



Predicting and Communicating Outcome of COVID-19 Hospitalizations with Medical Images and Clinical Data

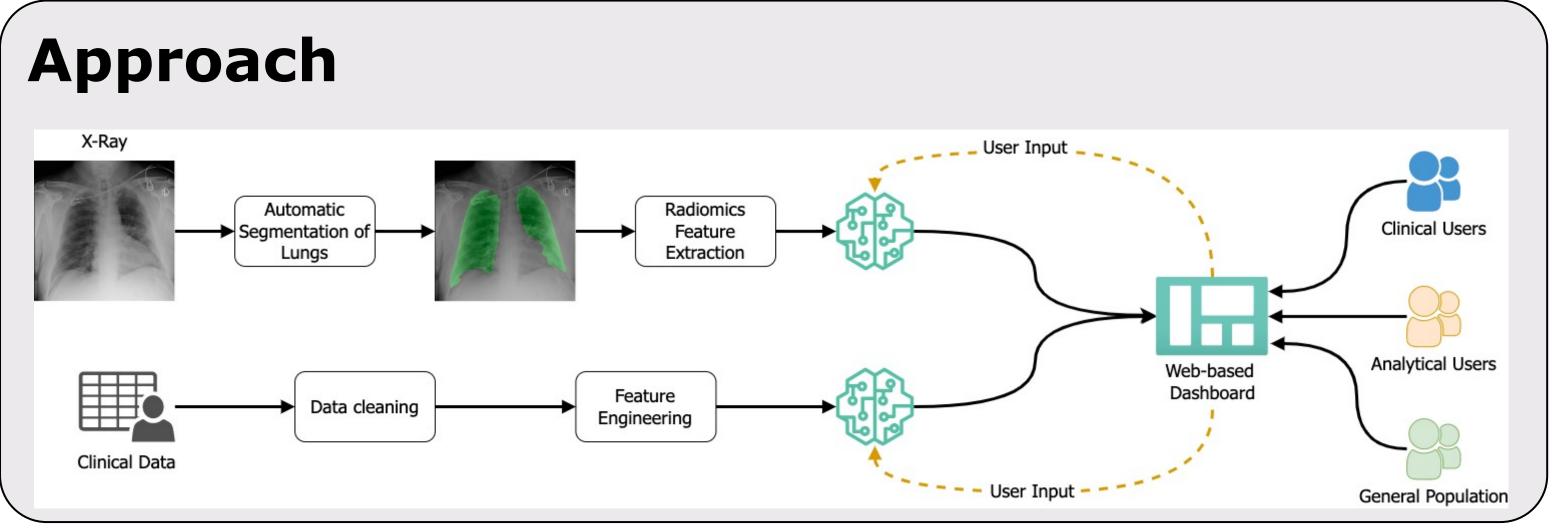
Oliver Stritzel Data Science TU Wien Informatics Institute of Visual Computing and Human-Centered Technology Research Unit of Computer Graphics Supervisor: Assistant Prof. Dr. Renata Raidou

Problem & Motivation

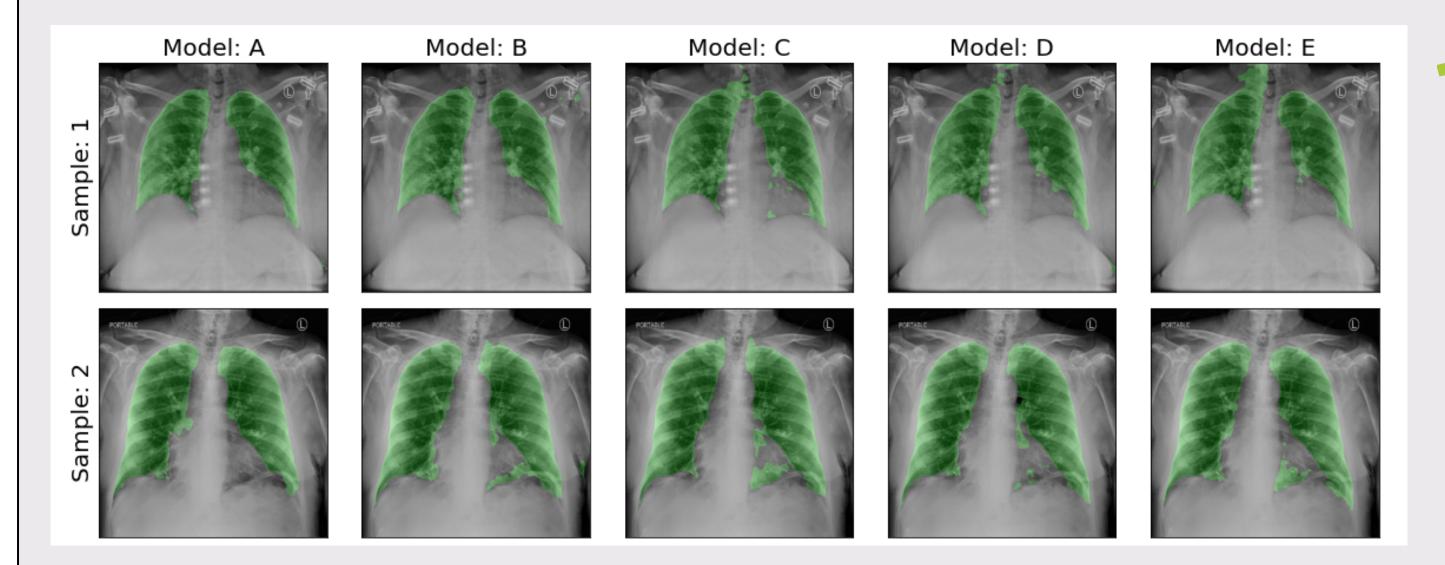
The COVID-19 pandemic put pressure on several public institutions, mainly the health-care system. Besides the danger of the actual illness, there also has been a rise of fake news and alternative facts. It is of increased importance to communicate novel scientific insights to numerous possible user groups, including laypeople as well as medical and analytical experts in an effort to **increase risk perception** and provide **decision making support**.

Goals and Research Questions

We propose a novel **visual analytics application** which communicates insights into COVID-19 hospitalization outcome prediction using the publicly available Stony Brooks COVID-19 dataset. The dataset includes electronic health data as well as medical image data in form of patients chest X-rays. We face

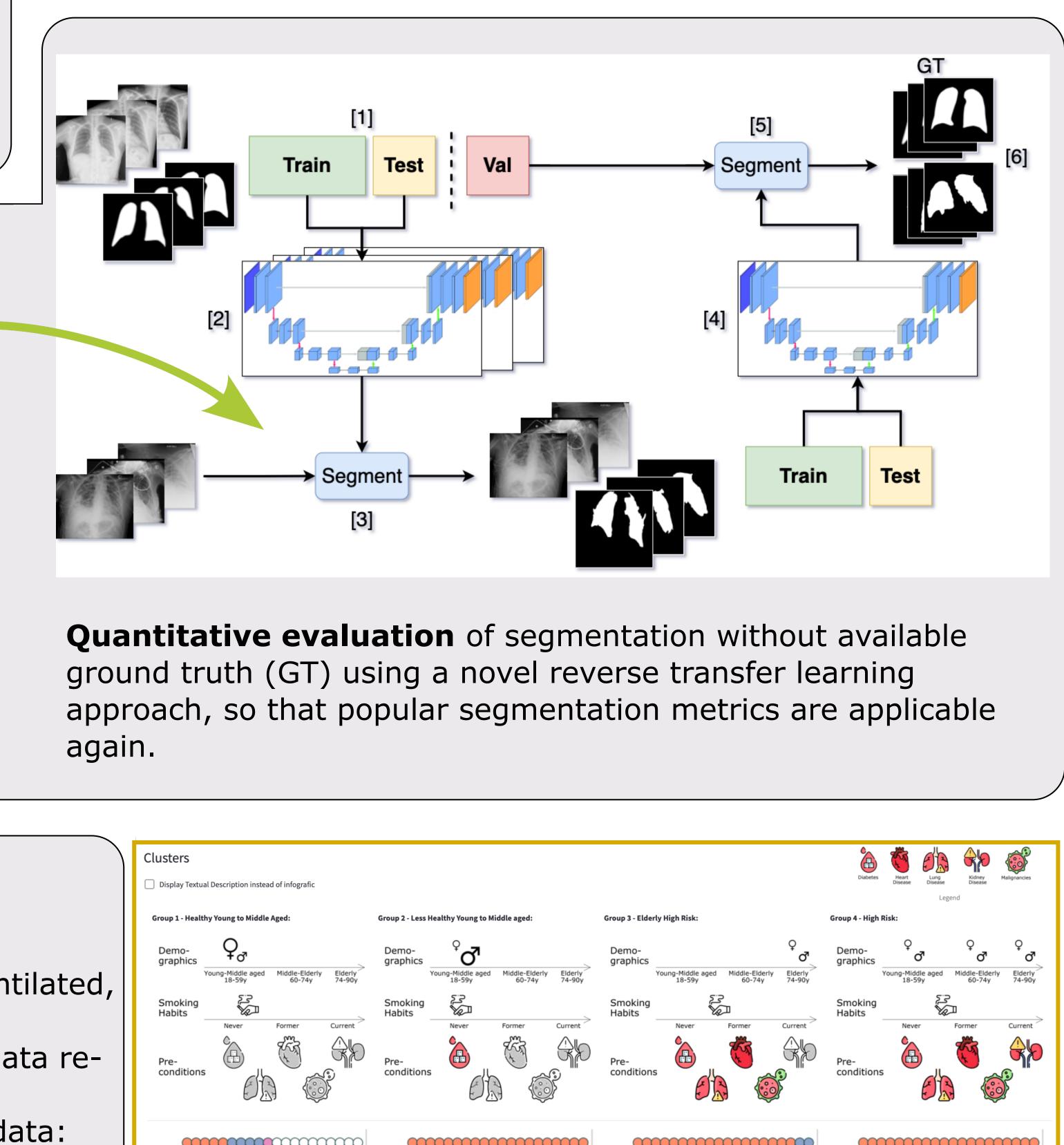


1. Automatic Lung Segmentation



several challenges that define our research questions:

- 1. Automatic Lung Segmentation using transfer learning to enable radiomics feature extraction.
- Prediction of COVID-19 hospitalized patients outcome
 Development of a visual analytics solution to support understanding into the prediction of COVID-19 outcomes for different user groups.



Qualitative evaluation of segmentation results from multiple U-Netlike deep learning models using transfer learning and different data augmentation and image pre-processing techniques.

2. Hospitalization Outcome Prediction& Data Analysis

- Multi-target prediction including combinations of targets: patient ventilated, admissioned to ICU, deceased

- Data preprocessing: imputation, scaling, dimensionality reduction, data resampling for imbalanced data

- Clustering patients into different groups based on electronic health data: Healthy Young to Middle aged, Elderly High Risk (old patients with certain less serious pre-conditions), Less Healthy Young to Middle aged (certain early pre-conditions like diabetes, increased blood pressure), High Risk (cancer patients, diabetes, heart issues, chronic kidney disease)

 Former
 Current
 Never
 Former
 Current
 Pre-conditions

 Pre-conditions
 Pre-conditions
 Pre-conditions
 Pre-conditions
 Pre-conditions



3. Visual Analytics

- **Requirement analysis** following multi-level typology of abstract visualization (Brehmer 2013)
- Implementation of interactive web-based dashboard
- **Evaluation** of design choices by conducting **case-studies** with laymen users and **usage scenarios** for remaining groups

Contributions

- Novel strategy to **evaluate segmentation** performance **without GT** - Indications that merging image and electronic health data increases outcome prediction performance slightly ($0.78 \rightarrow 0.785$ ROC)
- Visual Analytics dashboard, evaluating design choices with a small user group conducting case studies with lay users

	died.	25 died.	49 died.
to the ICU, 4 patients had to be Ventilated, and 1 died.	admissioned to the ICU, 24 patients had to be Ventilated, and 8	admissioned to the ICU, 12 patients had to be Ventilated, and	admissioned to the ICU, 46 patients had to be Ventilated, and
out of a rate in this group, 5 patients were admissioned	out of a 100 f attents in this group, 21 patients were	out of a roof adents in this group, ro patients were	out of a root attents in this group, so patients were

Outcome Prediction

Here you can input characteristis of yourself or somebody else, for instance a family member, or just play around with different inputs. From the given inputs a cluster is either *automatically* selected or can be chosen *manually*. This data is used to fill in laboratory observations in order to create a prediction. The underlying data consits of a total of **1279** patients all hospitalized during the first Covid-19 wave in 2020 in Stony Brooks, New York.

Sele	lect Age		Select gender		Patients h	neight (cm)			Patients weight (kg)	Smoking Status	Heart Failure	?	Select Group
1	L8 - 59	-	Male		- 184		-	+	90 - +	Current	- No	-	Automatic
) Diabetes		High Blood P	ressure	🗌 Kidn	ney replacem	ent ther	ару	Kidney transplant	Malignancies	Chronic Obstruct Pulmonary Disea		Predict
	tomatic group	o assignment suggests ients	Group [2] for your	inputs.					Predictions	ICU + Ventilated 6.4 Ventilated 11.6%	4%		
Sim		ients			Age 74-90	BMI	COPD	с	Hospitalized (only)		4%		
Sim	nilar Pat		Group [2] for your Age 18-59 Yes	inputs. Age 59-74 No	Age 74-90	BMI 26.9900		CN		Ventilated 11.6%	4%		
Sim	nilar Pat	ients Outcome	Age 18-59	Age 59-74	_		No		Hospitalized (only)	Ventilated 11.6%	4%		
Sir	nilar Pat Patient ID A840306	Outcome Hospitalized (only)	Age 18-59 Yes	Age 59-74	No	26.9900	No No	Ν	Hospitalized (only)	Ventilated 11.6%		italized (only) 78	8%
0 1 2	Patient ID A840306 A811941	Outcome Hospitalized (only) Hospitalized (only)	Age 18-59 Yes Yes	Age 59-74 No No	No No	26.9900 29.2900	No No No	N N	Hospitalized (only) ICU Ventilated ICU + Ventilated	Ventilated 11.6%		italized (only) 78	8%

Figure: Dashboad for laypeople users including overview of risk groups and outcome prediction enabled by interactive user input