Deep learning classification of cardiomegaly using combined imaging and non-imaging ICU data

Declan Grant[1], Bartłomiej W. Papież[2], Guy Parsons[3], Lionel Tarassenko[1], and Adam Mahdi[1,*]

[1] Department of Engineering Science, University of Oxford, UK
declan.grant@exeter.ox.ac.uk, lionel.tarassenko,adam.mahdi@eng.ox.ac.uk
[2] Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK bartlomiej.papiez@bdi.ox.ac.uk
[3] NIHR Academic Clinical Fellow at Oxford University, Kadoorie Centre and Intensive Care Registrar, Thames Valley Deanery, UK guy.parsons@nhs.net

[*] Corresponding author: Adam Mahdi, adam.mahdi@eng.ox.ac.uk

Abstract. In this paper, we investigate the classification of cardiomegaly using multimodal data, combining imaging data from chest radiography with routinely collected Intensive Care Unit (ICU) data comprising vital sign values, laboratory measurements, and admission metadata. In practice a clinician would assess for the presence of cardiomegaly using a synthesis of multiple sources of data, however, prior machine learning approaches to this task have focused on chest radiographs only. We show that non-imaging ICU data can be used for cardiomegaly classification and propose a novel multimodal network trained simultaneously on both chest radiographs and ICU data. We compare the predictive power of both single-mode approaches with the joint network. We use a subset of data from the publicly available MIMIC-CXR and MIMIC-IV datasets, which contain both chest radiographs and non-imaging ICU data for the same patients. The approach from non-imaging ICU data alone achieves an AUC of 0.684 and the standard chest radiography approach an AUC of 0.840. Our joint model achieves an AUC of 0.880. We conclude that non-imaging ICU data have predictive value for cardiomegaly, and that combining chest radiographs with non-imaging ICU data has the potential to improve model performance for the same subset of patients, with further work required to demonstrate a significant improvement.

Keywords: deep learning · chest X-ray · cardiomegaly · multimodal approach.

1 Introduction

Cardiomegaly is an abnormal enlargement of the heart usually indicating an underlying pathology warranting further investigation. In clinical practice, cardiomegaly can be detected visually by examining the size of the heart on a

postero-anterior chest radiograph. A standard parameter for diagnosis is the cardiothoracic ratio (CTR), which is the ratio of the cardiac width (maximum horizontal cardiac diameter) to the thoracic width (maximum horizontal distance between the inner edges of the ribs). Automated detection of cardiomegaly originally used edge detection to measure the CTR [10] but more recently there has been growing interest into deep learning approaches to the problem. This has been accelerated by the availability of large, publicly available clinical imaging datasets [24, 12]. Deep learning methods focus on using convolutional neural networks to either assign a binary cardiomegaly diagnosis based on the input image [19], or to use U-Net networks to segment the heart and lungs from the image and estimate the CTR [18, 23].

Until now, classification methods have been almost exclusively applied to imaging datasets (primarily radiographs), which occasionally provide some demographic data such as gender and age [12]. However, this is in contrast to how clinicians would diagnose cardiomegaly. A clinician may use a chest radiograph alone to diagnose the condition but would also consider other types of medical data, including demographics, results from a number of blood tests, other imaging information and vital sign data to assess severity and underlying pathology.

Previous approaches, which introduced multimodal network architectures, have combined medical images with basic demographics to predict the outcome of endovascular treatment from clinical metadata and imaging [20], or to classify skin lesions from dermoscopic images and patient age and sex [7]. Limited amount of work has been done in relation to cardiomegaly classification from combining imaging and non-imaging data. An example is the method proposed by [1], in which the neural network integrates the X-ray image with the limited non-imaging data (patient age, gender and acquisition type) in the classification process.

In this work, we aim to improve prediction of cardiomegaly by proposing a novel model capable of combining imaging with non-imaging Intensive Care Unit (ICU) data. To the best of our knowledge, this is the first study that combines chest radiographs and extensive non-imaging information collected during the ICU stay including vital sign values, laboratory measurements, and demographics. We use a combination of the MIMIC-IV and MIMIC-CXR, which constitutes one of the first publicly available datasets of X-rays and cardiomegaly labels with additional modalities of vital sign values, laboratory measurements, and patient metadata. We assess the relative predictive power of each modality, and then compare our joint framework with the approaches that utilize imaging or non-imaging data alone.

2 Data and Methods

2.1 Datasets

We used the publicly available MIMIC-CXR [12] and MIMIC-IV [11] datasets. Both of these databases used the same patient IDs, stays, admissions and the dates and times are consistent for each patient.

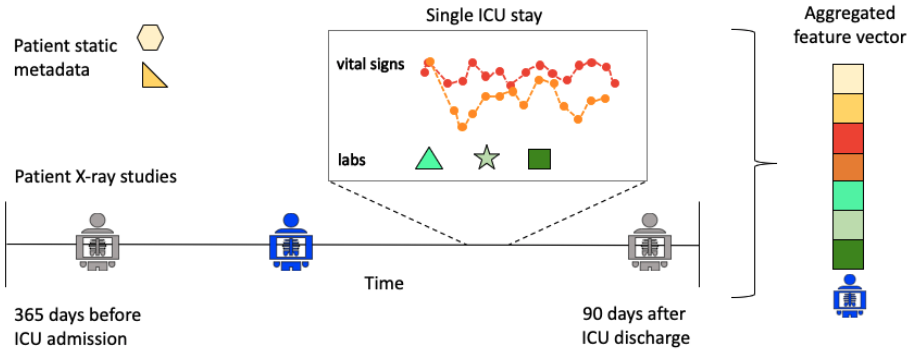

Fig. 1. Aggregating feature vectors from multimodal ICU data (see Sec 2.3). We built a consistent, high-quality dataset, which combines imaging and non-imaging data including patient static metadata (e.g. age, gender, ethnicity, length of stay) as well as laboratory values (e.g. glucose, magnesium, hemoglobin) and vital sign values (e.g. heart rate, respiration rate, oxygen saturation).

MIMIC-CXR Database. MIMIC-CXR is a large publicly available database of patient chest radiographs collected from the Beth Israel Deaconess Medical Center (BIDMC) emergency department between 2011 and 2017. It contains 227,835 X-ray studies for 64,588 patients. Each study may contain multiple images from different view positions and in total there are 377,110 radiographs. Every study also has an associated free-text radiology report, written at the time of the study. Here, we used MIMIC-CXR in JPG format provided with MIMIC-CXR-JPG [13] as it additionally contains structured labels derived from these reports.

MIMIC-IV Database. MIMIC-IV contains data from hospital stays for patients who were admitted to the BIDMC between 2008 and 2019. MIMIC-IV is separated into five modules: core (patient stay information), hospital (laboratories and microbiology), ICU data (ICU stays and events), emergency department and CXR (lookup tables to allow linking to MIMIC-CXR).

2.2 Preprocessing

Cardiomegaly was detected using posterior-anterior radiographs, to prevent the artificial enlargement of the cardiac silhouette that can occur with antero-posterior and other radiograph views. We therefore filtered the image datasets to only include posterior-anterior views. MIMIC-CXR-JPG provides four types of cardiomegaly labels (positive, negative, uncertain, no mention) derived using two natural language processing tools NegBio [15] and CheXpert [9]. We removed any images where these tools disagreed with each other over the label assigned, to prevent introducing error. We note that we cannot conclude that a report with

no mention of cardiomegaly indicates the absence of cardiomegaly. Therefore, we created a subset of images, which only have either a positive or negative label for cardiomegaly. These images were then re-sized to 244 x 244 pixels and underwent normalisation of the mean and standard deviation. We also performed standard data augmentation consisting of a random rotation of up to 10 degrees as well as random horizontal and vertical flips to improve robustness.

We grouped vital sign values (e.g. heart rate, respiration rate, oxygen saturation) and laboratory values (e.g. glucose, magnesium, hemoglobin) for each ICU admission, before averaging them across the stay to produce a single set of summary features. We then added patient metadata (e.g. age, gender, ethnicity, length of stay) to each feature vector.

2.3 Combining imaging and non-imaging data

There is a unique patient ID across MIMIC-CXR-JPG and MIMIC-IV. We linked ICU stays to radiographic studies, which took place up to 365 days before the patient entered the ICU and up to 90 days after leaving the ICU. Cardiomegaly is a chronic condition that usually slowly progresses if the underlying pathology remains untreated, and it does not resolve once present without considerable intervention. This period was therefore identified as providing a reasonable window of stability in the condition for assessment in our modelling [5]. For each unique ICU stay, we first collected the studies from our images subset which took place within the specified window for that same patient (See Fig. 1). We then checked for consistency between the image labels. We assumed cardiomegaly does not change from positive to negative within the chosen window. If a patient has conflicting (both positive and negative) image labels within this window, we cannot be confident which is the correct label to assign to the ICU stay, so we remove ICU stays surrounded by conflicting cardiomegaly labels. If the labels show consistency, we link the ICU stay and the image study which took place closest to ICU admission. While the labels generated from free-text reports [15, 9] can be inconsistent, the presented data curation process led us to create a high-quality dataset (both imaging and non-imaging), which was later used to train, validate, and test the models presented in the next section.

2.4 Models

Image classification. We implemented a ResNet [8] architecture for classifying cardiomegaly presence in X-ray images. ResNet architectures have been evaluated on similar radiograph databases and shown to achieve state-of-the-art results for this task. For example, ResNet-32 achieved 0.84 cardiomegaly accuracy [2] on the CheXpert database [9], which also uses labels automatically derived from the reports. All networks were pre-trained on ImageNet [6]. We used a binary cross entropy loss function, Adam optimizer [14] and cyclical learning rates [22]. Each image-based model was trained in two stages. First, we froze the convolutional layers of the ResNet and trained just the final fully connected layers. Then we unfroze all of the layers and continued training. Before each stage we

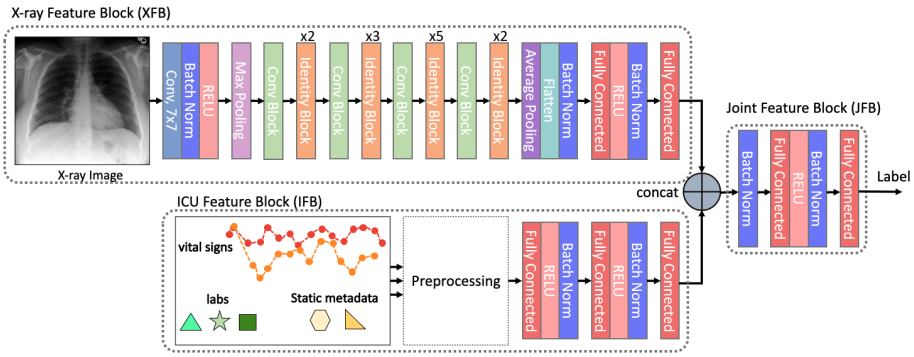

Fig. 2. Our approach for combining X-ray imaging and non-imaging ICU data.

conducted an initial learning rate test and determined optimal boundaries for the learning rate to cycle in during training [22].

ICU data classification. We implemented XGBoost [4] to classify cardiomegaly from ICU data consisting of vital sign values, laboratory measurements and demographics. XGBoost has been shown to perform well at similar classification tasks in clinical machine learning [3, 16, 17]. We use a binary cross entropy loss function.

Multimodal classification. We propose a multimodal network for the classification of cardiomegaly (see Fig. 2), which combines simultaneously the imaging data (chest radiographs) and non-imaging ICU data (vital sign values, laboratory values, static patient metadata including demographics). We use a ResNet-50 architecture in the X-ray Feature Extraction Block (XFB) to extract relevant features from the images. The ICU data is fed into the ICU Feature Block (IFB), which contains a neural network consisting of three fully connected layers. To join the learned image and ICU features, we concatenate the outputs of the XFB (32 nodes) and IFB (16 nodes) using a fully connected layer. After concatenation, we add two more fully connected layers to produce a label. Previous approaches have used similar network architectures to combine medical imaging with small amounts of patient metadata [20, 7, 1], however, here we extend on this to include all routinely collected ICU data including vital sign values and laboratory measurements.

The joint network was trained similarly to the ResNet model, using a binary cross entropy loss function, Adam optimizer [14] and cyclical learning rates [22]. We also pre-trained the XFB and IFB separately before concatenating them together. The model was then trained in two stages. First, we froze the early layers and trained only the fully connected layers of the JFB. Then we unfroze the entire network and continued training. Optimal bounds for the cyclical learning rate were determined before each stage from a learning rate test.

Table 1. Patient characteristics.

Variable	**Cardiomegaly**	**All patients**
ICU Stays	1,795	2,571
Male / Female	756 / 904	1077 / 1327
Age [years], mean (SD)	66.0 (14.9)	64.9 (15.6)
Ethnicity		
White, N (%)	1,084 (65.3)	1,601 (66.9)
Asian, N (%)	58 (3.5)	93 (3.9)
Black, N (%)	309 (18.6)	390 (16.3)
Hispanic/Latino, N (%)	94 (5.7)	127 (5.3)
Other/Mixed/Unknown, N (%)	78 (10.2)	183 (7.6)
LOS [days], median (IQR)	2.1 (2.5)	2.0 (2.4)

3 Results

The original MIMIC-CXR contains 227,835 studies (377,110 radiographs) for 64,588 patients including 53,565 X-ray studies, with positive or negative cardiomegaly labels, for 22,914 patients. After applying the inclusion and exclusion criteria described in Sec. 2.2, the dataset used in this work contains 2,571 ICU stays and 2,404 patients. A brief overview of the patient characteristics including Sex, Age, Ethnicity and length of stay (LOS) is given in Table 1.

We divided our curated dataset (see Sec. 2.2) into 5-folds for cross-validation. No patient appears in more than one fold and every fold has the same ratio of positive to negative labels. Each model was trained on the same 5-folds but only made use of the relevant modalities: ResNet (Images), XGBoost (ICU data) and the joint network (Images and non-imaging ICU data). The results of classification averaged across 5-folds are summarised in Table 2 using Accuracy, F1-score and AUC (area under the receiver operating characteristic curve).

XGBoost has a number of hyperparameters which we optimised through a grid search within the ranges defined in Table 3. The results of 5-fold cross validation for image-based ResNet-50 approach showed stronger predictive value (AUC 0.840) than the non-imaging ICU data-based XGBoost model (AUC 0.684) with the multimodal approach showed the strongest predictive skill (AUC 0.880). For comparison we showed the results for a model using demographics/metadata only (AUC 0.647) and laboratory and vital sign values only (AUC 0.671).

Examples of radiographs with correct and incorrect classification are given in Fig. 3. Finally, the heat maps from the image-based ResNet-50 model for true positive and true negative instances of cardiomegaly are shown in Fig. 4.

Table 2. Results for the three classes of models considered in this work on the combination of demographics/metadata (D), laboratory and vital sign values (LV) and imaging data (I). The multimodal network, which merges imaging and non-imaging data, has the strongest predictive power showing the overall moderate improvement over ResNet-50. The values show the mean Accuracy, F1-score and AUC (the area under the ROC curve) for the 5-fold cross validation with the corresponding standard deviation.

Model	**Data Type**	**Accuracy**	**F1-score**	**AUC**
Multimodal	D+LV+I	0.837 (0.012)	0.886 (0.009)	0.880 (0.011)
ResNet-50	I	0.797 (0.034)	0.857 (0.028)	0.840 (0.033)
XGBoost	D+LV	0.700 (0.084)	0.771 (0.029)	0.684 (0.046)
XGBoost	LV	0.694 (0.020)	0.769 (0.017)	0.671 (0.036)
XGBoost	D	0.642 (0.030)	0.712 (0.031)	0.647 (0.021)

Table 3. The hyperparameters for each XGBoost model were optimised within the ranges shown below.

Hyperparameters	**Range**
learning rate	[0.0001, 0.1]
max tree depth	[2, 8]
gamma	[0, 2]
colsample	[0, 1]
subsample	[0, 1]

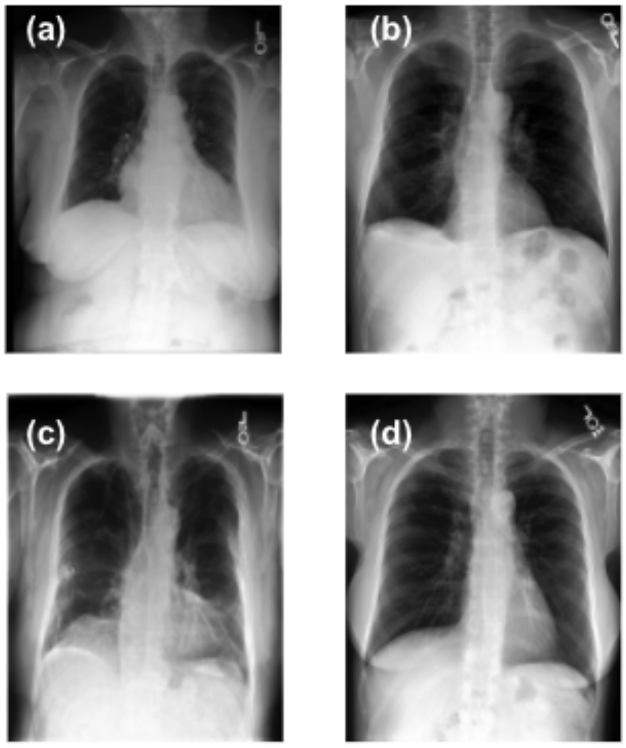

Fig. 3. Examples of X-ray images showing correct and incorrect classification by our proposed joint model: (a) true positive, (b) false negative (clinical review of a radiograph suggests cardiomegaly, based on pattern of cardiac silhouette), (c) false positive (clinical review of a radiograph suggests anatomical distortion by other pathology), (d) true negative.

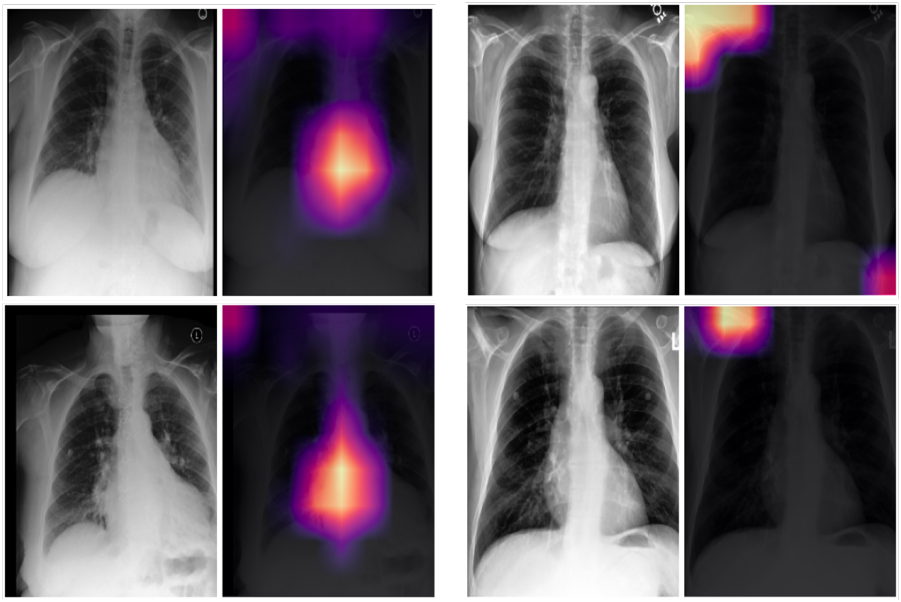

Fig. 4. Examples of heat maps from the image-based ResNet-50 model for true positive (left two columns) and true negative (right two columns) instances of cardiomegaly. The heat maps were created using a Grad-Cam [21] approach and show which areas of the X-ray image were important during prediction.

4 Discussion

4.1 Principal findings

In this work, we proposed a multimodal network which is able to use both images and ICU data to determine the presence of cardiomegaly. Further it provides a proof-of-concept for a multi-modal deep network approach for detecting other cardiopulmonary diseases, in general, from a combination of images, vital sign values, laboratory measurements and other patient metadata. We used two recently published datasets to uniquely combine various modalities for the same patients. This allowed us to build a multimodal approach with the aim to imitate the thought process of a clinician. Clinicians tend to use multiple sources of data to draw their conclusions, unlike most machine learning approaches which have relied exclusively on chest radiographs. Finally, as cardiomegaly usually indicates an underlying pathology that warrants further investigation and clinical management, and can convey significant ill health if unrecognised, the early detection of cardiomegaly is beneficial. The automatic detection of cardiomegaly on radiographs taken for the investigation of other pathologies using these tools could therefore assist as a form of preventative healthcare.

4.2 Strengths and weaknesses of the study

Advantages of using automated systems for X-ray labelling. Current clinical practice is for the requesting clinician or responsible radiologist to review and report the chest radiograph performed. Radiologist availability and clinician time is often a limited resource and tools with the potential to optimise the application of this limited resource have the potential to increase the efficiency of care pathways. A tool which can accurately label chest radiographs as containing either no elements of concern or with relevant labels warranting further review or clinical management would help to optimise the efficient use of clinician time.

Further, when a clinician/radiologist reviews a chest radiograph a potential cognitive bias can occur whereby once a pathology is identified secondary pathologies are not as readily recognised. Automated tools are not subject to this bias and so the use of such tools could help prevent the missed detection of further pathologies on radiographs with multiple elements of concern.

Finally, as cardiomegaly usually indicates an underlying pathology that warrants further investigation and clinical management, and can convey significant ill health if unrecognised, the early detection of cardiomegaly is beneficial. The automatic detection of cardiomegaly on radiographs taken for the investigation of other pathologies using these tools can therefore assist as a form of preventative healthcare.

Imitating an ICU doctor. We use a unique, recently published dataset that combines various modalities for the same patients. This allowed us to build a multimodal approach with the aim to imitate the thought process of a clinician.

Clinicians tend to use multiple sources of data to draw their conclusions, unlike most machine learning approaches which have relied exclusively on chest radiographs.

Automatic labelling accuracy. The labels for our data are automatically generated from free-text reports [15, 9]. It is well known that these types of automatic procedures result in noisy data, which may affect the estimation of the performance of the models [13].

Time between X-rays and ICU data. Although we choose the closest X-ray to each ICU admission the period of time in between is not a constant value. Some X-rays and associated ICU stays will take place closer together in time than other X-rays and ICU stays. This introduces inconsistency into our dataset.

Label bias. The cardiomegaly labels originate from free-text reports derived from X-ray images. This may bias the estimated predictive power in favour of images over signals.

Time series data averaged. As part of the preprocessing we averaged the ICU time-dependent data (e.g. vital signs, laboratory measurement) into summary vectors over the whole stay in order to construct our joint network.

Acknowledgements

GP is supported by an NIHR fellowship. AM and LT are supported by the NIHR Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. BWP acknowledges Rutherford Fund at Health Data Research UK (HDR UK) and Nuffield Department of Population Health (NDPH) Senior Research Fellowship.

References

1. Baltruschat, I.M., Nickisch, H., Grass, M., Knopp, T., Saalbach, A.: Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. Scientific Reports **9**(1), 1–10 (2019). https://doi.org/10.1038/s41598-019-42294-8
2. Bressem, K.K., Adams, L.C., Erxleben, C., Hamm, B., Niehues, S.M., Vahldiek, J.L.: Comparing different deep learning architectures for classification of chest radiographs. Scientific Reports **10**(1), 13590 (2020). https://doi.org/10.1038/s41598-020-70479-z
3. Chang, W., Liu, Y., Xiao, Y., Yuan, X., Xu, X., Zhang, S., Zhou, S.: A machine-learning-based prediction method for hypertension outcomes based on medical data. Diagnostics **9**(4) (2019). https://doi.org/10.3390/diagnostics9040178
4. Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, New York, NY, USA
5. Davies, M.J.: The cardiomyopathies: An overview (2000). https://doi.org/10.1136/heart.83.4.469

6. Deng, J., Dong, W., Socher, R., Li, L.J., Kai L, Li F-F: ImageNet: A large-scale hierarchical image database. pp. 248–255. Institute of Electrical and Electronics Engineers (IEEE) (2010). https://doi.org/10.1109/cvpr.2009.5206848
7. Gessert, N., Nielsen, M., Shaikh, M., Werner, R., Schlaefer, A.: Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data. MethodsX **7**, 100864 (2020). https://doi.org/10.1016/j.mex.2020.100864
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. vol. 2016-Decem, pp. 770–778. IEEE Computer Society (2016). https://doi.org/10.1109/CVPR.2016.90
9. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., 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.: CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In: 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019. vol. 33, pp. 590–597. AAAI Press (2019). https://doi.org/10.1609/aaai.v33i01.3301590, www.aaai.org
10. Ishida, T., Katsuragawa, S., Chida, K., MacMahon, H., Doi, K.: Computer-aided diagnosis for detection of cardiomegaly in digital chest radiographs. In: Medical Imaging 2005: Image Processing. vol. 5747, p. 914. SPIE (2005). https://doi.org/10.1117/12.595755
11. Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L.A., Mark, R.: MIMIC-IV v0.4. Tech. rep. (2020), https://physionet.org/content/mimiciv/0.4/
12. Johnson, A.E.W., Pollard, T.J., Berkowitz, S.J., Greenbaum, N.R., Lungren, M.P., Deng, C., Mark, R.G., Horng, S.: MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. Scientific Data **6**(1), 1–8 (2019). https://doi.org/10.1038/s41597-019-0322-0
13. Johnson, A.E.W., Pollard, T.J., Greenbaum, N.R., Lungren, M.P., Deng, C., Peng, Y., Lu, Z., Mark, R.G., Berkowitz, S.J., Horng, S.: MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs. arXiv (2019)
14. Kingma, D.P., Ba, J.L.: Adam: A method for stochastic optimization. In: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. International Conference on Learning Representations, ICLR (2015)
15. Peng, Y., Wang, X., Lu, L., Bagheri, M., Summers, R., Lu, Z.: NegBio: A high-performance tool for negation and uncertainty detection in radiology reports (2017)
16. Pimentel, M., Mahdi, A., Redfern, O., Santos, M., Tarassenko, L.: Uncertainty-Aware Model for Reliable Prediction of Sepsis in the ICU. In: 2019 Computing in Cardiology Conference (CinC). vol. 45. Computing in Cardiology (2019). https://doi.org/10.22489/cinc.2019.209
17. Pimentel, M.A.F., Redfern, O.C., Malycha, J., Meredith, P., Prytherch, D., Briggs, J., Young, J.D., Clifton, D.A., Tarassenko, L., Watkinson, P.J.: Detecting Deteriorating Patients in Hospital: Development and Validation of a Novel Scoring System. American Journal of Respiratory and Critical Care Medicine (2021). https://doi.org/10.1164/rccm.202007-2700oc
18. Que, Q., Tang, Z., Wang, R., Zeng, Z., Wang, J., Chua, M., Gee, T.S., Yang, X., Veeravalli, B.: CardioXNet: Automated Detection for Cardiomegaly Based on Deep Learning. In: Proceedings of the Annual International Conference of

the IEEE Engineering in Medicine and Biology Society, EMBS. vol. 2018-July, pp. 612–615. Institute of Electrical and Electronics Engineers Inc. (2018). https://doi.org/10.1109/EMBC.2018.8512374
19. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P., Ng, A.Y.: CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv (2017), http://arxiv.org/abs/1711.05225
20. Samak, Z.A., Clatworthy, P., Mirmehdi, M.: Prediction of Thrombectomy Functional Outcomes Using Multimodal Data. In: Communications in Computer and Information Science. vol. 1248 CCIS, pp. 267–279. Springer (2020)
21. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision. pp. 618–626 (2017)
22. Smith, L.N.: Cyclical learning rates for training neural networks. In: Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017. pp. 464–472. Institute of Electrical and Electronics Engineers Inc. (2017). https://doi.org/10.1109/WACV.2017.58
23. Sogancioglu, E., Murphy, K., Calli, E., Scholten, E.T., Schalekamp, S., Van Ginneken, B.: Cardiomegaly Detection on Chest Radiographs: Segmentation Versus Classification. IEEE Access **8**, 94631–94642 (2020). https://doi.org/10.1109/ACCESS.2020.2995567
24. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. vol. 2017-Janua, pp. 3462–3471. Institute of Electrical and Electronics Engineers Inc. (2017). https://doi.org/10.1109/CVPR.2017.369