AUGMENTED INTELLIGENCE FOR SCALABLE INTERACTIVE GROUPING OF PATIENT DATA TO SUPPORT CLINICIAN RESEARCHERS

Ala Abuthawabeh[#], Abdelkader Baggag*, Michaël Aupetit*, Dena Al-Thani[@], Odette Chagoury[+], Shahrad Taheri[+]

[#]Al-Ahliyya Amman University - *Qatar Computing Research Institute, HBKU - @College of Science and Engineering, HBKU - [+]Weill Cornell Medicine Qatar
Corresponding author: maupetit@hbku.edu.qa

sponsored by Qatar National Research Fund, Member of Qatar Foundation

Objective

Design and develop an **interactive visualization system** to help **clinician researchers explore patient data**.

Visual and interactive **grouping** is essential to:

- **Discover**, compare and understand **patterns**
- **Create groups** of patients with **similar patterns**
- **Identify** relevant **statistics of each group**
- Generate **hypotheses** for next clinical study
- Generate **guidelines** for other professionals.

Challenges

How to explore data from **hundreds of patient**?

Limited actionable visualization

Limited screen space while visual interaction is crucial to engage the expert to understand and decide grouping

Limited time budget while automation is not easy as patterns and relevant features are initially unknown by the researcher

Data

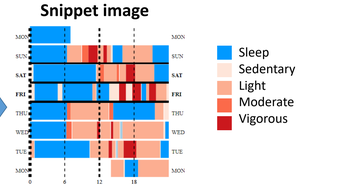
Each patient data comes as a list of **time-stamped levels of physical activity** measured through wearable sensors:

- **Visualized** as images to ease interactive grouping
- **Summarized** as statistical features for automation

Raw data

start_time	end_time	activity_level
2/9/2015 14:00	2/9/2015 17:12	light
2/9/2015 17:13	2/9/2015 18:12	light
2/9/2015 18:13	2/9/2015 20:07	light
2/9/2015 20:08	2/9/2015 20:17	sedentary
2/9/2015 20:18	2/10/2015 5:48	sleep
2/10/2015 5:49	2/10/2015 7:20	sedentary
2/10/2015 7:21	2/10/2015 8:55	light
2/10/2015 8:56	2/10/2015 10:15	moderate
2/10/2015 10:16	2/10/2015 11:04	light
2/10/2015 11:05	2/10/2015 13:34	light
2/10/2015 13:35	2/10/2015 15:45	light
2/10/2015 15:46	2/10/2015 16:41	vigorous
2/10/2015 16:42	2/10/2015 17:11	moderate
2/10/2015 17:12	2/10/2015 17:51	sedentary
2/10/2015 17:52	2/10/2015 18:28	moderate
2/10/2015 18:29	2/10/2015 19:10	vigorous
2/10/2015 19:11	2/10/2015 21:24	light
2/10/2015 21:25	2/10/2015 22:15	light
2/10/2015 22:16	2/11/2015 0:22	sedentary
2/11/2015 0:23	2/11/2015 4:27	sleep
2/11/2015 4:28	2/11/2015 5:09	moderate
2/11/2015 5:10	2/11/2015 6:25	light

Snippet image

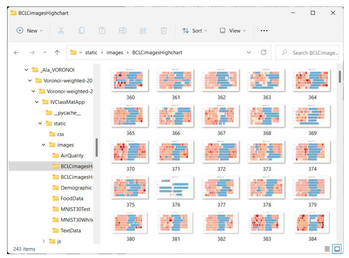


Summary statistics used by automatic **classifier**
WEEK = {full-week, week-day, week-end}
ACTIVITY = {sleep, light,..., vigorous}
- Avg ACTIVITY mid time over WEEK
- Avg ACTIVITY duration over WEEK
- Number of ACTIVITY segments during WEEK
- Etc.

Current tools

Standard image grouping tools are not adapted for visual pattern discovery at scale

- **Rigid** grid-like layout
- **Burdensome** interaction
- **No visual** grouping
- **No statistical** analysis
- **No automation**

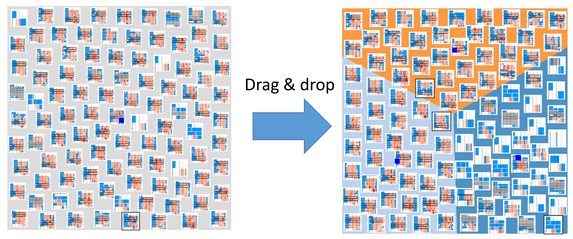


Actionable visualization

[EuroVis 2020, EuroVis 2021]

Use a **treemap visual metaphor**

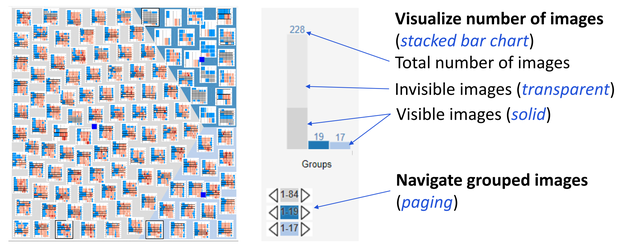
- **Subset** of the data **automatically spread** to avoid clutter and ease visual pattern discovery
- **Drag and drop** to create group interactively
- **Hierarchical arrangement** of snippet images matching with user-defined groups



Navigation to fit screen space

[EuroVis 2022]

Use of paging, navigation, and bar charts representing visible/invisible groups of patient data

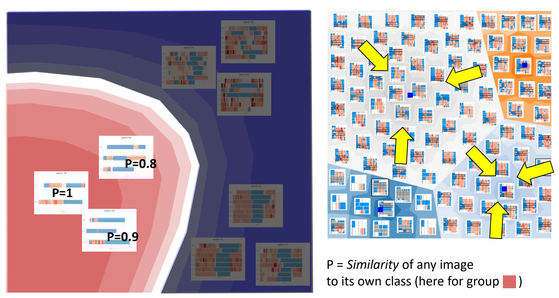


Automation to fit time budget

[EuroVis 2022]

Use of **metric learning** to compute similarity between patient data based on **statistical features** and **user-defined groups**

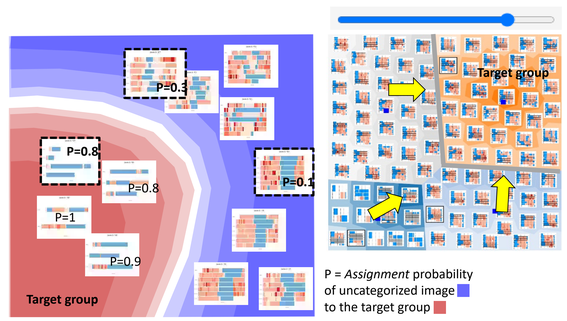
Re-arrangement and highlight of images matching with learned metric and group assignment probabilities to ease visual comparison and grouping decision



P = *Similarity* of any image to its own class (here for group)

Use of an **automatic classifier** trained on user-grouped data based on **statistical features** to compute assignment probability for yet uncategorized data (grey area)

Slider to push all data to the target group based on assignment probability threshold



P = *Assignment* probability of uncategorized image to the target group

Acknowledgment This research was made possible by the NPRP award [NPRP11C-0115-180010] from the Qatar National Research Fund. The statements made herein are solely the responsibility of the authors.

Outcomes

- **US Patent** 11/01/2022 US11222453 *Interactive visual data categorization systems and methods*
- **JMIR Human Factors 2022** K. Khowaja, W. Waheeda Syed, M. Singh, S. Taheri, O. Chagoury, D. Al-Thani, M. Aupetit. *Participatory Design Approach to develop Visualization of Wearable Actigraphy Data for Healthcare Professionals: A Case Study in Qatar*
- **EuroVis 2022** A. Abuthawabeh, A. Baggag, Michael Aupetit *Augmented Intelligence with Interactive Voronoi Treemap for Scalable Grouping: a Usage Scenario with Wearable Data*
- **Eurovis 2021** A. Abuthawabeh, M. Aupetit *Toward an Interactive Voronoi Treemap for Manual Arrangement and Grouping*
- **EuroVis 2020** A. Abuthawabeh, M. Aupetit *A Force-Directed Power Diagram Approach for Interactive Voronoi Treemaps*

Future work

Final evaluation with clinician researchers planned in Summer 2023.

The tool is generic. Contact ***maupetit@hbku.edu.qa*** to explore/analyze your clinical data.

Abstract

We propose to design and develop an interactive visualization to support the clinician researcher to explore patient data. Visual and interactive grouping is essential to discover, compare and understand patterns in patient data, to create groups of data with similar patterns, and to identify relevant statistics of each group.

Such groups are then used by the clinician researcher to generate hypotheses to design the next clinical study or to generate health guidelines for downstream healthcare professionals like physicians and educators. Here we focus on the analysis of wearable data to help clinician discover groups of patients with specific patterns of physical activity and sleep.

Our visualization is based on an interactive Voronoi treemap which supports arrangement and grouping of data with snippet image representations. We show how we address the scalability issue arising when a clinician attempts to group hundreds of patient data. We propose visualizations designed to manage images visibility, evaluate group homogeneity, and shorten grouping task completion time using metric learning and classification while keeping final decision under control. It is supported by an automatic classifier forming an augmented intelligence system that tackles arrangement and grouping tasks at scale.