Grupo de Supervisión y Diagnóstico de Procesos Industriales

Universidad de Oviedo

Interactive visual analytics for medical data: application to COVID-19 clinical information during the first wave

José M. Enguita[1], Diego García[1], María D. Chiara[2], Nuria Valdés[3], Ana González[1], Abel A. Cuadrado[1], Ignacio Díaz[1]

[1]University of Oviedo - Dept of Electrical Engineering
Edificio Torres Quevedo, módulo 2, Campus de Gijón 33204 - SPAIN

[2]Institute of Sanitary Research of the Principado de Asturias
Hospital Universitario Central de Asturias, Oviedo 33011 - SPAIN

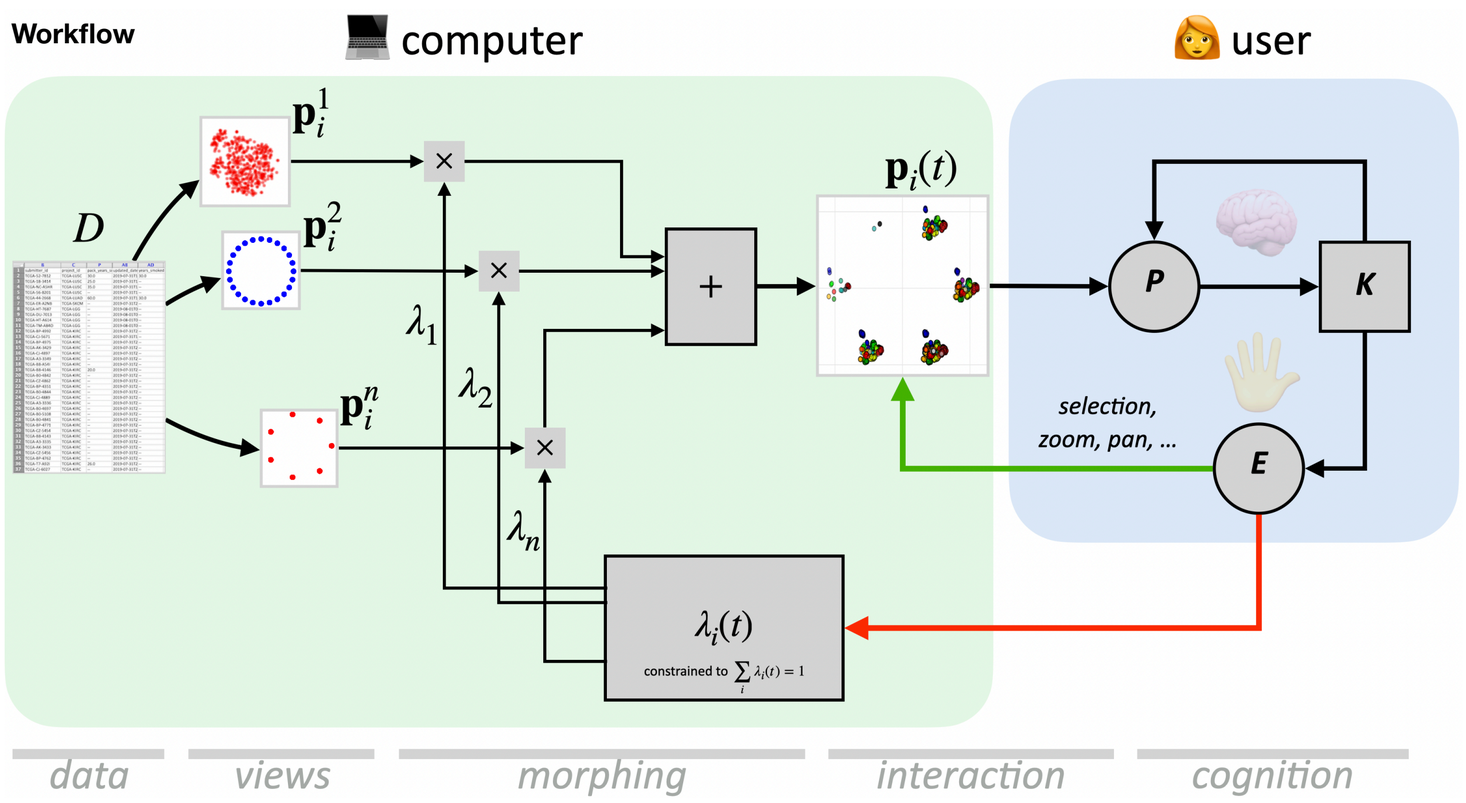
[3]Department of Internal Medicine, Section of Endocrinology and Nutrition Hospital Universitario de Cabueñes, Gijón 33204 - SPAIN

Abstract

Biomedical data recorded as a result of clinical practice are often multi-domain —involving lab measurements, medication, patient attributes, logistic information—, and also highly unstructured, with high rates of missing data and asynchronously sampled measurements. In this scenario, we need tools capable of providing a broad picture prior to more detailed analyses. We present here a visual analytics approach that uses the morphing projections technique to combine the visualization of a *t*-SNE projection of clinical time series, with views of other clinical or patient's information. The proposed approach is demonstrated on an application case study of COVID-19 clinical information taken during the first wave.

Methods

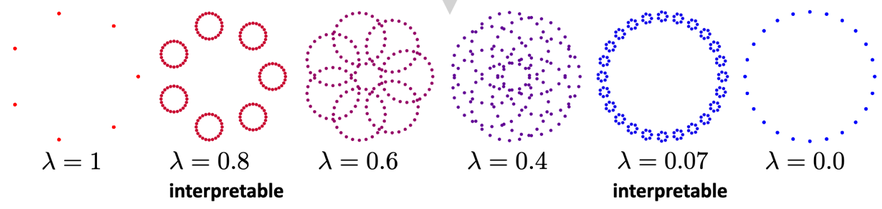
Workflow



Morphing operation $\mathbf{p}_i(t) = \sum_j \lambda_j(t)\mathbf{p}_i^j$

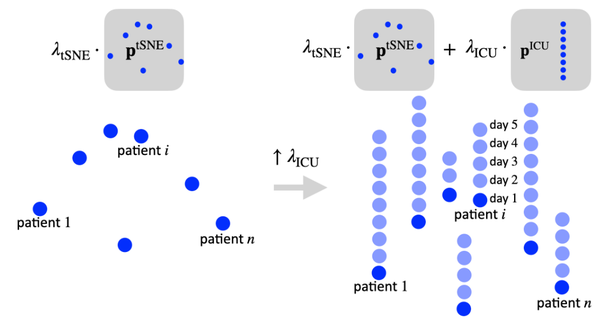
Example 1
week day and day hour

$\lambda(t) \cdot \mathbf{p}_i^{\text{weekday}} + [1-\lambda(t)] \cdot \mathbf{p}_i^{\text{day hour}}$



Example 2
tSNE + days in ICU

Patient	day	value 1	...	value *n*
⋮	⋮	⋮		⋮
patient *i*	*1*	value	...	value
patient *i*	*2*	value	...	value
⋮	⋮	⋮		⋮
patient *i*	*n*	value	...	value
⋮	⋮	⋮		⋮

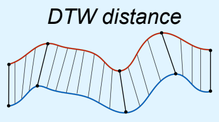


Timeseries: dynamic time warping (DTW)

Medical evolution, drug administration

d_{ij}^a DTW distance between timeseries *i* and *j* for measurement *a*

Overall distance $d_{ij}^{\text{overall}} = \sqrt{\sum_{k=1}^{5} \left(d_{ij}^k\right)^2}$

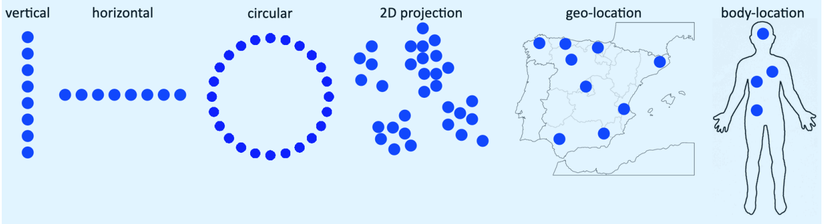


Dimensionality Reduction (*t*-SNE)

$d_{ij} \rightarrow [\text{tSNE}] \rightarrow \mathbf{p}_i^{\text{tSNE}}$

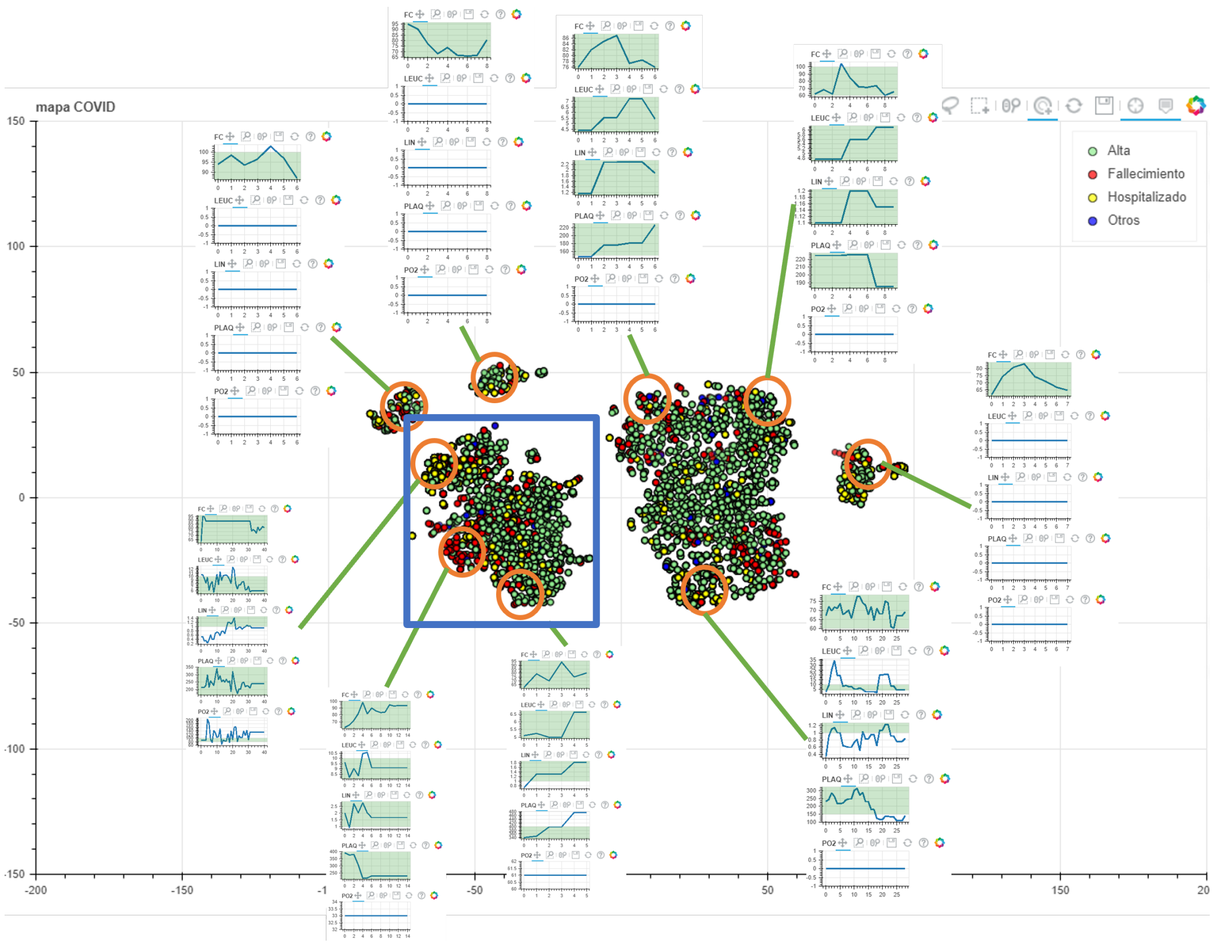
- close projections represent close (small d_{ij}) time series behavior
- clusters reveal groups of similar timeseries behavior
- results in a map revealing the *temporal* behavior of samples

Typical encodings

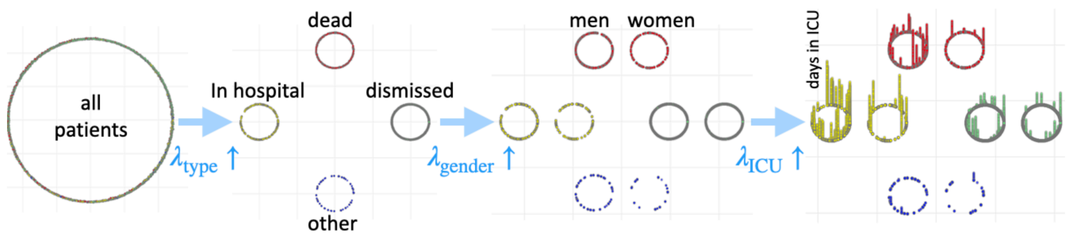


Results

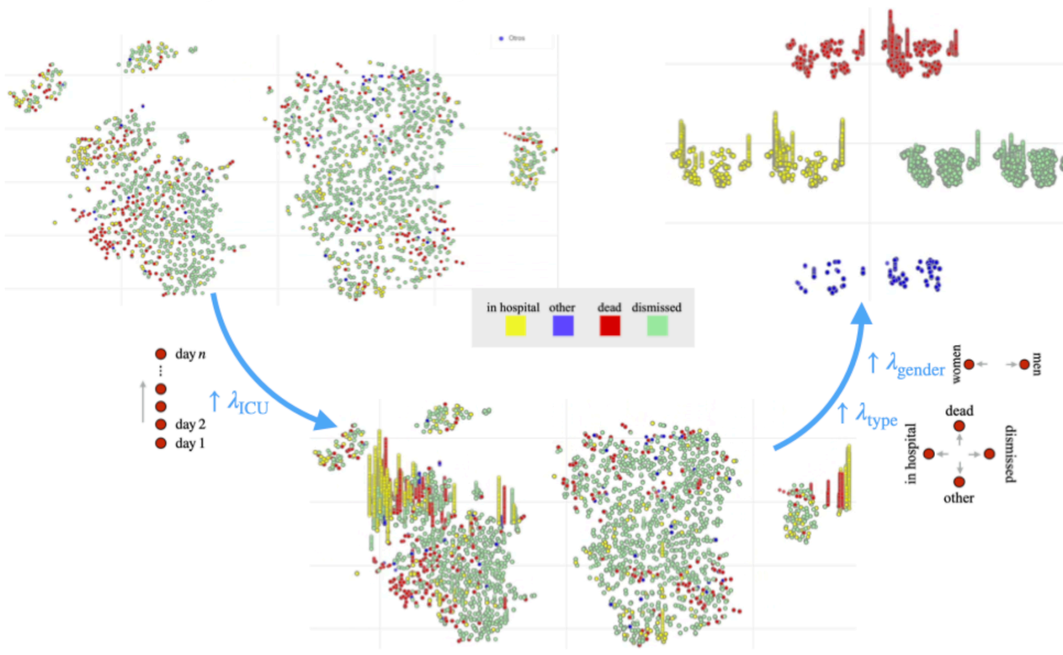
Analysis by evolution (map of timeseries)



Stratification of patients (status, sex, days in ICU)



Combined analysis: evolution + days in ICU + status



Conclusions

- Visual Analytics approach → synergy between computational intelligence and human-based reasoning
- Allows to combine knowledge from multiple domains: time evolution, clinical information (status, sex, other)
- Fluid interaction allows immediate feedback to the user → iterative question/answer exploration
- Tolerant to highly unstructured multifaceted data
- Smooth transitions foster the visual tracking of items, enhancing the discovery of connections between different domains

Acknowledgments

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
AGENCIA ESTATAL DE INVESTIGACIÓN

This work is part of Grant PID2020-115401GB-I00 funded by MCIN/AEI/ 10.13039/501100011033.

The authors also want to thank HM Hospitales for the dataset *"covid data saves lives"* used in this paper (https://www.hmhospitales.com/coronavirus/covid- data- save- lives)

Demo video