COV-AID: An Efficient and Interpretable Machine Learning Ensemble for COVID-19 Diagnosis, Prognosis, Abnormality Localization, and Severity Assessment

Garg, Aksh (School: Palos Verdes Peninsula High School)

COVID-19 has infected over 114 million people as of March 1, 2021, pushing the world towards social and economic collapse. With cases rising rapidly and the current gold-standard for diagnosis—RT-PCR—suffering from high false-negative rates, identifying alternate diagnosis and prognosis methods is pressing. This project designed, evaluated, and compared 40 convolutional neural network (CNN) architectures for COVID-19 diagnosis from chest-CT images, serving as the first to employ the EfficientNet family and intermediate activation maps for COVID-19 diagnosis. Subsequently, innovative methods for localizing 14 radiological abnormalities were developed using YOLOv5 and a weighted-boxes fusion algorithm. Shallow 3D CNNs were designed to triage patients into five categories based on severity. Finally, gradient boosting machines, random forests, decision trees, SVMs, and neural networks were built to forecast patient mortality using biomarkers from blood samples. EfficientNetB5 was the best radiological model with an accuracy of 0.9931±0.0021, F1 score of 0.9931±0.0020, sensitivity of 0.9952±0.0020, and specificity of 0.9912±0.0048. Intermediate activation maps and gradient-weighted class activation mappings offered human-interpretable evidence of the model's perception of ground-glass opacities, consolidations, and paving-patterns. YOLOv5 localized thoracic abnormalities with an intersection-over-union score 0.84, the shallow CNN attained a triaging accuracy of 0.823±0.028, and XGBoost got accuracies and F1-scores of 0.974±0.018 and 0.972±0.020, respectively. COVAID offers a scalable and interpretable tool with diagnosis, prognosis, and severity assessment capacities. Access to assembled datasets and computation boosts can further optimize triaging and localization.

Awards Won:

Second Award of $2,000