

Severity Quantification of Pediatric Viral Respiratory Illnesses in Chest X-ray Images

Kazunori Okada¹, Marzieh Golbaz¹, Awais Mansoor², Geovanny F Perez³, Krishna Pancham³, Abia Khan², Gustavo Nino³, and Marius George Linguraru^{2,4}

Abstract—Accurate assessment of severity of viral respiratory illnesses (VRIs) allows early interventions to prevent morbidity and mortality in young children. This paper proposes a novel imaging biomarker framework with chest X-ray image for assessing VRI’s severity in infants, developed specifically to meet the distinct challenges for pediatric population. The proposed framework integrates three novel technical contributions: a) lung segmentation using weighted partitioned active shape model, b) obtrusive object removal using graph cut segmentation with asymmetry constraint, and c) severity quantification using information-theoretic heterogeneity measures. This paper presents our pilot experimental results with a dataset of 148 images and the ground-truth severity scores given by a board-certified pediatric pulmonologist, demonstrating the effectiveness and clinical relevance of the presented framework.

I. INTRODUCTION

Viral respiratory infections (VRIs) are a leading cause of morbidity and mortality in the pediatric population worldwide [1]. Although most pediatric VRIs only affect the upper airways (common colds), severe VRIs may involve the lungs and rapidly lead to life-threatening complications. Accordingly, robust tools for severity quantification of lung disease in pediatric VRIs are critically needed to guide early interventions that prevent mortality in this age group. In addition, pediatric lung markers of disease progression in VRIs could also be used as novel phenotypical tools for research and as end-points in clinical trials [2], [3]. It is noteworthy that the development of lung biomarkers in the pediatric population poses distinct challenges because objective pulmonary function testing (i.e. spirometry) is not reliable in young individuals given their inability to follow instructions [4]. Similarly, imaging biomarkers of lung disease based on chest CT have been successfully used in adults [2], [5], [6] but CT scans entail heightened risks for children due to cumulative radiation and need for sedation [7]. In the literature, we are not aware of any previous studies that have investigated the use of lung imaging biomarkers for VRIs in children.

This paper proposes a novel imaging biomarker framework with chest X-ray (CXR) image for assessing VRI’s

severity in infants. We chose CXR as a non-invasive imaging modality because of its lower radiation dosage and wider availability than CT [7]. The proposed framework is designed to quantify the level of *heterogeneity* between intensity distributions from different lung areas, caused by pulmonary air-trapping which is a surrogate of airway obstruction in VRIs [2], [5], [6]. In X-ray images, air-trapping commonly appears as irregularly-shaped areas with intensities darker than surroundings. In order to efficiently quantify such signatures, our method first segments both lung fields using weighted partitioned active shape model and subdivides each field into quadruple areas automatically. Then it quantifies the heterogeneity in each area by computing maximum Kullback-Leibler (KL) divergence of intensity distributions from the target to the other quadruple areas. To further improve the accuracy, we propose a graph cut-based solution with asymmetry constraint to automatically remove large obtrusive objects, such as mechanical support devices, which are often included in CXR images of infants admitted for VRIs. Our implementation is validated by using a dataset that includes 148 CXR images with ground-truth segmentation and the severity scores based on manual assessment of imaging phenotypes due to hyperaeration, demonstrating the effectiveness and clinical relevance of the presented framework.

II. METHOD

A. Lung Segmentation with Weighted Partitioned ASMs

Accurate delineation of lung fields from CXR is challenging due to ambiguous boundaries of lung field, existence of pathologies, superposition of non-target rib bones and heart, anatomical variation of lung shapes and size across subjects, and technical variations (rotation, respiratory phase), especially in children. Previous attempts in the literature for the segmentation of lung field from CXR struggle to accommodate large anatomical and pathological variations found in pediatric CXRs. In addition, state-of-the-art existing methods, such as [8], [9], do not delineate parts of lung field behind aortic arch and apex of heart in CXR and therefore annotate the lung field only partially.

To address these shortcomings, we propose a solution that extends the weighted partitioned active shape model [10] to partition a shape into a set of partial shapes by using the joint shape and appearance sparse learning proposed in [9]. To increase their robustness, we introduce a two-tier hierarchy of landmarks placed along the contour of the lung field border: six primary landmarks defined at salient

¹Department of Computer Science, San Francisco State University, San Francisco, CA

²Sheikh Zayed Institute for Pediatric Surgical Innovation and ³Division of Pulmonary and Sleep Medicine, Children’s National Health System, Washington, DC

⁴School of Medicine and Health Sciences, George Washington University, Washington, DC

locations and secondary landmarks that are automatically set between the primary ones by equidistance interpolation. We initialize the active shape models (ASM) by placing the primary landmarks by using cascade learning classifiers trained with the histogram of oriented gradients (HOG) and the local binary patterns (LBP) features, following [9].

To derive local partitions, the lung shape is divided into overlapping segments with consistent shape variations by performing soft-thresholding using fuzzy c-means clustering, extending the shape sparse learning in [9] so that the variation homogeneous overlapping segments can be derived in a principled fashion without the heuristic used in [9]. The new cost function $\psi_i - |\nu_i| - \gamma |\mu_i|$ for the i -th landmark is given with the number of clusters c^* determined by,

$$c^* = \underset{c}{\operatorname{argmin}} \sum_{i=1}^n \sum_{j=1}^c w_{ij} \|\psi_i - c_j\| \quad (1)$$

where c and n denote the number of partitions and landmarks, c_j and w_{ij} denote the j -th cluster center and the j -th fuzzy membership for i -th landmark, and ν_i μ_i denote spatial mean and variance of i -th landmark across training shapes.

A local appearance model consisting of three features is obtained for each landmark: i) normalized derivatives, ii) tissue intensity probability (second class probability within three class fuzzy c-mean), and iii) elongated rib structure probability based on the vesselness filter (lungs are within the rib cage). Depending upon the image properties, presence of pathologies, and relevant discriminative information in the neighborhood, landmark locations may not be reliable; therefore, each landmark is assigned a weight based on a confidence metric. Landmarks with higher confidence weights have greater contribution in shape deformation. Primary landmarks are assigned the maximum confidence of 1. For secondary landmarks the confidence weight is assigned based on local covariance of the normalized derivatives, tissue intensity probability, and vesselness: $w_i = \frac{1}{1+\operatorname{tr}(\Sigma_i)}$. The model fitting is performed individually for each partition and the optimal position of each landmark is determined by minimizing the Mahalanobis distance. The shape parameters of overlapping landmarks are calculated as mean shape parameters from the two overlapping partitions.

B. Information Theoretic Heterogeneity Quantification

Given binary masks of segmented lung fields, we apply a series of mathematical morphology operations to divide each field into quadruple areas (i.e., upper-proximal (UP), upper-distal (UD), lower-proximal (LP), and lower-distal (LD)) by deriving horizontal and vertical separating lines. The lung border is extracted by subtracting an eroded mask from the original one with a 3-by-3 structuring element. The horizontal line is determined by finding a line parallel to the horizontal image axis which goes through the gravity mass of each field. The vertical line is derived by the thinning operation, finding a medial axis that is equidistance from left and right field borders. At the bottom of each field, the medial axis branches out due to the concavity of diaphragm. This is corrected by determining the lower end-point of

the major medial axis then connect it to the nearest border pixel below it. As a result, a pair of the left and right lung fields are subdivided into 8 regions: $S = \{R_i\} = \{UP_L, UD_L, LP_L, LD_L, UP_R, UD_R, LP_R, LD_R\}$.

We assess the heterogeneity of intensity distributions for each of the 8 regions by using KL divergence $D_{KL}(P||Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$. First, we derive a probability distribution $P_i(x)$ as a normalized intensity histogram for each region R_i . We then compare each region against other three regions in the same field (3-Way) or seven regions from the both fields (7-Way) by using the KL divergence, resulting in either 3 or 7 divergence values. We consider two versions of heterogeneity measure $H_a(R_i)$ and $H_m(R_i)$ by computing arithmetic mean (AVE) and maximum (MAX) of the 3 or 7 divergence values,

$$H_a(R_i) = \operatorname{mean}\{D_{KL}(P_i, P_j) | j \neq i, j \in Q\} \quad (2)$$

$$H_m(R_i) = \operatorname{max}\{D_{KL}(P_i, P_j) | j \neq i, j \in Q\} \quad (3)$$

where Q denotes i) $\{1, \dots, 8\}$ for the 7-Way measure, ii) $\{1, \dots, 4\}$ for the 3-Way measure when R_i belongs to the left field, or iii) $\{5, \dots, 8\}$ when R_i belongs to the right field. Additionally, we define three field-wise measures that aggregate the above region-wise measures over the left field, right field and both fields,

$$H_{a,m}(Left) = \sum_{i=1}^4 H_{a,m}(R_i) \quad (4)$$

$$H_{a,m}(Right) = \sum_{i=5}^8 H_{a,m}(R_i) \quad (5)$$

$$H_{a,m}(All) = \sum_{i=1}^8 H_{a,m}(R_i) \quad (6)$$

where these measures indicate whether air trapping appears or not in the left field, the right field, or this patient case regardless of location, respectively.

C. Large Obtrusive Objects Removal with Asymmetry Constraint

CXR images of pediatric VRI patients often include mechanical diagnostic/treatment devices, such as clips and wires for cardiorespiratory monitoring, tubes for mechanical ventilatory support and aerosolized therapies, and surgical implants (e.g., pacemakers and spinal rods). The following describes our method to remove, from the segmented lung fields, relatively large objects that may strongly influence the quality of the severity quantification to be measured.

Our solution consists of two steps: 1) deriving spatial likelihood weights of the objects via template matching with asymmetric constraints and 2) graph cut segmentation initialized with the weights as its data cost as shown in Fig1. As our working assumptions, we employ an asymmetry constraint in that the target object appears only in one of the two lung fields of a patient, as well as a symmetry constraint in that left and right lung fields are similar. Let S_L and S_R denote respective sets of image coordinates within the segmented left and right lung fields. Two sets of k by k

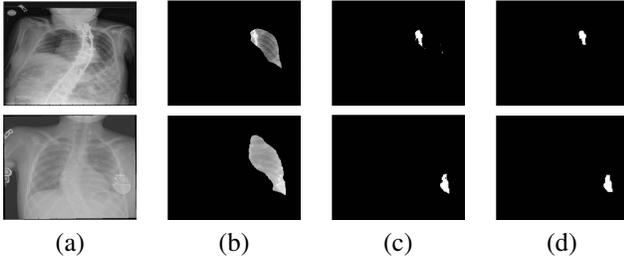


Fig. 1. Illustrative examples of obtrusive object removal. Top row: spinal rods, bottom row: pace maker. (a) original CXR images, (b) segmented lung field, (c) asymmetric thresholded weights, (d) graph-cut segmentation.

patch templates $\{T_L(\mathbf{x}_L)\}$ and $\{T_R(\mathbf{x}_R)\}$ are first generated centered at each location $\mathbf{x}_L \in S_L$ for the left lung and $\mathbf{x}_R \in S_R$ for the right lung. To account for the symmetry constraint, we create another sets of templates $\{\tilde{T}_L(\mathbf{x}_L)\}$ and $\{\tilde{T}_R(\mathbf{x}_R)\}$ by horizontally flipping T_L and T_R across their vertical axis. To compute the likelihood for the left field, each $T_L(\mathbf{x}_L)$ is matched against the set of the flipped templates from the opposite lung $\{\tilde{T}_R(\mathbf{x}_R)\}$. The likelihood weights $w_L(\mathbf{x})$ is then given by the Euclidean distance of the intensity differences with the best matched template,

$$\begin{aligned} w_L(\mathbf{x}_L) &= d(T_L(\mathbf{x}_L), \tilde{T}_R(\mathbf{x}_R^*)), \\ \mathbf{x}_R^* &= \underset{\mathbf{x}_R}{\operatorname{argmin}} d(T_L(\mathbf{x}_L), \tilde{T}_R(\mathbf{x}_R)), \end{aligned} \quad (7)$$

where $d(T_L, T_R)$ denotes Euclidean distance between a pair of k by k intensity arrays. The weights for right lung $w_R(\mathbf{x})$ can then be similarly computed by,

$$\begin{aligned} w_R(\mathbf{x}_R) &= d(T_R(\mathbf{x}_R), \tilde{T}_L(\mathbf{x}_L^*)), \\ \mathbf{x}_L^* &= \underset{\mathbf{x}_L}{\operatorname{argmin}} d(T_R(\mathbf{x}_R), \tilde{T}_L(\mathbf{x}_L)), \end{aligned} \quad (8)$$

Once these weights are computed, we create an initial segmentation of the obtrusive objects by thresholding $w(\mathbf{x})$ exploiting the asymmetry constraint with an empirically determined value $TH = 900$. This is then used to set the data cost of the graph cut segmentation [11] using a matlab wrapper proposed in [12]. The result of this segmentation is used to adjust the heterogeneity measures described in Sec II-B by ignoring pixels within the segmented area for the intensity histogram construction.

III. EXPERIMENTS

A. Image Data

We prepared a dataset of 148 posterior-anterior CXR images with 46 patients, having dimensions 1607 x 1320 pixels with 0.143 mm/pixel and 16-bit gray levels. All images were obtained from a database of cases of acute viral respiratory illnesses in children (0-12 years of age). We selected cases with rhinovirus (n=31) and human metapneumovirus (HMPV; n=15), which are viral pathogens that cause significant respiratory morbidity in young children [1].

B. Ground-Truth

A set of ground-truth (GT) is made both for lung field segmentation and for imaging phenotype scores. For evaluation of lung segmentations, the manual segmentation of lung fields is performed for all 148 cases under the supervision of the two expert pulmonologists. In order to quantify severity of hyperaeration leading to air trapping, basic lucency differentiation across quadrants, representing proximal and distal areas of the upper and lower regions, is evaluated by a board-certified pediatric pulmonologists with expertise in the phenotypical characterization of viral respiratory infections in children [13]. A total of the 8 regions (4 each lung) are independently scored with either 0 (no air trapping), 1 (mild cases), or 2 (severe cases) for a subset of our dataset, comprising of 60 randomly selected cases. The designed three-level score follows the primary features of validated imaging scales for pediatric lung diseases [14], [15].

C. Results

Our weighted partitioned ASM segmentation is evaluated in comparison with a conventional ASM. Each lung field is divided into 6 segments using the fuzzy c-means based solution described in Sec II-A. Average overlap scores against the ground-truth manual segmentation are 0.9091 ± 0.068 and 0.8578 ± 0.608 for the proposed solution and the ASM with mean shape, respectively. Similarly, Euclidean distances of landmark positions between the estimated and manually annotated ground-truth are $1.8821 \pm 0.8612\text{mm}$ and $4.8381 \pm 1.9126\text{mm}$ for our method and the ASM, respectively. For both metrics, we observed significant improvement by the proposed method over the ASM with p-value < 0.001 .

Our obtrusive object detection/removal is also evaluated by using the same dataset. Our dataset contained seven cases that included large mechanical devices (pacemakers and spinal rods, see Fig1) appeared within the lung fields. Our method ($k = 3$) successfully removed 6 out of these 7 cases. Total accuracy out of 148 cases were 89.26% with 85.71% sensitivity and 89.43% specificity.

Finally, we evaluate the proposed severity measures with the 60 cases that come with the severity GT scores. First we evaluate linear correlation for each of the 8 regions, summarized in Table I. Wilcoxon signed-rank test indicates that *MAX* measure in Eq(3) significantly improves *AVE* measure in Eq(3) for both *3-Way* ($p < 0.05$) and *7-Way* ($p < 0.05$) configurations, while no statistically significant improvements are observed between *3-Way* and *7-Way* measures. With *MAX* measures, the best correlation is achieved at the LD region for both left (0.40 for *7-Way* and 0.48 for *3-Way*) and right (0.28 for *7-Way* and 0.25 for *3-Way*) fields, while the worse is achieved at the UD region with values indicating no correlations. On average, both *3-Way* and *7-Way MAX* measures performed similarly. Low to moderate correlations are in part due to highly skewed GT labels where the number of cases without air trapping ranges between 40 to 55 out of 60 total cases across the 8 regions.

We also evaluate aggregating measures shown in Eqs(5-6) by correlating the GT scores that are also aggregated in the

TABLE I

LINEAR CORRELATION COEFFICIENTS OF THE PROPOSED REGION-WISE MEASURES.

	AVE,3-Way	MAX,3-Way	AVE,7-Way	MAX,7-Way
LUP	0.21	0.24	0.18	0.22
LUD	-0.25	-0.15	-0.18	0.05
LLP	0.18	0.23	0.17	0.19
LLD	0.25	0.48	0.18	0.40
RUP	0.18	0.18	0.18	0.18
RUD	-0.04	-0.01	-0.06	-0.03
RLP	0.16	0.21	0.15	0.24
RLD	0.12	0.25	0.05	0.28
Mean	0.10	0.18	0.08	0.19
Median	0.17	0.23	0.16	0.20
Stdv	0.17	0.19	0.14	0.13

TABLE II

LINEAR CORRELATION COEFFICIENTS OF THE PROPOSED FIELD-WISE MEASURES.

	Left	Right	All
3-Way	0.32	0.17	0.31
7-Way	0.28	0.18	0.35

same way over the left, right, and both fields. Each measure can be interpreted as an indicator for air trapping in specific lung field. Table II summarizes the results. Overall, the proposed measure was more effective to detect air trapping in the left field than the right; and the 7-WayMAX measure correlated best with the GT at 0.35 when we aggregated over both fields. In figure 2, illustrative examples of the air trapping cases that are matched with high values of the proposed measures are shown for four exemplar regions. The numbers shown next to the region names are relative region-wise 7-WayMAX measure normalized by the maximum value among the 60 cases used. For areas correlated the best (LLD) to the least (RUD) with GT scores, these examples show reasonable matches with the measure and the clinical GT scores.

IV. CONCLUSIONS AND DISCUSSION

This paper proposed an novel imaging biomarker framework to assess severity of VRIs in infants by quantifying intensity heterogeneity in patients' CXR images. Our pilot experimental results show potential of the proposed methods to offer non-invasive imaging biomarkers that can be used as diagnostic tool. The exact mechanism by which types of respiratory viruses induce pro-asthmatic changes is still largely unknown. This is an important direction of our future work to improve the proposed methods as end-points for clinical trials to investigate this further. Our future work also includes considering more flexible area divisions, evaluating other heterogeneity measures with different divergence metrics, applying machine learning classifiers to develop a predictive system, suppressing non-parenchymal objects in lungs, and experimenting with larger datasets.

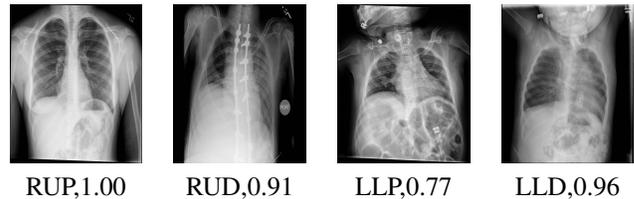


Fig. 2. Illustrative examples of severe air trapping cases with high values of the proposed measures. All cases are rated as level-2 severity except for LLP whose highest GT severity was 1. The numbers shown next to region names are relative region-wise MAX measures normalized by the largest value among the 60 cases used.

ACKNOWLEDGMENTS

This work was partially supported by Grants NHLBI-HL090020 (K12 Genomics of Lung), NICHC-HD001399 (K12 Child Health Research Career Development Award), UL1TR000075 KL2TR000076 Awards from the NIH National Center for Advancing Translational Sciences, SFSU Center for Computing for Life Sciences seed grant, and CTSI-CN Pilot Research Award at Clinical and Translational Science Institute at Children's National.

REFERENCES

- [1] K Hasegawa and et al., "Trends in bronchiolitis hospitalizations in the United States, 2000-2009," *Pediatrics*, vol. 132, no. 1, pp. 28-36, 2013.
- [2] A Busacker and et al., "A multivariate analysis of risk factors for the airtrapping asthmatic phenotype as measured by quantitative CT analysis," *Chest*, vol. 135, pp. 48-56, 2009.
- [3] JJ Smith and et al., "Biomarkers in imaging: realizing radiology's future," *Radiology*, 2003.
- [4] B Vogt and et al., "Pulmonary function testing in children and infants," *Physiol Meas*, vol. 35, no. 3, pp. R59-90, 2014.
- [5] CJ Galban and et al., "Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression," *Nat Med*, vol. 18, no. 11, pp. 1711-5, 2012.
- [6] V Kim and et al., "Clinical and computed tomographic predictors of chronic bronchitis in COPD: a cross sectional analysis of the COPDGene study," *Respir Res*, vol. 15, no. 1, pp. 52, 2014.
- [7] Diana L Miglioretti and et al., "The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk," *JAMA pediatrics*, vol. 167, no. 8, pp. 700-707, 2013.
- [8] S Candemir and et al., "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," *IEEE Trans. Medical Imaging*, vol. 33, no. 2, pp. 577-90, 2014.
- [9] Yeqin Shao and et al., "Hierarchical lung field segmentation with joint shape and appearance sparse learning," *IEEE Trans. Medical Imaging*, vol. 33, no. 9, pp. 1761-1780, 2014.
- [10] Xue Yang and et al., "Weighted Partitioned Active Shape Model for Optic Pathway Segmentation in MRI," in *Clinical Image-Based Procedures. Translational Research in Medical Imaging*, pp. 109-117. Springer, 2014.
- [11] Y Boykov and et al., "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 23, no. 11, pp. 1222-1239, 2001.
- [12] Shai Bagon, "Matlab wrapper for graph cut," December 2006.
- [13] SM Restrepo-Gualteros and et al., "Characterization of cytomegalovirus lung infection in non-HIV infected children," *Viruses*, vol. 6, no. 5, pp. 2038-51, 2014.
- [14] S Terheggen-Lagro and et al., "Correlation of six different cystic fibrosis chest radiograph scoring systems with clinical parameters," *Pediatr Pulmonol*, vol. 35, no. 6, pp. 441-5, 2003.
- [15] SM Shin and et al., "Bronchopulmonary dysplasia: new high resolution computed tomography scoring system and correlation between the high resolution computed tomography score and clinical severity," *Korean J Radiol*, vol. 14, no. 2, pp. 350-60, 2013.