

DIRECTORY FILE STRUCTURE

covid_siteIDs.txt

Text file listing each dataset with its assigned ID number used throughout the project

- Each site has a Quality Control (QC) file listing images that may require review due to low quality.
- Each site has an ID mapping file that maps the file names on the COVID-ARC server to that of the original datasets.

DATA STRUCTURE

Original data aggregated from various sites

	Location	Modality	File Format	Metadata	Number of Images		Number of Patients		Public	
					COVID-19	Non-COVID-19	COVID-19	Non-COVID-19		
Site 1	Sao Paulo, Brazil	CT	.png		2168	Other pulmonary illnesses 1247 Healthy 758	80	Other pulmonary illnesses 80 Healthy 50	Yes	
Site 2	China	CT	.png	Patient details including age, sex, city, medical history, time of scan, CT scan abnormalities, and comorbidities	349	463	216	55	Yes	
Site 3	Italy	CT	.nii	Patient details including age, sex, clinical symptoms, hospital, medical history, and CT scan abnormalities and comorbidities	110		60		Yes	
Site 4	Wenzhou, China and Netanya, Israel	CT	.nii		20				Yes	
Site 5	Valencian Region Medical Bank	CT X-Ray	.tgz		CT 900 X-Ray 16800				Yes	
Site 6	Moscow, Russia	CT	.nii	Patients are split into 5 categories based on severity of disease	1110		1110		Yes	
Site 7	Sari, Iran	CT	.tif	Patient details including age, sex, diagnosis, date of scan, and location of scan	15589	48260	95	282	Yes	
					2282	Subset for Machine Learning		95	282	Yes
Site 8	Various locations	X-Ray	.png	Patient details including age, sex, diagnosis, severity, modality, date of scan, and location of scan abnormalities and comorbidities	CT 84 X-Ray 857				Yes	
Site 9	University of Electronic Science & Technology, China	CT	.nii		120		120		No	
Site 10	China Consortium of Chest CT Image	CT	.png	Patient details including age, sex, critical illness, liver function, lung function, and time of progression.	159702	Pneumonia 156071 Normal 95756	408 *Patient group not specified		Yes	
Site 11	Hannover, Germany	X-Ray	.nii	Patient details including sex, time of progression, and clinical findings. Age, size, and weight will be published in the future.	243				Yes	
Site 12	Various locations	X-Ray	.png	Patient details including sex, time of progression, and clinical findings. Age, size, and weight will be published in the future.	3,616	Viral Pneumonia 1345 Normal 1341 Lung Infection 6,012			Yes	
Site 13	Spain	None available	.csv	Patient details including age, sex, diagnosis, progression of disease, lab results, treatment, and patient outcome.					No	
Site 14	Various locations	None available	.csv	Daily levels of COVID-19 cases, recoveries, and deaths worldwide.					Yes	
Site 15	South Korea	None available	.csv	Patient epidemiological data, time series data, and regional data based on affected cases.					Yes	
Site 16	Various locations	Ultra-sound	Videos .gif mov, mp4 Images .png, .jpg	Patient details including diagnosis, symptoms, sex, age, and lung abnormalities. Image details including framerate, frame lengths, and resolution.	Bacterial/Viral Pneumonia Healthy				Yes	
					Videos 50 Images 22	57 22	73 15			
Site 17	Radiological Society of North America	CT	.nii		632		632		Yes	
Site 18	Arkansas, USA	CT, DX, X-Ray	.dcm	Patient details including age, weight, height, race, sex, comorbidities, lab data, and patient outcome	31,935		105		Yes	
Site 19	Radiological Society of North America	CT	.dcm		31,856		110		Yes	
Site 20	Radiological Society of North America	X-Ray	.dcm			29,684			Yes	
Site 21	Guangzhou, China	X-Ray	.jpg			Pneumonia 4,273 Normal 1,583			Yes	
Site 22	Los Angeles, CA								No	
Site 23	Various locations	None available	.cvs	Neurological complications accompanying COVID-19 as well as demographic information of patients.		581			Yes	
Site 24	China Consortium of Chest X-ray Image Investigation	X-Ray	.csv			161,398 includes suspected COVID-19			Yes	
Site 25	Catanzaro, Italy	CT	.nrrd	Type of scanner used, patient age, clinical score for baseline and follow-up.	51,247		50		Yes	
Site 26	Pakistan	EEG	.jpg					Cardiovascular Disease 828 Normal 859	Yes	
Site 27	University of California, San Diego	X-Ray	.h5 .ipynb			Pneumonia 524 Normal 752		Pneumonia ~400 Normal ~576	No	

ANALYSIS STRUCTURE

Image Processing Outputs

	Location	Available Analysis	File Format	Number of Files		Number of Patients	
				COVID-19	Non-COVID-19	COVID-19	Non-COVID-19
Site 1	Sao Paulo, Brazil	Lung masks Segmented images	.nii .png	1248 1248	1219 1199		
Site 2	China	Lung masks	.png	349	463	216	44
Site 3	Italy	COVID-19 masks Lung masks	.nii .nii	110 110		60 60	
Site 4	Wenzhou, China and Netanya, Israel	COVID-19 masks Lung masks Lung + COVID-19 masks	.nii .nii .nii	20 20 20		20 20 20	
Site 5	Valencian Region Medical Bank	Annotated COVID-19 masks	.tgz	23			
Site 6	Moscow, Russia	COVID-19 masks	.nii	50		50	
Site 7	Sari, Iran	None Available					
Site 8	Various locations	Lung masks	.png	CT 84 X-Ray 857			
Site 9	University of Electronic Science and Technology, China	Annotated COVID-19 masks	.nii	120		120	
Site 10	China Consortium of Chest CT Image Investigation	COVID-19 masks	.nii	750		150	
Site 11	Hannover, Germany	None Available					
Site 12	Various Locations	None Available					
Site 13	Spain	None Available					
Site 14	South Korea	None Available					
Site 15	Various Locations	None Available					
Site 16	Various Locations	None Available					
Site 17	Radiological Society of North America	Detailed annotation	.json				
Site 18	Little Rock, Arkansas	Detailed annotation	.json				
Site 19	Radiological Society of North America	Detailed annotation	.json				
Site 20	Radiological Society of North America	None Available					
Site 21	Guangzhou, China	None Available					
Site 22	Los Angeles, CA	None Available					
Site 23	Various locations	None Available					
Site 24	China Consortium of Chest X-ray Image Investigation	None Available					
Site 25	Catanzaro, Italy	None Available					
Site 26	Pakistan	None Available					
Site 27	UCSD	Heat maps			Pneumonia 524 Normal 752		Pneumonia ~400 Normal ~576