, B.S., Prorok, P.C., Purdue, M.P., Berg, C.D., Gohagan, J.K., 2013. The Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial and Its Associated Research Resource. JNCI Journal of the National Cancer Institute 105, 1684–1693. doi:10.1093/jnci/djt281.

Zhu, J., Shen, B., Abbasi, A., Hoshmand-Kochi, M., Li, H., Duong, T.Q., 2020. Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs. PLOS ONE 15, e0236621. doi:10.1371/journal.pone.0236621.

Zhu, J.Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks, in: 2017 IEEE International Conference on Computer Vision (ICCV), IEEE. pp. 2223–2232. doi:10.1109/iccv.2017.244.

Zou, X.L., Ren, Y., Feng, D.Y., He, X.Q., Guo, Y.F., Yang, H.L., Li, X., Fang, J., Li, Q., Ye, J.J., Han, L.Q., Zhang, T.T., 2020. A promising approach for screening pulmonary hypertension based on frontal chest radiographs using deep learning: A retrospective study. PLOS ONE 15, e0236378. doi:10.1371/journal.pone.0236378.

Zucker, E.J., Barnes, Z.A., Lungren, M.P., Shpanskaya, Y., Seekins, J.M., Halabi, S.S., Larson, D.B., 2020. Deep learning to automate Brasfield chest radiographic scoring for cystic fibrosis. Journal of Cystic Fibrosis 19, 131–138. doi:10.1016/j.jcf.2019.04.016.

Zunair, H., Hamza, A.B., 2021. Synthesis of COVID-19 chest X-rays using unpaired image-to-image translation. Social Network Analysis and Mining 11, 23. doi:10.1007/s13278-021-00731-5.