SIIM CMIMI24
Oct 21-22 | Boston, MA
Conference On Machine Intelligence In Medical Imaging

Machine Learning Clustering of Qualitatively Assessed Lung Computed Tomography Scans to Distinguish Nonhuman Primate Models of Respiratory Virus Infection: A Pilot Study

Edmond Adib, PhD, National Institute of Health; Shiva Singh, PhD; Marcelo Castro, PhD; Mark Rustad, PhD; Winston Chu, PhD; Maryam Homayounieh, MD; Gabriella Worwa, PhD; Daniel Chertow, MD, MPH; Reed Johnson, PhD; Michael Holbrook, PhD; Yu Cong, MD, MS; Ian Crozier, MD; Ashkan Malayeri, MD; Jeffrey Solomon, PhD

Introduction/Background

We explore unsupervised machine learning (ML) to analyze expert-generated qualitative assessments of lung computed tomography (CT) scans to differentiate nonhuman primate (NHP) models of experimental respiratory virus infections.

Methods/Intervention

We utilized CT scans from four distinct experiments evaluating NHP infection models after cowpox virus (CPV), influenza A virus (IAV; with and without superimposed methicillin-resistant staphylococcal [MRSA] exposure), Nipah virus (NiV), and SARS-CoV-2 exposures (Figure 1). N=19 subjects with imaging abnormality across multiple time points were selected. While CT protocols were controlled, NHP species, age, weight, and dose/route of inoculation varied across experimental groups (Table 1). Using a standardized evaluation questionnaire, a radiology specialist qualitatively graded each CT lung-lobe. Features were one-hot encoded, and Uniform Manifold Approximation and Projection (UMAP) was applied for dimensionality reduction followed by k-means clustering.

Results/Outcome

A UMAP plot (Figure 2) demonstrates the grouping of CT scan qualitative features from different NHP models into clusters revealing key insights: • # Clusters: Six-clustering analysis generated the highest silhouette score (SS= 0.575). Two clusters express peak vs. non-peak disease and another contains an individual (subject 12) which had pre-existing lung abnormality. • Exposure Clustering: CT abnormalities from IAV +/- MRSA and SARS-CoV-2 virus models cluster together suggesting similar qualitative lung features across these models. CT abnormalities after CPV exposure consistently clustered separately, suggesting distinct qualitative features in this model. • Longitudinal Variability: after a specific viral exposure (e.g. IAV + MRSA), peak CT abnormality (cluster 0 = blue) clustered distinctly versus non-peak (cluster 3 = red) abnormality, consistent with the expected time-series analysis.

Conclusion

Lung lesion phenotypes likely vary across viral infections, routes of inoculation, dose and other factors. Using a radiologist's lobe-based qualitative assessment, ML methods can effectively distinguish differences. Future efforts with more subjects will explore fully automated methods (e.g. lung segmentation, radiomic feature extraction) as input to machine learning-based classification.

Statement of Impact

Differentiating qualitative CT lung abnormality across NHP models of viral infections provides initial proof-of-principle that urges ML approaches of user-independent radiomic feature analysis in the future.

Virus Type	*Route of Inoculation*	*Species*	*Age Range (y)*	*Number of Subjects*	*Days Post-Exposure*
Cowpox virus	Aerosol	Rhesus monkey	3-8	6	4-29
Nipah virus	Aerosol	African Green monkey	6-8	2	3-56
SARS-CoV-2	Intrabronchial + Aerosol	Crab-eating macaque	4-10	6	2-28
Influenza A virus +/- MRSA	Intrabronchial	Rhesus monkey	2-4	5	2-14

Table 1 - Experimental Conditions

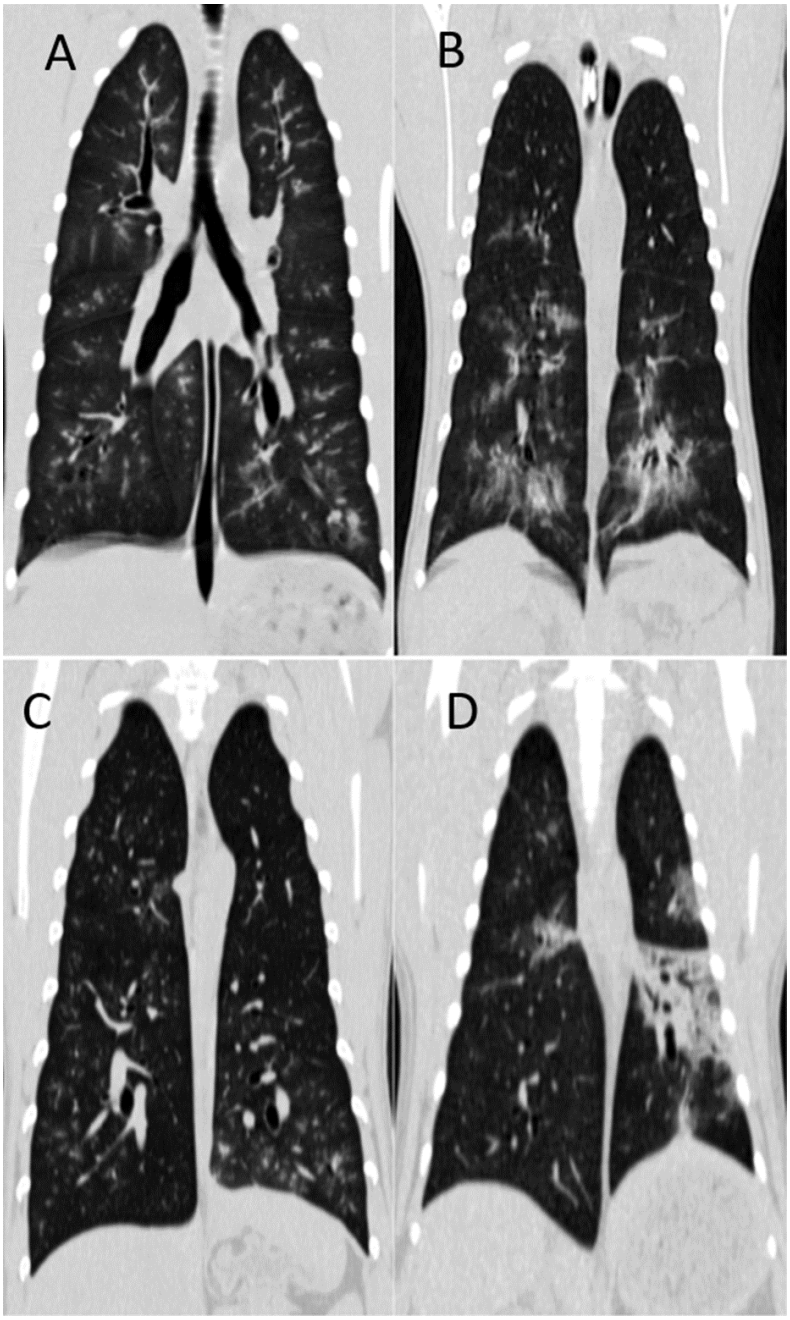

Figure 1 – Representative NHP Lung CT Scans from with four different infection models: A) Cowpox virus (DPE=7), B) influenza virus + MRSA (DPE=10), C) Nipah virus (DPE=3), D) SARS-CoV-2 (DPE=4)

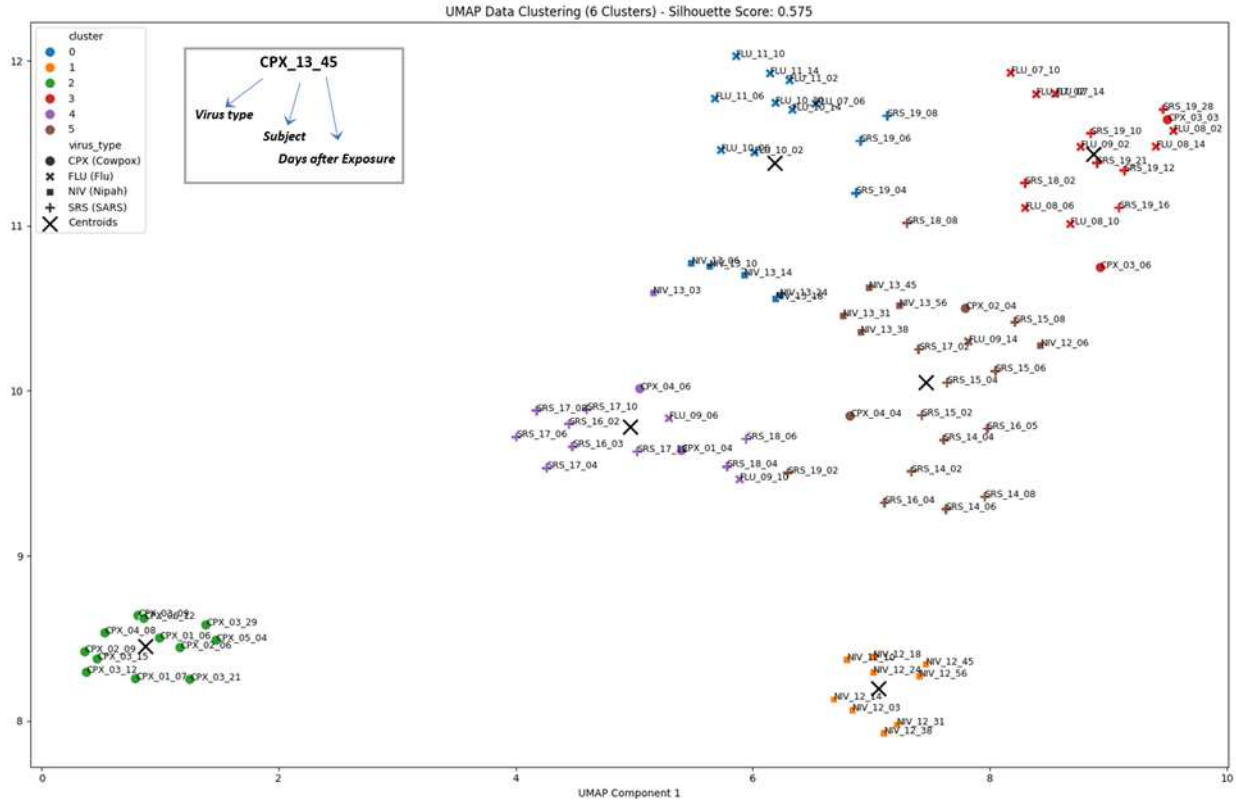

Figure 2 - UMAP and k-means Clustering of Qualitative CT Scan Features by NHP model type

Keywords

SARS-CoV-2; Nipah virus; cowpox virus; influenza virus; UMAP; k-means clustering