IMAGE CLASSIFICATION IN THE CLOUD HELPS FIGHTING THE EFFECTS OF HYPEROXIA

Christine Bui[1,2], Steven Cho[1,2], Ina Rudloff[1,2], Marcel F. Nold[1,2], Claudia A. Nold-Petry[1,2], Kirstin Elgass[3]

[1] Ritchie Centre, Hudson Institute of Medical Research, Australia

[2] Department of Paediatrics, Monash University, Australia

[3] Carl Zeiss Pty Ltd, Research Microscopy Solutions, Australia

Contact: Kirstin Elgass, Service and Application Engineer, Carl ZEISS Pty Ltd, Research Microscopy Solutions, Australia, kirstin.elgass@zeiss.com, www.apeer.com

ABSTRACT

Bronchopulmonary dysplasia (BPD) is a chronic inflammatory lung disorder which can develop upon prolonged mechanical ventilation and supplemental oxygen therapy as often required by prematurely born infants. These vital interventions can lead to injury and inflammation of the underdeveloped lungs of preterm babies, resulting in disruption of normal alveolar and pulmonary vascular growth [1]. A severe complication that can arise due to the reduction of blood vessels in the lung is pulmonary hypertension secondary to BPD (BPD-PH), resulting in increased pulmonary vascular resistance and vessel remodeling [2]. Both, BPD and BPD-PH, are grave diseases but despite years of medical research, no safe and effective treatments for BPD and BPD-PH are currently available.

INTRODUCTION

The mechanisms of impaired pulmonary vascular development in preterm infants leading to BPD-PH is poorly understood. Therefore, researchers at Monash University and the Hudson Institute of Medical Research in Melbourne aim to assess the characteristics and degree of damage in BPD- and BPD-PH-affected lungs along with screening for drugs that can potentially prevent the development of both diseases. To this end, a clinically relevant mouse model is being used, as previously published [3-5]. Briefly, pregnant mice are injected with the bacterial cell wall component lipopolysaccharide (LPS) to induce maternal inflammation, and then immediately after birth the pups are raised either under normal oxygen conditions (21% oxygen; hereafter called AIR) resulting in healthy lungs, or under hyperoxia conditions (65% oxygen; hereafter called HYP) resulting in damaged lungs with a phenotype similar to BPD/BPD-PH affected lungs of babies. The lungs are then stained in Lugol's iodine solution and imaged at the

synchrotron using x-ray computed tomography (CT) to assess the 3D structure of the pulmonary vascular bed.

Image analysis of lung CT data sets is still challenging though, as there is no commercial software available that addresses all needs, and the data sets are large (tens to hundreds of Gb) due to a large volume (up to 1 cm^3) being recorded at very high resolution (up to 1 µm). Especially for drug screening purposes, automated classification into healthy and sick lungs would be a suitable, time-efficient and unbiased approach.

Machine learning algorithms are designed to handle huge amounts of data automatically, and assign - after an initial learning process - a specific state to each of the analyzed data sets. Combined with a cloud-based digital solutions platform such as APEER, it is the ideal solution for classification of 3D CT data sets for screening purposes. APEER allows the user to develop and combine multiple modules (generally performing a single task only) into workflows that can be customized to address challenging research problems for which no commercial solution is currently available. APEER is equipped with a powerful GPU processing unit suitable for processing extremely large data sets (hundreds of Gb).

DEEP LEARNING IN BPD RESEARCH

To advance BPD research, we developed a machine learning algorithm that classifies lung CT images (Fig. 1A) into a healthy or sick status, and made it publicly available on APEER. The modules use APEER's GPU processing units and can be trained in just a few minutes on 2000 images to an accuracy of >99.8%. Once trained, the model provides near-instant classification of the analyzed lung's health state, thus completely omitting the need to pre-process, visualize and analyze the images.

Lung Classifier and Prediction Module on APEER

Our APEER modules 'Image Classifier – Train' and 'Image Classifier – Predict' are based on Python's Tensorflow [6] and Keras [7] packages, thus using well-established open-source neural network libraries. Keras neural networks generally consist of various layers. The Image Classifier modules use convolutional layers to extract features of the input images in combination with pooling layers to down-sample and summarize these features (Fig. 1B). A subsequent flattening step then reduces the dimensionality of the resulting feature maps before a dense layer can be applied which mathematically corresponds to a rotation, scaling and translation transformation of the input images. Finally, a dropout layer is used in

which some dimensions are randomly set to zero. Dropout is used during training to prevent overfitting by artificially adding noise to the input data.

Quality Control of the Trained Network

The Image Classifier Training module only uses 67% of the provided training data for training. The remaining 33% of the training data is set aside for validation of the trained network. After each training iteration (epoch), the validation data is being used to determine the current accuracy of the network (Fig. 1C, validation accuracy or val_acc). Thus, the validation accuracy values are a measure of how good the predictions of a trained model are (0.5 = 50% accuracy, random classification; 1.0 = 100% accuracy, perfectly trained model). Training should be stopped when the validation accuracy stops increasing, otherwise overfitting of the model is likely to occur. A model becomes overfitted when during training the neural network algorithm extracted too many parameters from the training data (considering noise as real data). Consequently, while the model will perform well on the training data set, it will perform poorly when exposed to new, unknown data.

PROSPECT

Discovering and Developing Drugs against BPD

The described BPD/BPD-PH research project has already identified one potential drug against BPD: interleukin-1 receptor antagonist (IL-1Ra). IL-1Ra is a protein which exhibits anti-inflammatory effects by inhibiting the activity of interleukin-1. We successfully demonstrated that IL-1Ra protects against murine BPD [3,4] and BPD-PH [5]. With the combination of artificial intelligence (machine learning) and the cloud-based computational power of APEER, BPD/BPD-PH research can be taken to the next level by analyzing a wide range of lung images (including CT images but also electron microscopy, histological or immunofluorescence images) with high-throughput in a correlative approach.

Potential Applications of the APEER Image Classifier Modules

APEER as an open-source platform brings high-end computational image analysis to all researchers in an easy-to-use, flexible and time-efficient manner. Although our Image Classification modules were originally designed for classification of 3D lung CT data sets into 'healthy' or 'sick' states, they can be used to classify any image type (2D, 3D, monochrome, color) into any two distinct states. Therefore, other applications that require image classification, such as quality control of materials and detection of damaged materials/parts, can immediately use and profit from the Image Classifier modules on APEER.

1) Speer CP. *Pulmonary inflammation and bronchopulmonary dysplasia*. J. Perinatol. (2006) 26 (Suppl. 1):S57–62; discussion S63–54. doi: 10.1038/sj.jp.7211476

2) Mourani PM, Abman SH. *Pulmonary vascular disease in bronchopulmonary dysplasia: pulmonary hypertension and beyond*. Curr Opin Pediatr. 2013;25:329–337

3) Nold MF, Mangan NE, Rudloff I, Cho SX, Shariatian N, Samarasinghe TD, Skuza EM, Pedersen J, Veldman A, Berger PJ, Nold-Petry CA. *Interleukin-1 receptor antagonist prevents murine bronchopulmonary dysplasia induced by perinatal inflammation and hyperoxia*. Proc Natl Acad Sci USA. (2013) 110:14384–9. doi: 10.1073/pnas.1306859110

4) Rudloff I, Cho SX, Bui CB, Mclean C, Veldman A, Berger PJ, Nold MF, Nold-Petry CA. *Refining anti-inflammatory therapy strategies for bronchopulmonary dysplasia*. J Cell Mol Med. (2017) 21:1128–38. doi: 10.1111/jcmm.13044

5) Bui CB, Kolodziej M, Lamanna E, **Elgass K**, Sehgal A, Rudloff I, Schwenke DO, Tsuchimochi H, Kroon MAGM, Cho SX, Maksimenko A, Cholewa M, Berger PJ, Young MJ, Bourke JE, Pearson JT, Nold MF, Nold-Petry CA. *Interleukin-1 Receptor Antagonist Protects Newborn Mice Against Pulmonary Hypertension.* Front Immunol. (2019) Jul 11;10:1480. doi: 10.3389/fimmu.2019.01480.

6) https://www.tensorflow.org/

7) https://keras.io/

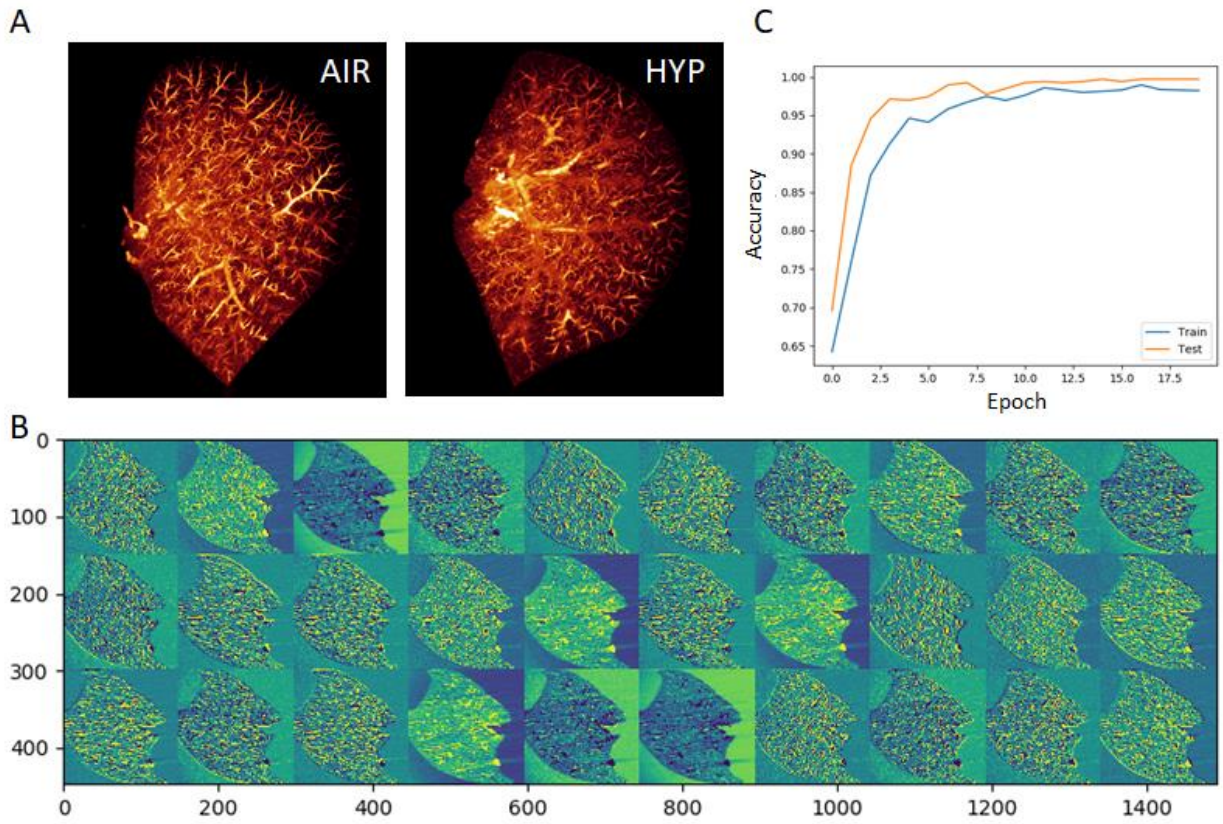

Figure 1: Training the Image Classifier Model. A) 3D Lung CT data sets of healthy (AIR) and sick (HYP) lungs. B) Visualization of Keras layers the neural network algorithm uses for learning. C) Validation Accuracy of the trained model after each epoch of training.

Scan to see the workflows on APEER