2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) | 979-8-3503-7149-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/EMBC53108.2024.10781834

Computer-Aided Lung Auscultation Screening and Radiographic Evaluation of Pediatric Pneumonia

Annapurna Kala[†,1], Daniel Chong[†,1], Abdullah H. Baqui[2], Salahuddin Ahmed[3], ASMD Ashraful Islam[3], Nabid Chowdhury[3], Arunangshu Roy[2], Eric D. McCollum[2,4], Mounya Elhilali[1]
[†] Both authors contributed equally
[1] Johns Hopkins University, Dpt. Electrical & Computer Engineering, Lab. for Computational Audio Perception
[2] Johns Hopkins University, Dpt. International Health, Bloomberg School of Public Health
[3] Projahnmo Research Foundation, Dhaka, Bangladesh
[4] Global Program in Pediatric Respiratory Sciences, Eudowood Division of Pediatric Respiratory Sciences, Johns Hopkins School of Medicine, Dpt. Pediatrics, Baltimore, USA

***Abstract*—While chest auscultations provide an accessible and low-cost tool for pediatric pneumonia diagnosis, its subjectivity and low reliability continues to hinder its inclusion in global pneumonia guidelines; eventhough more robust tools like chest radiography also suffer from cost and accessibility issues. Advances in computer-aided analytics is offering more robust tools for interpreting digital auscultation signals though little has been done to explore variations of lung sounds across different chest positions and the correspondence between auscultations and specific radiographic findings. The present study explores interpretation of lung auscultations across chest positions in a pediatric pneumonia population, using a deep neural network classification of normal and abnormal breathing patterns. The results reveal a strong alignment between computer-aided auscultation findings and radiographic interpretation not only in terms of presence of adventitious lung patterns, but also in terms of localizing the abnormality along the left or right lung. Though evaluated in a small clinical population, this research underscores the potential of computer-aided auscultation analysis as a cost-effective substitute for radiography in resource-limited settings.**

I. Introduction

Listening to lung sounds or chest auscultations has often been used to identify pathological acoustic markers that manifest due to the presence of underlying respiratory conditions. Consequently, stethoscopes emerge as a valuable low-resource modality for screening pulmonary pathologies compared to more intricate and costly modalities. However, lung auscultation is notably absent from current World Health Organization pneumonia guidelines[1] due to the considerable subjectivity in sound interpretation. This interpretation is not only hindered because of the complex nature of normal and abnormal breathing sounds, but also the masking effect of ambient sounds that are often picked up by stethoscopes and result in distorting the quality of the auscultation signal. In addition, auscultation from the left side of the chest often has to contend with masking by heart sounds, making positioning of the stethoscope an important factor to access clearer breathing sounds that are not overshadowed by potentially loud heartbeats.

Chest Radiography on the other hand gives a static image with a more established ground truth. While the usual issues of inter-reader variability persists to a certain extent amongst this modality too, its static nature renders it less susceptible to subjectivity compared to auscultations. However, acquiring chest images is generally more challenging and importantly more costly requiring not only expensive equipment but also highly trained technicians making wide access to radiographic modalities a major challenge across the globe [2]. Extensive datasets of chest radiography have paved the way for the application of deep learning techniques in Computer Vision for diagnosing pulmonary diseases [3], [4], [5], [6] . Nevertheless, this progress has not extended seamlessly to pediatric populations [7], highlighting the necessity for population-based transfer learning [8], [9], [10]. Furthermore, within existing radiography collections, efforts to automate diagnosis face numerous challenges stemming from the lack of standardization across devices, technician practices, and reader interpretations [11], [12]

In recent years, the emergence of computerized auscultation analysis has spurred numerous endeavors to automate screening and diagnostics using recorded auscultations [13], [14], [15]. Important contributions have been made to provide more nuance in computer-aided interpretation and develop algorithms that are clinically transparent and explainable [16], as well as address the issues of domain generalization that comes with a medical setting [17]. However, the interpretation of lung sound using mathematical models is often set against a gold standard of expert listeners, typically in the form of a panel of physicians providing their assessment and expert opinion on each auscultation signal. By treating computer-aided auscultation analysis and chest radiography in silo, a question remains regarding the alignment of computerized auscultation models with chest radiography findings. While early results suggest an association between digital recording of lung sounds and radiographic pneumonia particularly in pediatric populations in limited-resource [18], establishing an association between computer-aided auscultation assessment and radiographic findings remains an open question.

This study proposes a pipeline to explore association between data-driven Computer-Aided Auscultation (CAA) models and chest radiography. We conduct a first evaluation on a small dataset and demonstrate promising outcomes in terms of association between radiographic pneumonia findings and uncertainty in CAA outputs as reflected in the posterior probabilities of auscultation models. These findings are then extended to assess the localization of lung abnormalities solely based on digital auscultations. This investigation addresses the specificity of lung auscultations across various chest positions and their capacity to provide informative insights into the presence of localized lung abnormalities within right/left lobes of the lung. The proposed study examines the variability in outcomes of auscultation models and establishes a significant correlation between chest position and radiographic findings in terms of lobe specificity.

II. Data

Clinical Setting: The data used in this work has been collected as part of a study conducted at Projahnmo research site, Sylhet, Bangladesh to robustly assess the impact of introduction of 10-valent Pneumococcal Conjugate Vaccine (PCV10) on Invasive Pneumonical Disease (IPD) in children 0-59 months of age [19]. The goal of this effort was to benchmark the population incidence of IPD prior to the introduction of PCV10 utilizing multiple surveillance modalities such as lung ultrasound, blood work, digital auscultations, and chest radiography. For the current study, we analyzed a subset of 407 subjects with verified records of chest radiography and digital auscultations. All patients analyzed were identified by study physicians as suspected of IPD.

Digital Auscultation: Each subject had their auscultations recorded at 44.1 KHz across two left and two right chest positions using the JHU Stethoscope, a low-cost digital auscultation device that improves lung signal strength by uniformly distributing highly sensitive microphone arrays across the stethoscope diaphragm [20]. These sound recordings were annotated by a panel of expert physicians to indicate the presence or absence of auscultation abnormalities such as wheezes, crackles or both. Each individual recording was randomly assigned to two reviewers, and further assigned to an arbitrator in case of disagreement. In case of abnormality, the reviewers further identified the onset and offset of the anomaly within the recording. Together, the data consisted of 122.83 minutes of normal and 150.52 minutes of abnormal data.

Chest Radiography: Corresponding analogue chest radiography images from the same subjects were collected using portable POLYMOBIL Plus (Siemens, Erlangen, Germany) units and digitised them with CR Fujifilm cassette readers (Tokyo, Japan). These images were again annotated by two primary radiography experts, and then further assigned to a third and fourth reader in case of mismatch [21]. As per definitions of WHO chest radiography findings [22], the present study focused on the final readout of Primary Endpoint Consolidation (PEP) indicating the presence of alveolar consolidation or pleural effusion that is associated with any type of consolidation. Out of the 407 patients, reader panelists agreed on a left, right, both, and no PEP for 360 patients with XRay categorizations as listed in TABLE I . They disagreed in their final annotation on the rest of the 47 patients.

Primary End Point Consolidation	Number of Patients
Left	12
Right	46
Both	6
None	296

TABLE I: Patient Count across different XRay Categories with reader agreement

III. Computer-aided Auscultation Model

Data processing: All digital auscultations included in this analysis were resampled at 8 KHz. The signal underwent noise-cancellation using [23]. Reviewer annotations of auscultation signals varied in length across patients. In order to standardize the computer-aided analysis, all segments were unified at 4 seconds in length. For long-enough segments, the 4 seconds were centered relative to the original onset and offset identified by the expert listeners. For segments that are shorter than 4 seconds, zero-padding was used.

Auscultation data was then divided into a train and test set to train classification models. The train set consisted of 80 percent of patients with no Xray PEP as well as all patients on which the radiography panel disagreed. 20 percent of the patients with no Xray PEP and all the other patients with positive PEP were used for evaluating the model performance. This accounts to 209.83 minutes of training data and 63.52 minutes of testing data.

Classification Models: The study develops a deep learning model to analyze auscultation signals and automatically classify normal lung signals from abnormal breathing patterns such as wheezes and crackles. In this work, we employ an autoencoder model with Convolutional Neural Network (CNN) encoder and Long Short Term Memory network (LSTM) decoder. Given the limited amount of data available in this study, we keep the structure of the network relatively light in order to limit the number of tunable parameters to a manageable range given the training data. The CNN portion of the neural network consists of two 1-dimensional convolutional layers with average pooling after each layer. The first convolutional layer has a filter length of 32, and the second convolutional layer has a filter length of 16. Both layers have a stride of 1 and do not use any additional padding. The output of these mappings is a set of embedding features that are processed through a 2-layer LSTM decoder followed by a fully connected linear layer. The final layer employs a sigmoid activation function and results in a classification posterior value. The network is trained using Binary Cross Entropy loss, and Adam optimizer with a learning rate of .01. All components of the neural network are implemented using PyTorch, and a grid search is performed to determine the optimal set of parameters for the classification task.

The model takes as input a mel-scaled spectrogram of each audio sample, using the python librosa library [24]. The spectrogram is derived using a fast Fourier-transform window

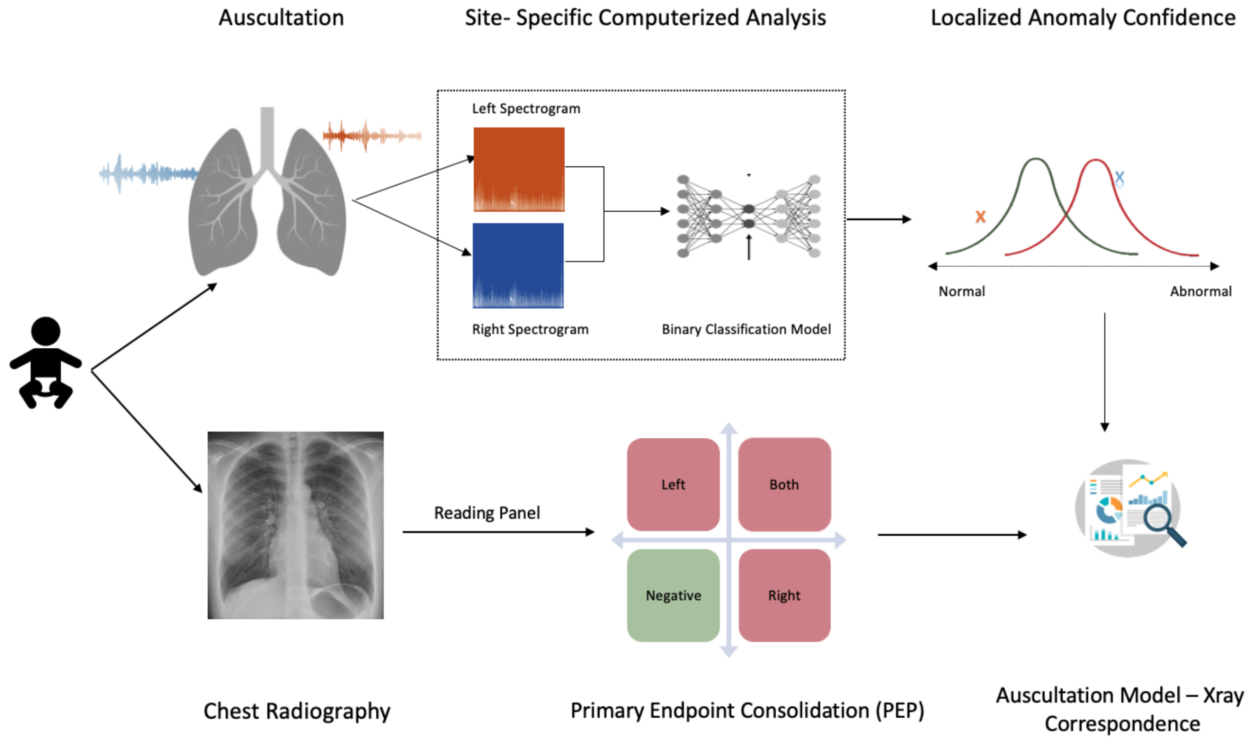

Fig. 1: Overview: In this study, we work with pediatric patients enrolled in a PCV10 incidence study with multiple surveillance modalities. We specifically focus on their auscultation recordings and chest radiography. The bottom branch depicts the chest radiography modality. This modality is annotated by an expert reading panel into one of the four Primary Endpoint Consolidation categories: 1) Left 2) Right 3) Both and 4) None. The top branch visualizes the computerized auscultation analysis. Stethoscope recordings are collected from two left (orange) and two right (blue) chest positions. We map these acoustic recordings to a spectro-temporal representation and pass it through an auscultation classifier model that outputs an abnormality probability for each recording. We establish the correspondence the location specific diagnosis as evaluated by an auscultation based model with the patient's corresponding Xray reading.

of 2048 samples and hop length of 512 and then normalize the data.

Auscultation Site Analysis: The present study explores the specificity of auscultation across site positions and the ability to localize lung abnormalities from breath sounds only. In order to evaluate specificity of chest auscultations from the left versus the right, we train three classifiers: a first model trained on audio samples obtained from the right chest positions only, a second model trained on audio samples from the left chest positions only, and a third model trained on *all* audio samples. Given the limited dataset used for training, particularly with regards to abnormal sound patterns, we augment the dataset using Constrained Synthetic Sampling, an extension to the synthetic minority over-sampling technique [25]. This technique is used to create additional audio samples that are statistically indistinguishable from the original data. By using this augmentation methodology, we are able to increase our training samples to 1800 normal Samples and 1827 abnormal samples for left side samples, 792 normal and 795 abnormal samples for right side samples, and 1990 normal samples and 2104 abnormal samples for the third all-site model. Note that data augmentation is used only for training the models, while a held-out test set is used to evaluate the performance of all models.

IV. Results

A. Agreement between Computer-Aided Auscultation Analysis and XRay findings

While there have been a number of studies that assessed the sensitivity and specificity of digital auscultations relative to a panel of listeners [26], [27], we wanted to evaluate whether a model trained on these audio recordings agrees with radiographic findings in terms of normal/abnormal conclusions regardless of their chest positions. Using the held-out set, we examine the auscultation model trained on both sites and derive aggregate average posterior probabilities obtained for patients flagged as abnormal based on Xray findings versus those identified as normal. Figure 2 shows that the median of auscultation recordings for patients with no primary endpoint consolidation is about 0.336. On the other hand, patients whose XRay findings show consolidation in either/both of the lungs have a higher aggregate auscultation abnormality probability, around 0.803. Given the nonuniform nature of the posteriors distribution, the non-parametric Mann Whitney U

Test confirms that the two posterior distributions for positive and negative end point Consolidation patients is statistically significant, with a p-value of 2.3e-05.

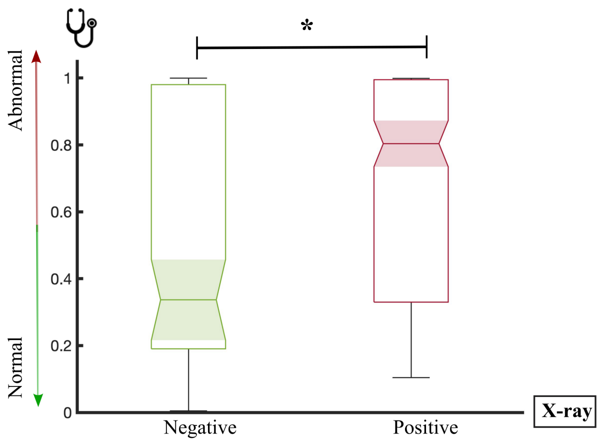

Fig. 2: Average Abnormal Probability of Auscultation Recordings with patients across different Xray Categorizations regardless of chest position. These probability samples of positive and negative Xray categories are statistically significant on performing Mann Whitney U test with a p-value of 2.3e-05.

B. Objective Comparison between Lung Auscultation across sites

One of the open questions explored in the present work is whether left and right chest positions carry sufficiently distinct information about adventitious events in the lung if the abnormality is emanating from a different position relative to the position where the signal is recorded. To first address this question at the signal level, we compare the spectral profiles of left versus right auscultation signals. This analysis follows established signal comparison methods [28] where the average spectral density of left and right auscultations is evaluated over different 500msec windows. Figure 3 reveal a very close correspondence of average signal profiles across both positions. Signal characteristics extracted from these two signals include maximal peak amplitude of the spectrum, the frequency of the peak, the peak width as well as the spectral slope indicating the decay in spectral energy at high frequencies. Performing a statistical comparison (t-test) across both left and right auscultations confirm there is no statistically significant difference between the two (Table II)

Feature	p-value	Mean (Left)	Mean (Right)
Peak Max	.3897	25.064	23.179
Peak Frequency	.0176	108.0 Hz	108.0 Hz
Peak Width	.7834	178.683 Hz	177.882 Hz
Spectral Slope	.5183	-61.006 dB/Hz	-61.012 dB/Hz

TABLE II: Average spectral values and statistical comparison (p-value) of left versus right auscultation sites

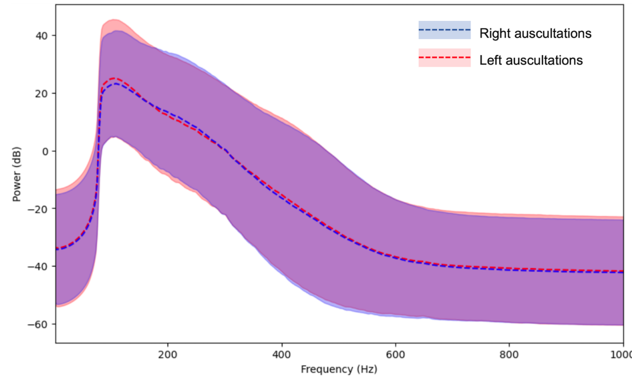

Fig. 3: Comparison of Spectra of left and right auscultations averaged over all patients. Shaded area reflects the standard deviation across all data in each side.

C. Localized Site Specific Analysis between Auscultation and Chest Radiography

Next, focusing on the actual specificity of chest positions, we compare the model outputs separately for left and right auscultations and evaluate how well they reflect the radiographic findings in terms of localized or non-localized consolidation. Note that the auscultation model is still blind to the location-specific information of chest positions. The analysis evaluates the average posterior of left and right chest sites within each Xray category that indicates consolidation and test their agreement with positional Xray findings. We observe that the average posterior abnormal probability of left chest auscultations (0.678) is higher than the right chest auscultations (0.631) for patients with primary end point consolidation in left lung as observed in Fig 4. The same is observed for the patients with right primary end point consolidation (right: 0.725, left: 0.656). We also report the values aggregate left (0.574) and right (0.653) posterior values for Xray Category ”Both”. While the sample distributions are not statistically significant between left and right chest positions across all three categories, we would like to note that the right chest positions show a higher aggregate difference as compared to left chest positions within their respective categorizations. Moreover, the right chest positions also show a higher aggregate behaviour when positive PEP is noted in both the lungs.

We now explore the effect of auscultation models that are attuned to specific chest positions and see if they can better correspond with Xray categories. We work with two different models, one trained only on left auscultations and one trained on right auscultations. We present our findings in Figure 5. In this analysis, we look at the auscultation abnormality probabilities of both left and right chest postions of Xray Left PEP Category patients as inferred by Left Chest Site Classifier. Similarly, probabilities of both chest positions of Xray Right PEP Category are inferred by Right Chest Site Classifier. We now find that the probability samples of left chest positions (mean: 0.88) are significantly higher than the right chest postions (mean:0.603) for PEP left patients with a p-value of 0.0341. Similarly, for patients with Right PEP findings, we note a similar significant difference between right chest samples (mean: 0.74) over left chest samples (0.657) with a p-value of 0.0242.

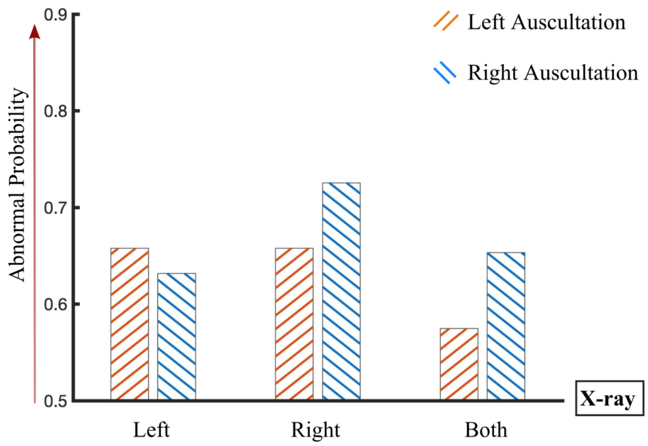

Fig. 4: Comparison of Left (orange forward stripes) and Right (blue backward stripes) Auscultation Abnormal Probabilities across different localized XRay findings as inferred by classifier trained on both side auscultation recordings. Abnormal Probability for the 6 patients with positive XRay findings in both lungs are also reported in the third bar graph (Both).

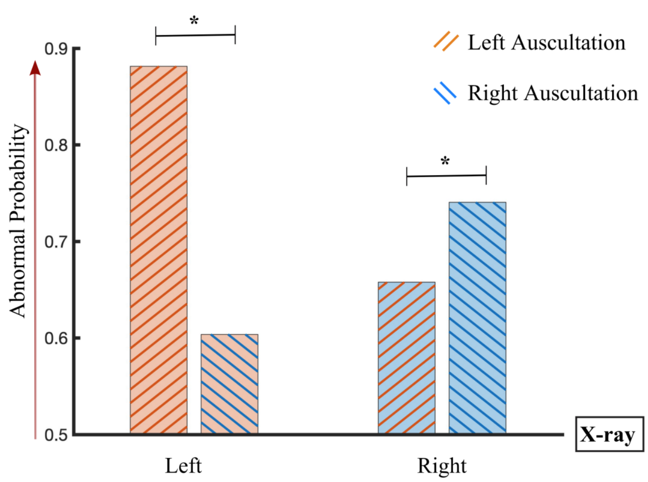

Fig. 5: Left and Right Chest Site comparison in Left Xray Category patients when analyzed by Left-Attuned Auscultation Classifier (highlighted with orange).Left XRay comparative analysis is statistically significant with a p-value of 0.0341. The Chest Site specific comparison in Right Xray Category patients analyzed by Right-Attuned Classifier (highlighted in blue) shows the statisitcal significance with a p-value of 0.0242

V. Discussion

Overall, the results indicate an preliminary positive agreement between X-ray findings and Computer-Aided Auscultation (CAA) along two main outcomes. First, overall predictions of auscultation models reveal a statistically significant trend in normal/abnormal decisions that align strongly with positive and negative readings of an expert radiology panel where positive readings indicate a Primary Endpoint Consolidation (PEP) along the left, right or both lung lobes indicating the presence of alveolar consolidation or pleural effusion. This result strengthens earlier studies in terms of association between digitally recorded lung sounds with WHO-defined radiographic primary end point pneumonia [18]. While establishing this link is an important step towards accepting digital auscultation as a possible tool in child pneumonia care, the present study takes a step further by using computer-aided models that standardize the analysis of lung signals all while maintaining a more nuanced interpretation of the results. This latter point is a critical one given the general approach in machine learning to adopt a correct/incorrect or accuracy-based measure. As discussed in other studies [29], there is a need for robust computer-aided models in diagnostics, but importantly a more graded outlook in model output interpretation. In the present work, we focus on model output as a continuum reflecting a posterior or confidence of evaluating the incoming signal as normal or abnormal and show that this confidence statistically fluctuates towards the abnormal class for patients with positive chest Xray outcomes and fluctuates towards the normal class for patients with negative chest Xray outcomes.

Furthermore, when exploring the localization capabilities of this model, the analysis reveals different outcomes between lung sounds recorded from left versus right chest positions. While an initial evaluation using a generic site-agnostic model was not statistically significant, more refined models trained specifically on right versus left chest position do reveal a clear difference. The results in Figure 5 clearly show that left lung auscultations tend to carry different information about lung abnornmality compared to right positions for patients with radiographic findings indicating a left primary endpoint consolidation; and vice versa. This result is by no means a trivial outcome given that the recorded lung sound using a stethoscope is capturing sound waves that are propagating through the entire chest cavity and as such reflect a much more global view of the wave patterns within the chest [30], [31], [32]. While the presence of an abnormality (e.g. obstruction or closed airways, fluid or secretion) colors the pattern of the sound emanating from the chest, the nature of sound propagation and laws of physics cause these patterns to emanate from all positions. The results in the current study suggest a more pronounced signature of these abnormalities closer to the chest position where the abnormality is identified, though focused primarily on just left versus right positions.

While these results are promising, it is important to emphasize that they are very preliminary. One of the challenges with the current study and studies along this direction is the limited number of subjects undergoing chest radiography, particularly which makes it challenging to utilize evaluation metrics based on strict binary classification. Although the patient numbers underscore the discrepancy in data curation convenience between auscultations and X-rays, especially within the context of large-scale, diverse data collections, our study focuses on analyzing the soft decisions made by computerized auscultation models across different X-ray categories. This analysis aims to provide preliminary insights into the correspondence between computerized models trained on a low-cost, easily collected mode of auscultation surveillance and the relatively more complex and expensive chest radiography modality. Additionally, we delve into the location-specific agreement between the proposed auscultation models and X-rays, underscoring the potential of these acoustic models to objectively assess lung sounds from both left and right chest sites for pathological indicators.

ACKNOWLEDGMENT

Funding for this work was supported by NIH 1R01HL163439.

REFERENCES

[1] WHO, "World Health Organization Integrated Management of Childhood Illness (IMCI) Chart Booklet—Standard," *Geneva (Switzerland): World Health Organization*, no. March, pp. 1–80, 2014.

[2] P. S. Ngoya, W. E. Muhogora, and R. D. Pitcher, "Defining the diagnostic divide: an analysis of registered radiological equipment resources in a low-income African country," *Pan African Medical Journal*, vol. 25, 2016.

[3] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus *et al.*, "CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 590–597, 7 2019.

[4] P. Rajpurkar, J. Irvin, R. L. Ball, K. Zhu, B. Yang *et al.*, "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," *PLOS Medicine*, vol. 15, no. 11, p. e1002686, 11 2018.

[5] R. Singh, M. K. Kalra, C. Nitiwarangkul, J. A. Patti, F. Homayounieh *et al.*, "Deep learning in chest radiography: Detection of findings and presence of change," *PLOS ONE*, vol. 13, no. 10, p. e0204155, 10 2018.

[6] C. Qin, D. Yao, Y. Shi, and Z. Song, "Computer-aided detection in chest radiography based on artificial intelligence: A survey," *BioMedical Engineering Online*, vol. 17, no. 1, pp. 1–23, 8 2018.

[7] S. Padash, M. R. Mohebbian, S. J. Adams, R. D. Henderson, and P. Babyn, "Pediatric chest radiograph interpretation: how far has artificial intelligence come? A systematic literature review," *Pediatric Radiology*, vol. 52, no. 8, pp. 1568–1580, 7 2022.

[8] Y. X. Tang, Y. B. Tang, Y. Peng, K. Yan, M. Bagheri *et al.*, "Automated abnormality classification of chest radiographs using deep convolutional neural networks," *npj Digital Medicine 2020 3:1*, vol. 3, no. 1, pp. 1–8, 5 2020.

[9] E. H. Pooch, P. Ballester, and R. C. Barros, "Can We Trust Deep Learning Based Diagnosis? The Impact of Domain Shift in Chest Radiograph Classification," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12502 LNCS, pp. 74–83, 2020.

[10] A. K. Dubey, M. T. Young, C. Stanley, D. Lunga, and J. Hinkle, "Computer-aided abnormality detection in chest radiographs in a clinical setting via domain-adaptation," *BIOIMAGING 2021 - 8th International Conference on Bioimaging; Part of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2021*, pp. 65–72, 12 2020.

[11] L. A. Eisen, J. S. Berger, A. Hegde, and R. F. Schneider, "Competency in Chest Radiography," *Journal of General Internal Medicine*, vol. 21, no. 5, pp. 460–465, 5 2006.

[12] S. Rajaraman, S. Sornapudi, P. O. Alderson, L. R. Folio, and S. K. Antani, "Analyzing inter-reader variability affecting deep ensemble learning for COVID-19 detection in chest radiographs," *PLOS ONE*, vol. 15, no. 11, p. e0242301, 11 2020.

[13] S. Li, F. Li, S. Tang, and W. Xiong, "A Review of Computer-Aided Heart Sound Detection Techniques," *BioMed Research International*, vol. 2020, 2020.

[14] S. B. Shuvo, S. N. Ali, S. I. Swapnil, T. Hasan, and M. I. H. Bhuiyan, "A Lightweight CNN Model for Detecting Respiratory Diseases from Lung Auscultation Sounds Using EMD-CWT-Based Hybrid Scalogram," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 7, pp. 2595–2603, 7 2021.

[15] G. Petmezas, G. A. Cheimariotis, L. Stefanopoulos, B. Rocha, R. P. Paiva *et al.*, "Automated Lung Sound Classification Using a Hybrid CNN-LSTM Network and Focal Loss Function," *Sensors 2022, Vol. 22, Page 1232*, vol. 22, no. 3, p. 1232, 2 2022.

[16] A. Chaddad, J. Peng, J. Xu, and A. Bouridane, "Survey of Explainable AI Techniques in Healthcare," *Sensors*, vol. 23, no. 2, p. 634, 1 2023.

[17] T. Nguyen and F. Pernkopf, "Lung Sound Classification Using Co-Tuning and Stochastic Normalization," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 9, pp. 2872–2882, 9 2022.

[18] E. D. McCollum, D. E. Park, N. L. Watson, N. S. S. Fancourt, C. Focht *et al.*, "Digital auscultation in PERCH: Associations with chest radiography and pneumonia mortality in children," *Pediatric Pulmonology*, vol. 55, no. 11, pp. 3197–3208, 11 2020.

[19] A. H. Baqui, E. D. McCollum, S. K. Saha, A. K. Roy, N. H. Chowdhury *et al.*, "Pneumococcal Conjugate Vaccine impact assessment in Bangladesh," *Gates Open Research 2018 2:21*, vol. 2, p. 21, 4 2018.

[20] E. West, I. McLane, D. McLane, D. Emmanouilidou, M. Elhilali *et al.*, "Introducing Feelix, a digital stethoscope incorporating active noise control and automatic detection of lung sound abnormalities," *The Journal of the Acoustical Society of America*, vol. 145, no. 3, p. 1923, 4 2019.

[21] E. D. Mccollum, S. Ahmed, N. H. Chowdhury, S. J. R. Rizvi, A. M. Khan *et al.*, "Chest radiograph reading panel performance in a Bangladesh pneumococcal vaccine effectiveness study Paediatric lung disease," *BMJ Open Resp Res*, vol. 6, p. 393, 2019.

[22] T. Cherian, E. K. Mulholland, J. B. Carlin, H. Ostensen, R. Amin *et al.*, "Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies." *Bulletin of the World Health Organization*, vol. 83, no. 5, p. 353, 5 2005.

[23] D. Emmanouilidou, E. D. McCollum, D. E. Park, and M. Elhilali, "Adaptive Noise Suppression of Pediatric Lung Auscultations With Real Applications to Noisy Clinical Settings in Developing Countries," *IEEE Transactions in Biomedical Engineering*, vol. 62, no. 9, pp. 2279–88, 9 2015.

[24] B. Mcfee, C. Raffel, D. Liang, D. P. W. Ellis, M. Mcvicar *et al.*, "librosa: Audio and Music Signal Analysis in Python," *PROC. OF THE 14th PYTHON IN SCIENCE CONF*, 2015.

[25] A. Kala and M. Elhilali, "Constrained Synthetic Sampling for Augmentation of Crackle Lung Sounds," in *Proceedings of the 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2023, p. in press.

[26] Z. Smith, N. Hoekstra, N. McClane, A. Kala, C. Verwey *et al.*, "Computerized lung sound analysis for child pneumonia in low-resource settings: agreement with expert listening panel and bedside auscultation in malawian children," Tech. Rep., 2023.

[27] D. E. Park, N. L. Watson, C. Focht, D. Feikin, L. Hammit *et al.*, "Digitally recorded and remotely classified lung auscultation compared with conventional stethoscope classifications among children aged 1–59 months enrolled in the Pneumonia Etiology Research for Child Health (PERCH) case–control study," *BMJ Open Respiratory Research*, vol. 9, no. 1, p. e001144, 5 2022.

[28] S. Graceffo, A. Husain, S. Ahmed, E. D. McCollum, and M. Elhilali, "Validation of Auscultation Technologies using Objective and Clinical Comparisons," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 7 2020, pp. 992–997.

[29] T. J. Loftus, B. Shickel, M. M. Ruppert, J. A. Balch, T. Ozrazgat-Baslanti *et al.*, "Uncertainty-aware deep learning in healthcare: A scoping review," *PLOS Digital Health*, vol. 1, no. 8, p. e0000085, 8 2022.

[30] P. J. Bishop, "Evolution of the stethoscope," *Journal of the Royal Society of Medicine*, vol. 73, no. 6, pp. 448–456, 6 1980.

[31] L. J. Hadjileontiadis, *Lung Sounds: An Advanced Signal Processing Perspective*. San Rafael, California: Morgan & Claypool Publishers, 1 2008, vol. 3, no. 1.

[32] S. Leng, R. S. Tan, K. T. C. Chai, C. Wang, D. Ghista *et al.*, "The electronic stethoscope," *BioMedical Engineering OnLine*, vol. 14, no. 1, p. 66, 12 2015.