t-viSNE: Interactive Assessment and Interpretation of t-SNE Projections

Angelos Chatzimparmpas

Assistant Professor

April 29, 2025

Visualization and Graphics Group

Utrecht University

ELLIIT
Excellence Center at Linköping – Lund in Information Technology

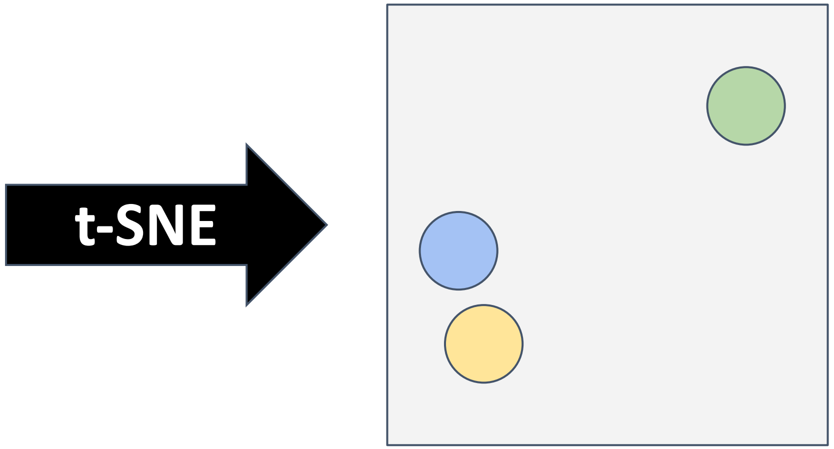
What is t-SNE?

t-SNE is a method for visualizing **high-dimensional** data, typically projected in 2D or 3D (scatterplot with meaningless axes).

4 attributes / dimensions

3 data points

	Age	Weight	Height	Eyes
John	32	78kg	1,67m	Brown
Lars	18	95kg	2,03m	Blue
Sarah	31	77kg	1,68m	Brown

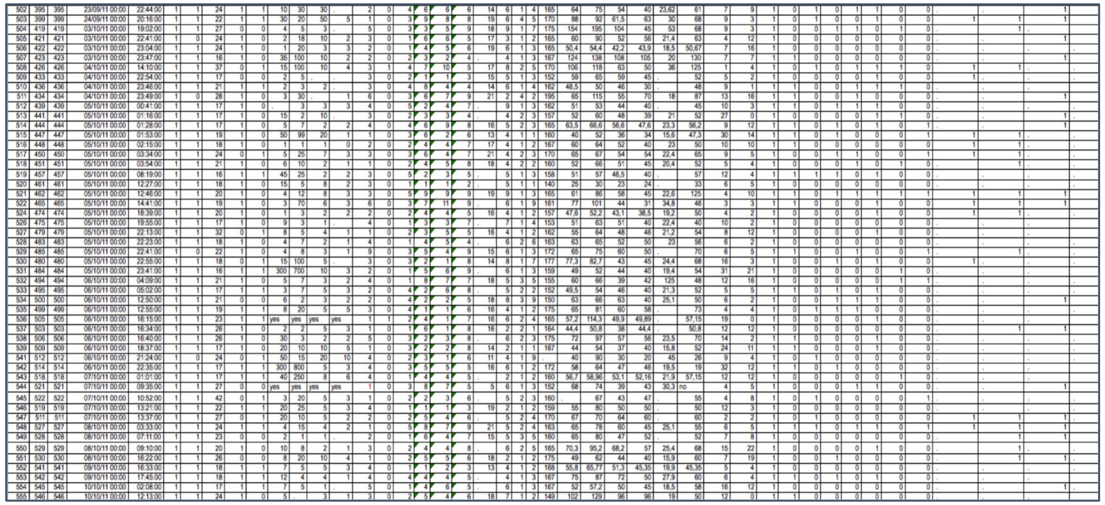


John and Sarah are quite similar;
Lars is different.

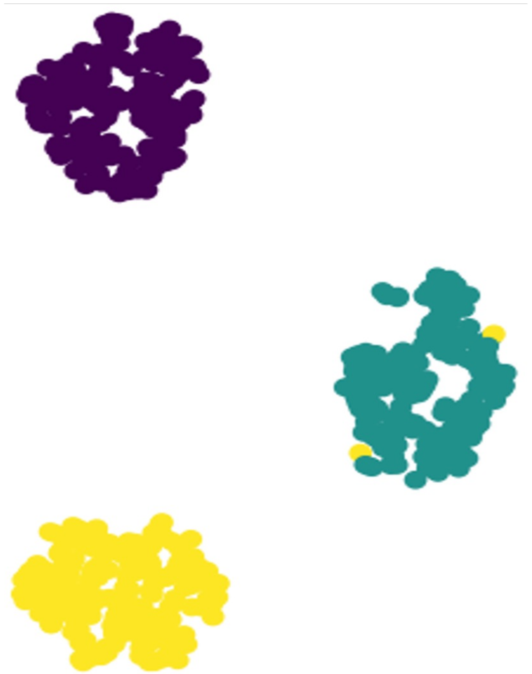
Why is t-SNE useful?

Hundreds of attributes / dimensions

Thousands of data points



t-SNE



Similar???

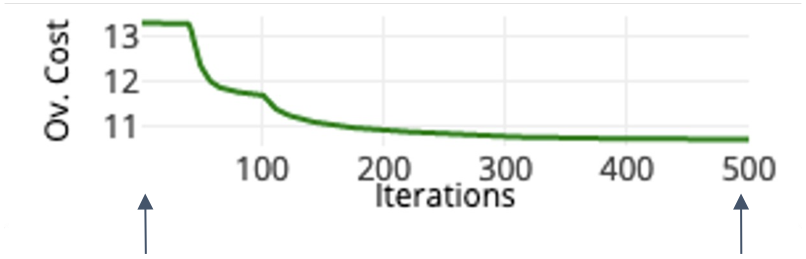
Different???

t-SNE algorithm in a nutshell

- Step 1: Calculate all *pairwise distances* between data points in the original, N-dimensional space.
- Step 2: Extract *neighborhoods* based on the distances from Step 1, i.e., we only care about the *nearest* points.
- Step 3: Try to match in the visualization, in 2D, *as much as possible*, the original neighborhoods, improving one point at a time (many times per second).

Each iteration of the optimization reduces the overall ***cost*** of the visualization, i.e., how much it differs from the original data.

Illustration of Step 3



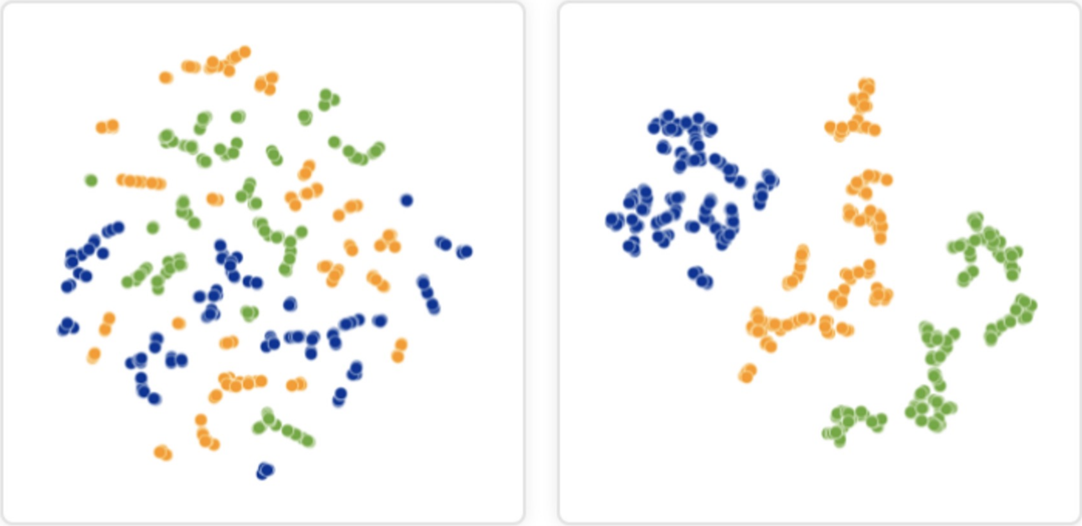
<u>Start (random):</u>
High cost

<u>End (optimized):</u>
Low cost

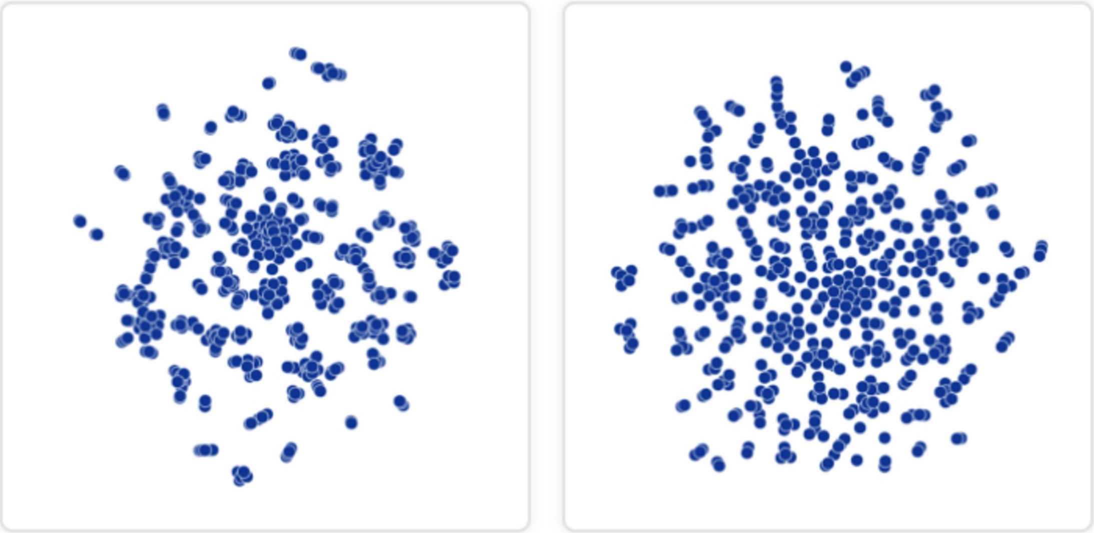
Animated GIF from: https://www.oreilly.com/learning/an-illustrated-introduction-to-the-t-sne-algorithm

Main problems with “pure” t-SNE plots

It is not always that easy, however…



Three tight clusters
(?)



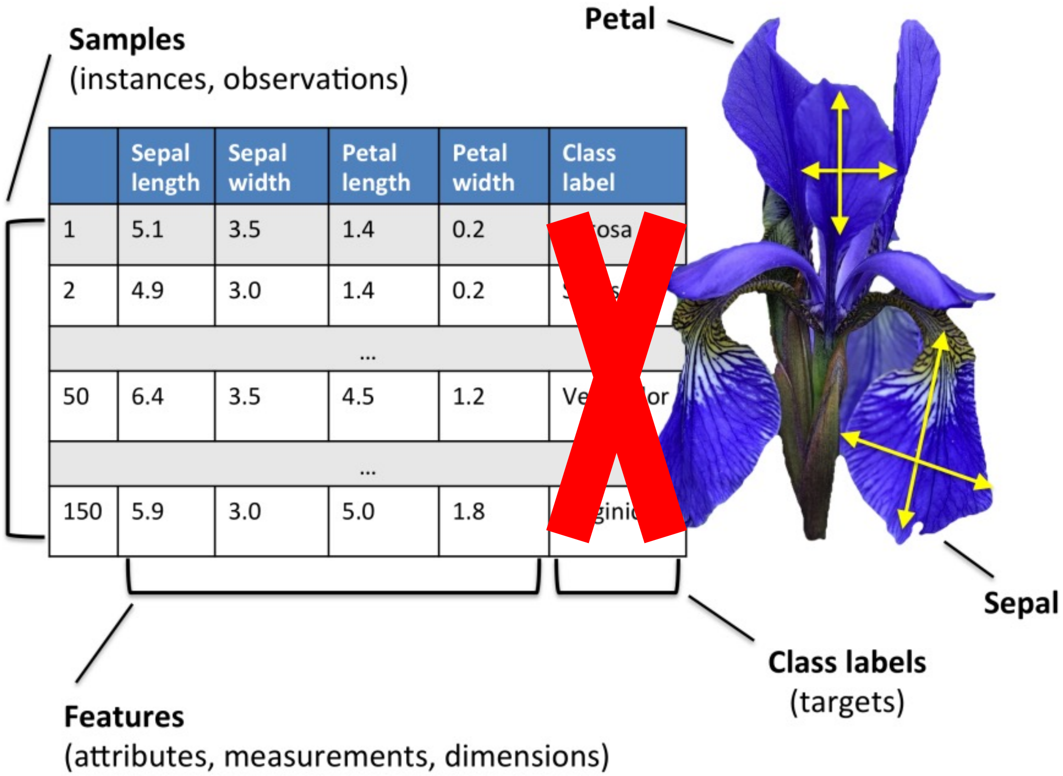
100% random data
(!)

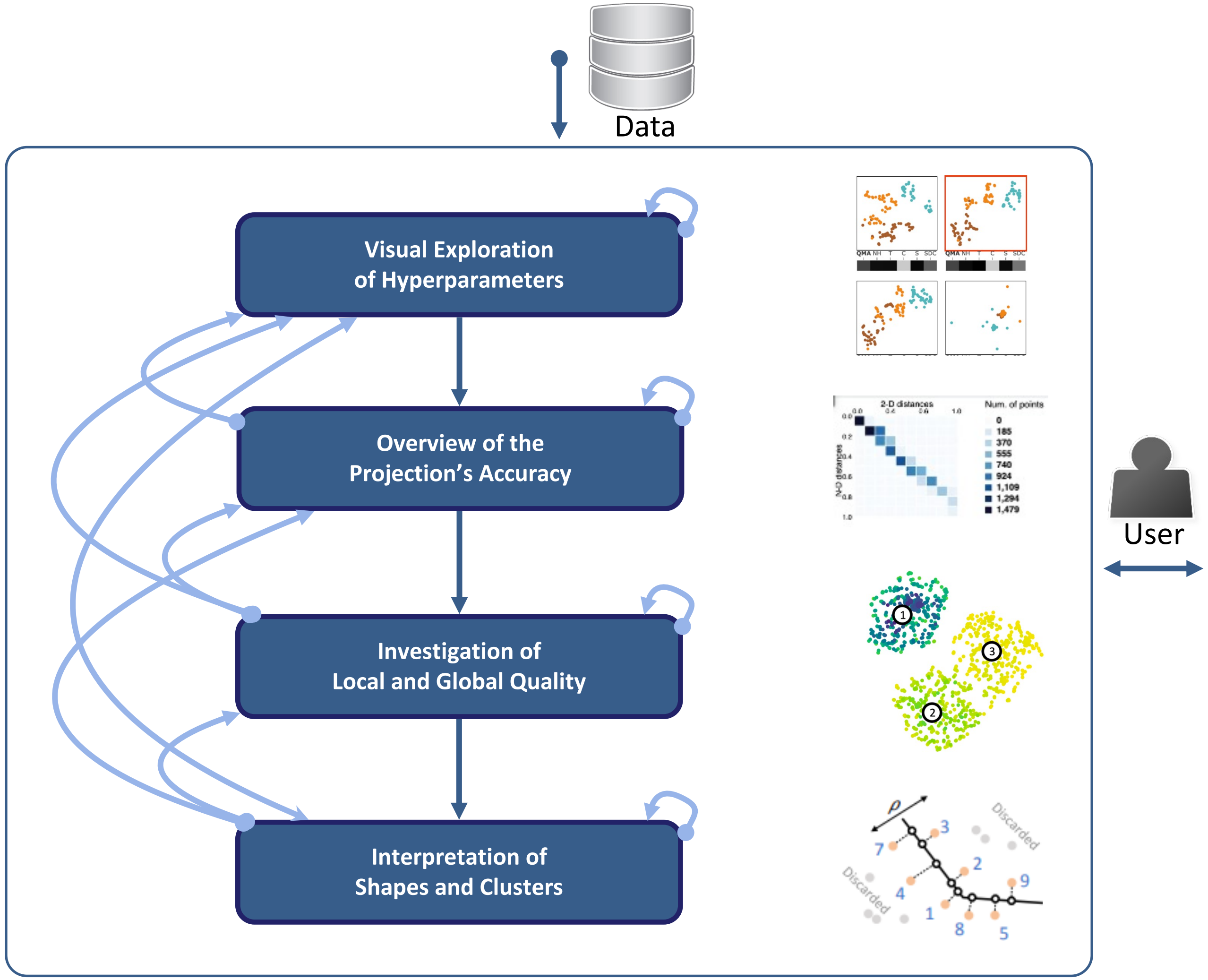
Previous work

Tool	Features/Capabilities							Available
	Overview			Quality		Dimensions		
	Multiple DR Algor. Support	t-SNE	Visual Param. Expl.	Global Quality Assess.	Local Quality Assess.	Rank Dim.	Interact. Shape Analysis	
t-viSNE		✓	✓	✓	✓	✓	✓	T+SC
Clustervision [51]	✓	✓		(✓)	(✓)	✓		T
VisCoDeR [22]	✓	✓	✓		✓			T
Clustrophile 2 [23]	✓	✓	(✓)	(✓)				
AxiSketcher [47]		✓				✓	✓	
GEP [63]	✓	✓			(✓)			T+SC
ccPCA [44]		✓				✓		T+SC
DimReader [45]		✓	(✓)					SC
Coimbra *et al.* [42]	✓			✓	✓	✓		T+SC
Praxis [46]	✓				(✓)	✓		
FocusChanger [50]					✓	✓		
Probing Proj. [36]				✓	✓			T+SC
ProxiLens [34]					✓			
Stress Maps [29]					✓			

The last column indicates if the tool (T) and/or its source code (SC) are available online (last checked: January 15, 2020).

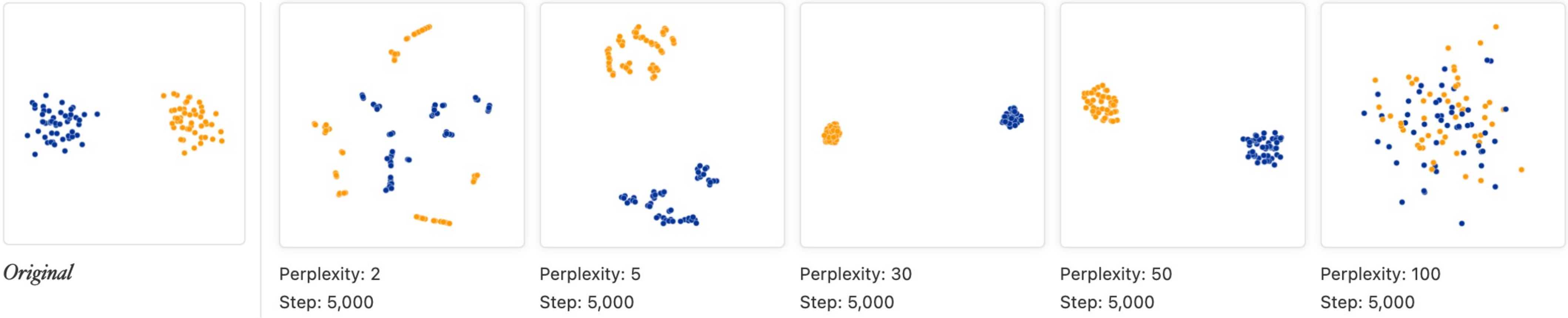
Deep Dive into t-viSNE's Workflow





Fisher R. A. The use of multiple measurements in taxonomic problems. Annals of Eugenics. 1936;7(2):179-188.

The t-SNE's **hyperparameters** really matter



Wattenberg, et al., "How to Use t-SNE Effectively", Distill, 2016. http://doi.org/10.23915/distill.00002

t-viSNE: Hyperparameters and Quality Metrics

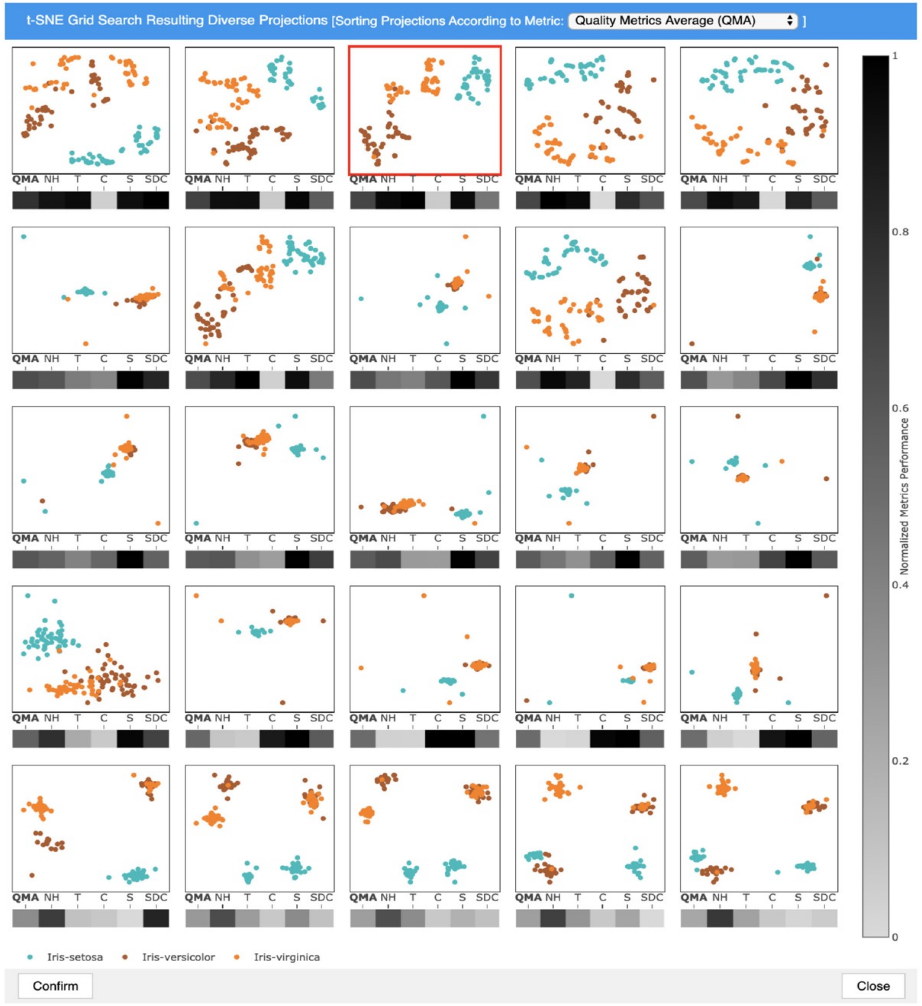
25 representative projections from a pool of 500

Hyperparameters visual exploration:

- perplexity
- learning rate
- max iterations

Quality metrics:

- neighborhood hit (NH)
- trustworthiness (T)
- continuity (C)
- normalized stress (S)
- Shepard diagram correlation (SDC)

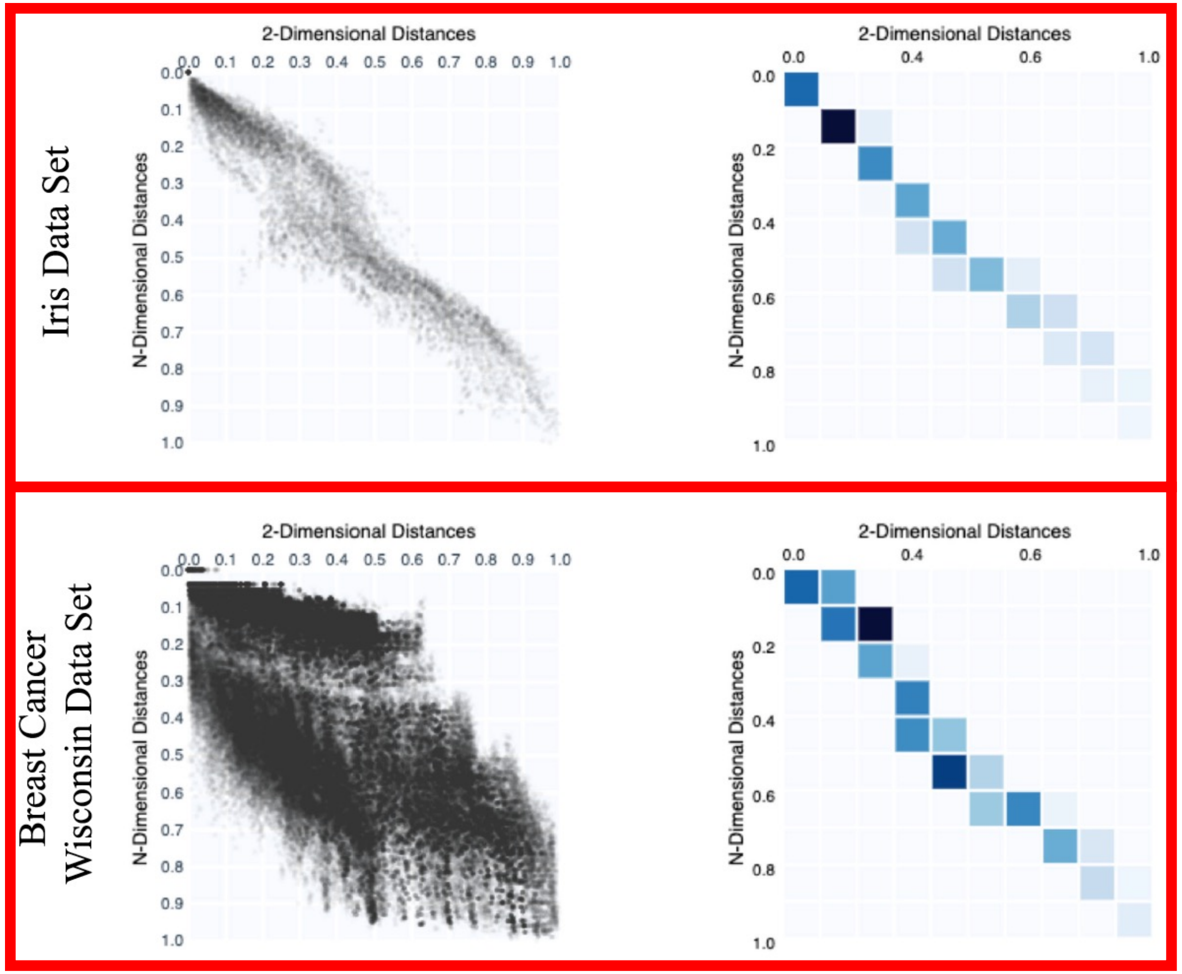


t-viSNE: Inspection of Global Projection Quality

Overall accuracy of a projection by comparing 2-D vs. N-D distances

Shepard Heatmap (SH) vs. Shepard Diagram (SD):

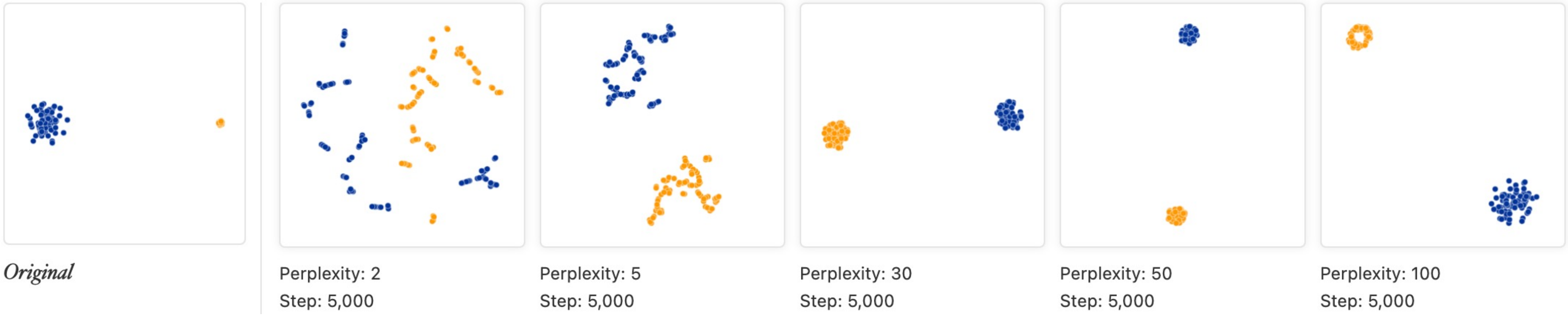
- SH = SD for smaller data sets
- SH > SD for larger data sets



Shepard Diagram

Shepard Heatmap

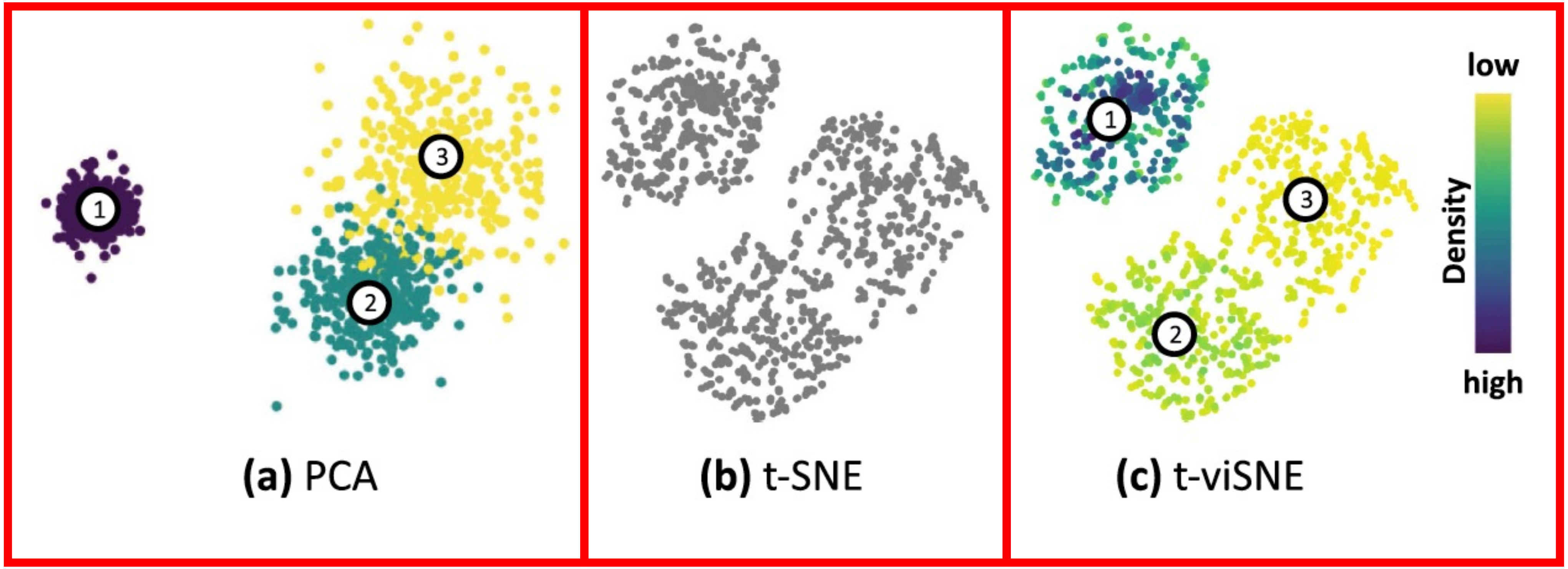
Cluster sizes in a t-SNE plot mean nothing



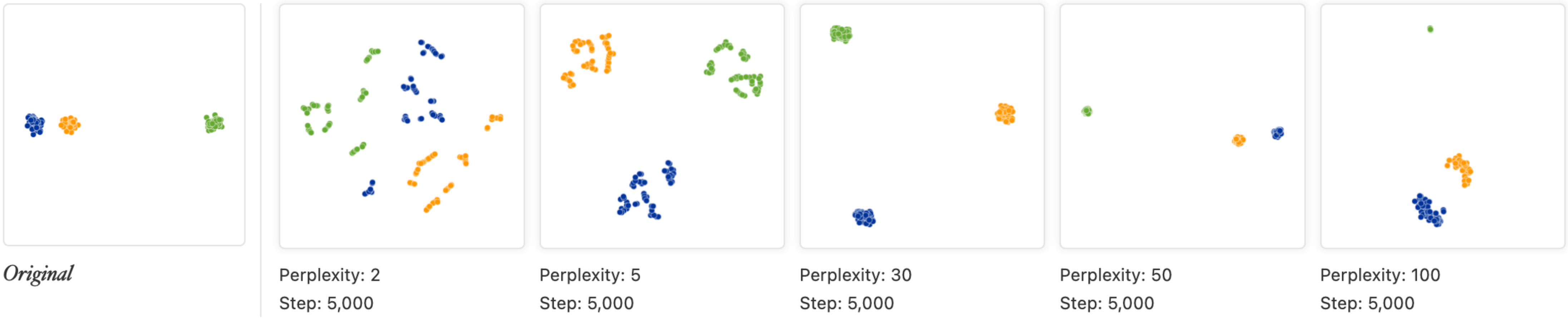
Wattenberg, et al., "How to Use t-SNE Effectively", Distill, 2016. http://doi.org/10.23915/distill.00002

t-viSNE: Inspection of Local Projection Quality

Extracting the **density** variable from the internal processes of t-SNE algorithm



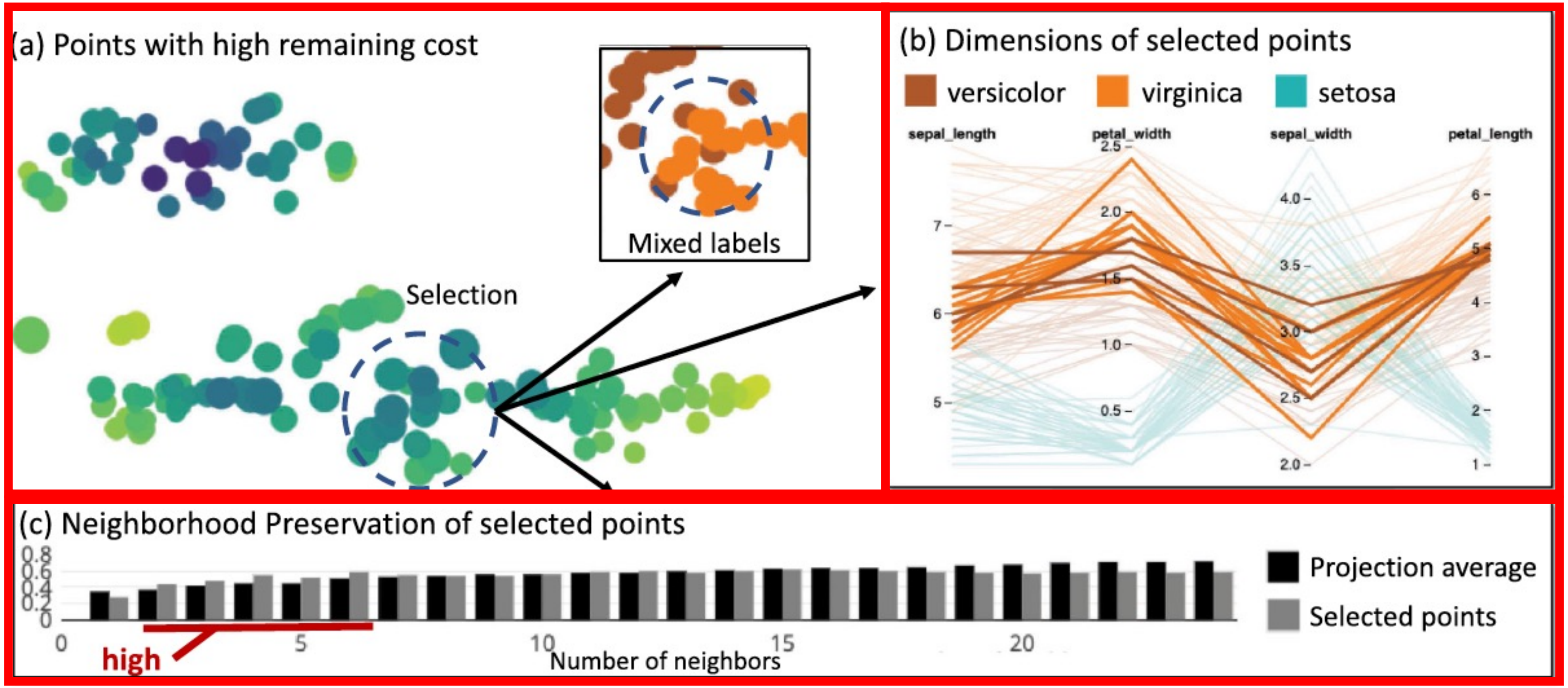
Distances between clusters may mean nothing



Wattenberg, et al., "How to Use t-SNE Effectively", Distill, 2016. http://doi.org/10.23915/distill.00002

t-viSNE: Inspection of Local Projection Quality

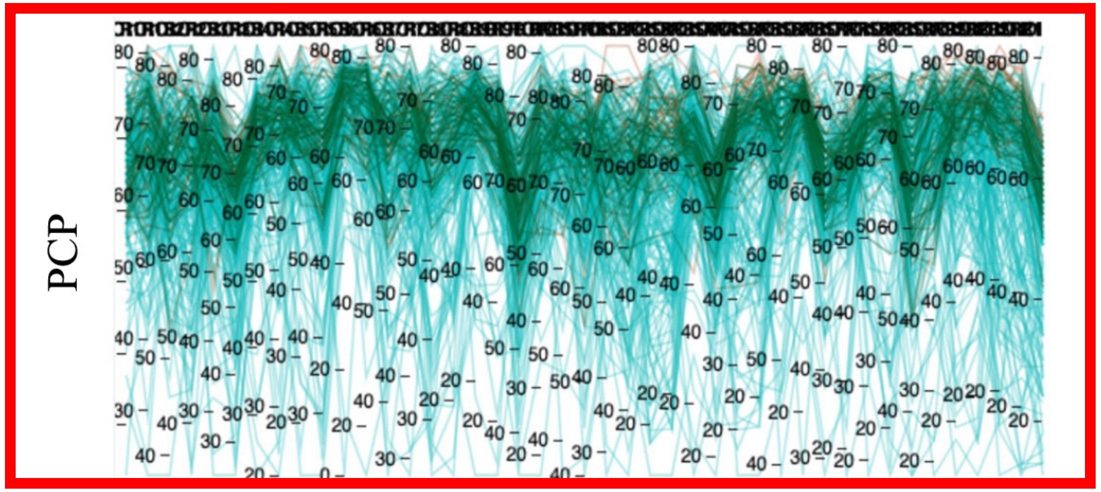
Investigation of local quality for specific data points

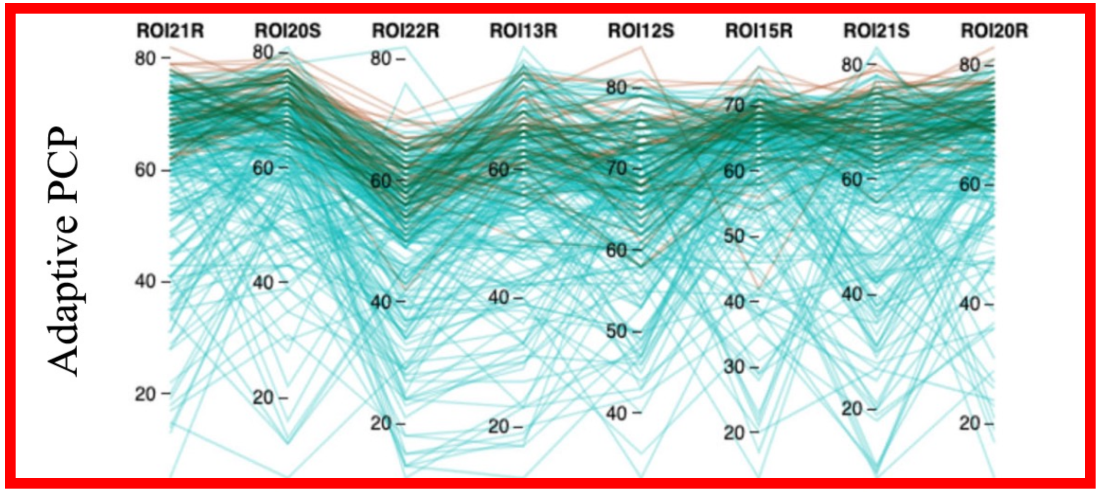


t-viSNE: Data Inspection with PCP

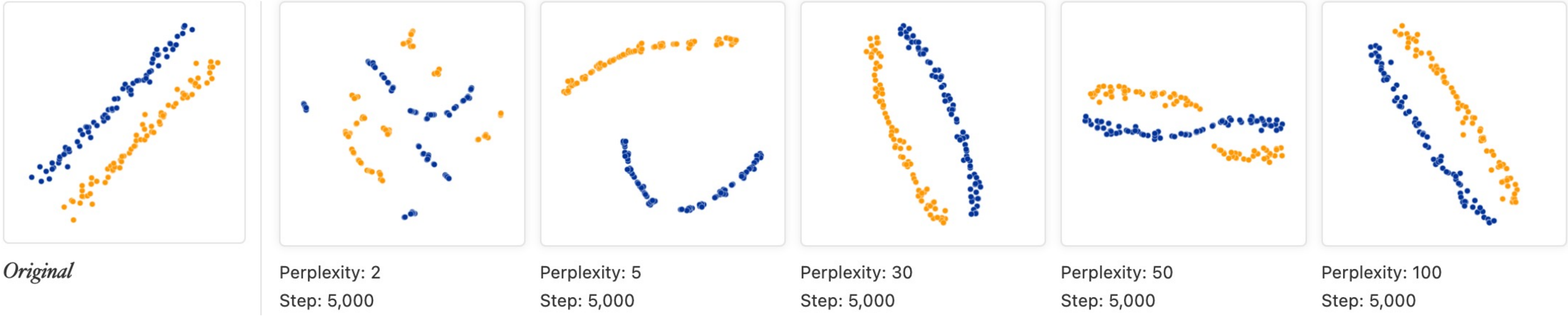
How does the **Adaptive** parallel coordinates plot work (see steps below)?

- Step 1: Use local Principal Component Analysis (PCA) algorithm for the selection of points
- Step 2: Compute the most important dimensions with the largest variance
- Step 3: Reorder the axes (i.e., dimensions) according to Step 2
- Step 4: Limit the dimensions to the 8 most informative





You can see some **shapes**, sometimes

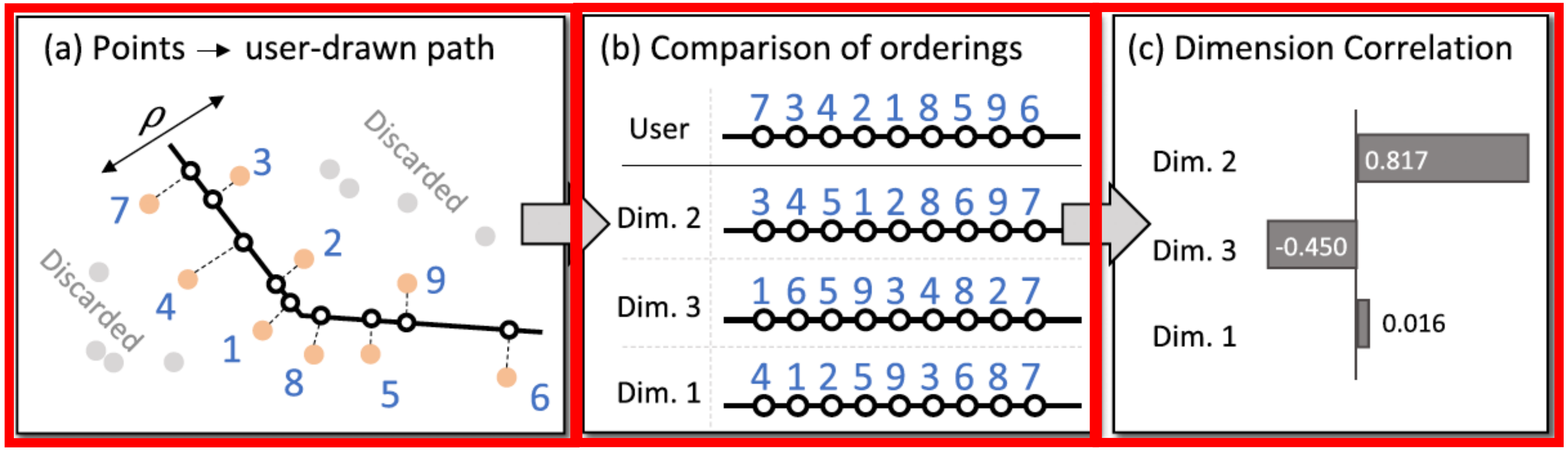


Wattenberg, et al., "How to Use t-SNE Effectively", Distill, 2016. http://doi.org/10.23915/distill.00002

t-viSNE: Investigation of User-Defined Patterns

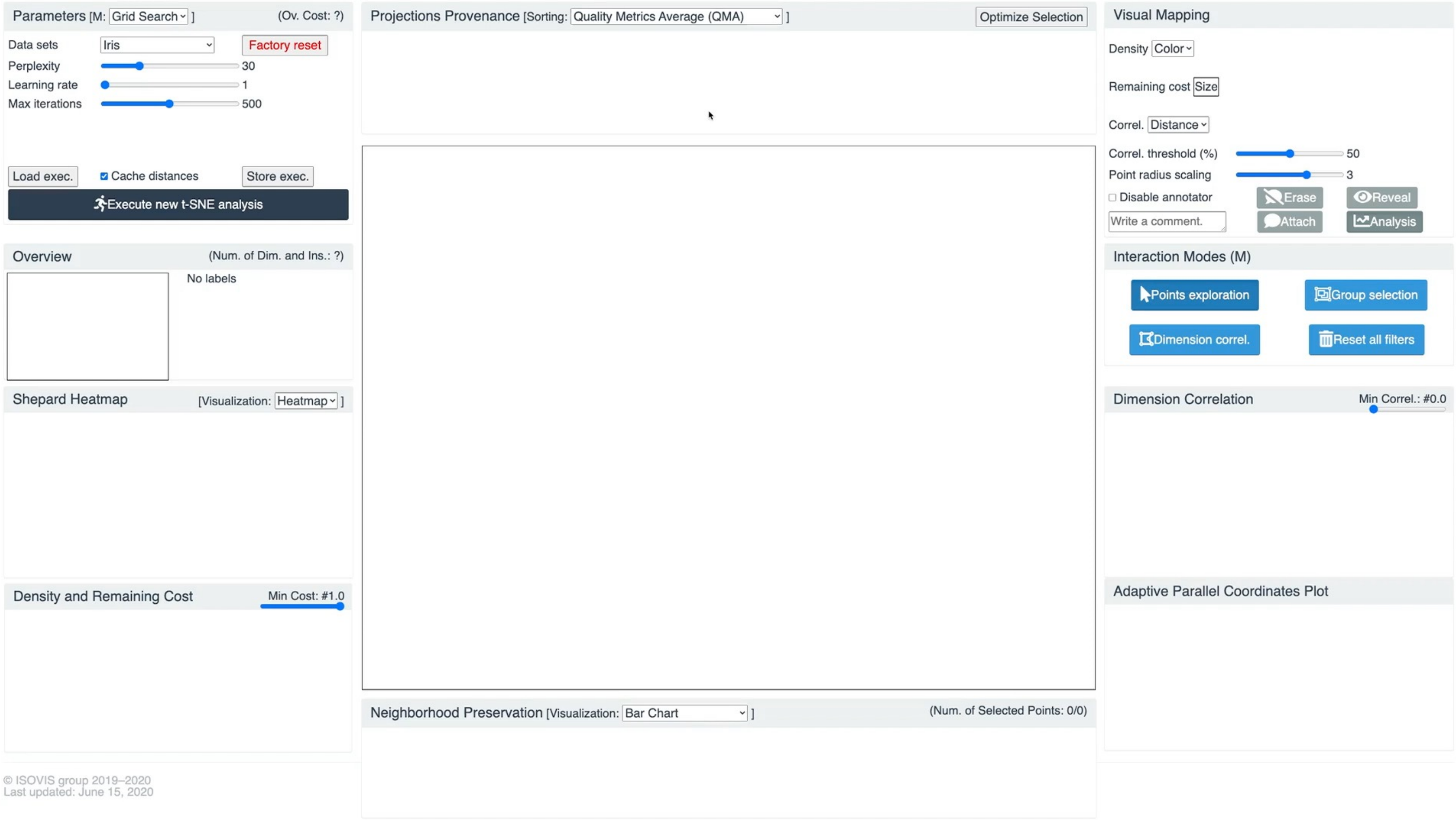
How does the **Dimension Correlation** tool work (see steps below)?

- Step 1: User draws a path
- Step 2: Automatically comparing user's path vs. all dimensions' ordering
- Step 3: Calculate the correlations between the user-drawn path and all dimensions



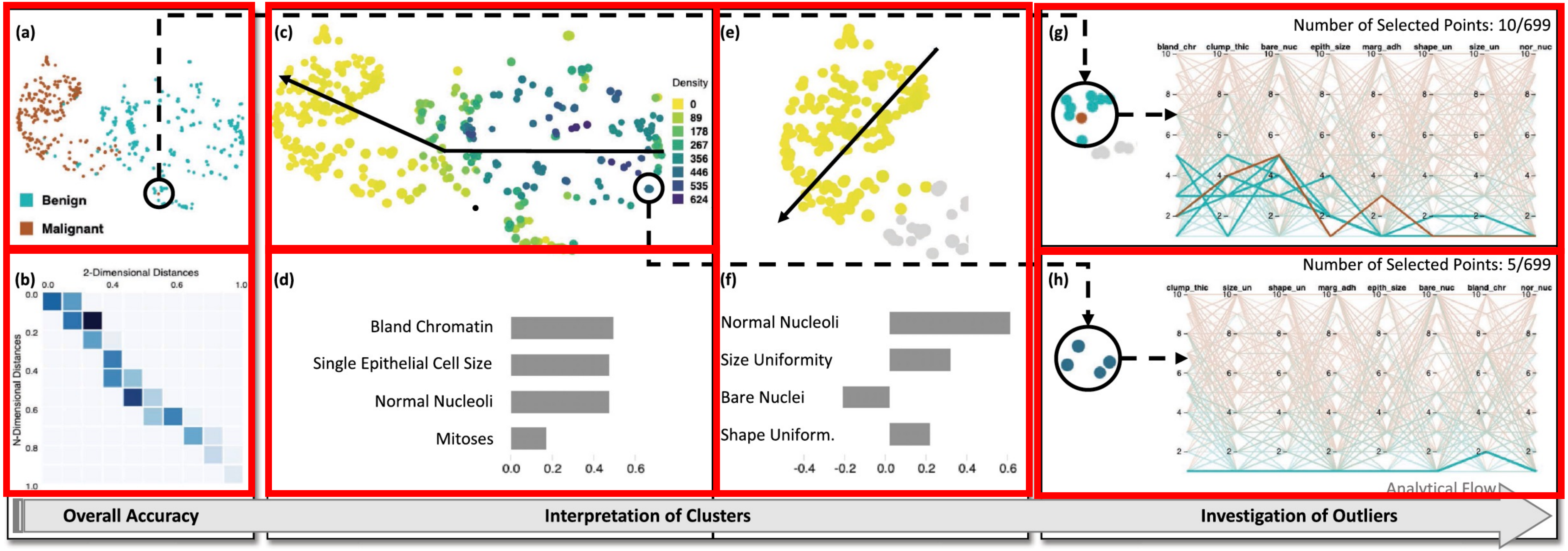
Demo Video

https://sheldon.lnu.se/t-viSNE/



Use Case: Understanding a Cancer Classifier

- Dataset: Breast Cancer Wisconsin (Benign or Malignant Cancer); 699 data points (cancer cases); 9 dimensions (cytological characteristics)



Evaluation

- ICE-T* qualitative users' feedback
- Further task-specific qualitative results present in the paper

Components	t-viSNE				
	Insight	Time	Essence	Confidence	Average
Participant 7	6.63	6.60	7.00	7.00	6.81
Participant 9	6.88	6.80	7.00	6.50	6.79
Participant 6	6.88	7.00	6.75	6.50	6.78
Participant 5	6.63	6.00	6.00	6.67	6.32
Participant 12	6.25	6.20	6.50	6.25	6.30
Participant 8	6.63	5.80	6.75	6.00	6.29
Participant 13	6.25	6.40	6.25	6.25	6.29
Participant 11	6.63	6.00	6.50	6.00	6.28
Participant 14	6.88	6.60	5.75	5.75	6.24
Participant 4	5.71	6.00	6.50	5.75	5.99
Participant 10	6.13	5.20	6.50	5.50	5.83
Participant 3	6.00	5.80	6.00	5.50	5.83
Participant 1	6.13	6.20	4.75	5.50	5.64
Participant 2	5.63	5.40	5.50	5.25	5.44
95% C.I.	**6.37 ± 0.24**	**6.14 ± 0.29**	**6.27 ± 0.36**	**6.03 ± 0.30**	**6.20 ± 0.24**

Components	GEP				
	Insight	Time	Essence	Confidence	Average
Participant 24	6.00	5.80	5.75	6.33	5.97
Participant 17	6.00	6.00	5.67	5.33	5.75
Participant 21	5.83	6.40	6.25	3.75	5.56
Participant 15	5.00	5.40	6.00	5.33	5.43
Participant 26	6.13	5.60	5.50	4.25	5.37
Participant 25	5.50	5.40	5.75	4.75	5.35
Participant 23	6.13	5.40	4.50	4.75	5.19
Participant 22	5.50	5.40	3.25	4.75	4.73
Participant 18	4.75	5.80	4.25	3.75	4.64
Participant 19	4.75	5.20	4.75	3.67	4.59
Participant 20	4.88	4.80	4.00	4.25	4.48
Participant 16	4.50	4.60	3.75	3.67	4.13
Participant 27	5.00	4.20	3.75	2.00	3.74
Participant 28	3.88	5.00	3.25	2.25	3.59
95% C.I.	**5.27 ± 0.40**	**5.36 ± 0.33**	**4.74 ± 0.61**	**4.20 ± 0.67**	**4.89 ± 0.43**

Legend: 7 6 5 4 3 2 1

* E. Wall et al., **"A heuristic approach to value-driven evaluation of visualizations,"** IEEE Trans. Vis. Comput. Graphics, vol. 25, no. 1, pp. 491–500, Jan. 2019.

Limitations and Future Work of t-viSNE

Limitations with regard to:

- Computational Efficiency (e.g., t-SNE-CUDA?)
- Other DR Methods (e.g., UMAP?)
- Observational User Study (e.g., Log Analysis?)
- **Progressive Quality Analysis (e.g., HyperNP*?)**

The above limitations $\Longrightarrow$ future work

* E. Appleby et al., **"HyperNP: Interactive Visual Exploration of Multidimensional Projection Hyperparameters,"** Computer Graphics Forum, vol. 41, no. 3, pp. 169-181, June 2022.

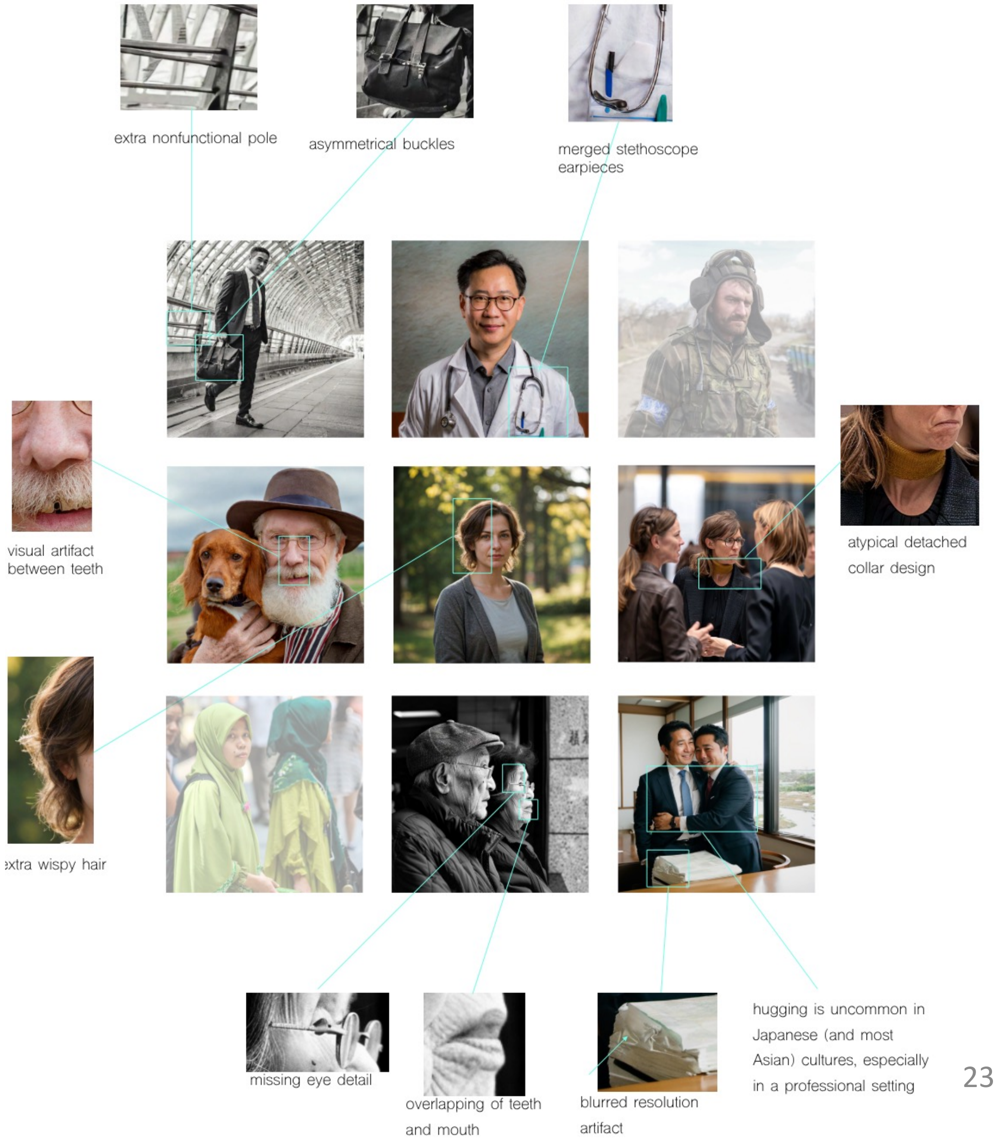
Okay, Okay, … We Get it!

You Develop Visual Analytics Tools,
Now What?

Detect DeepFakes Project

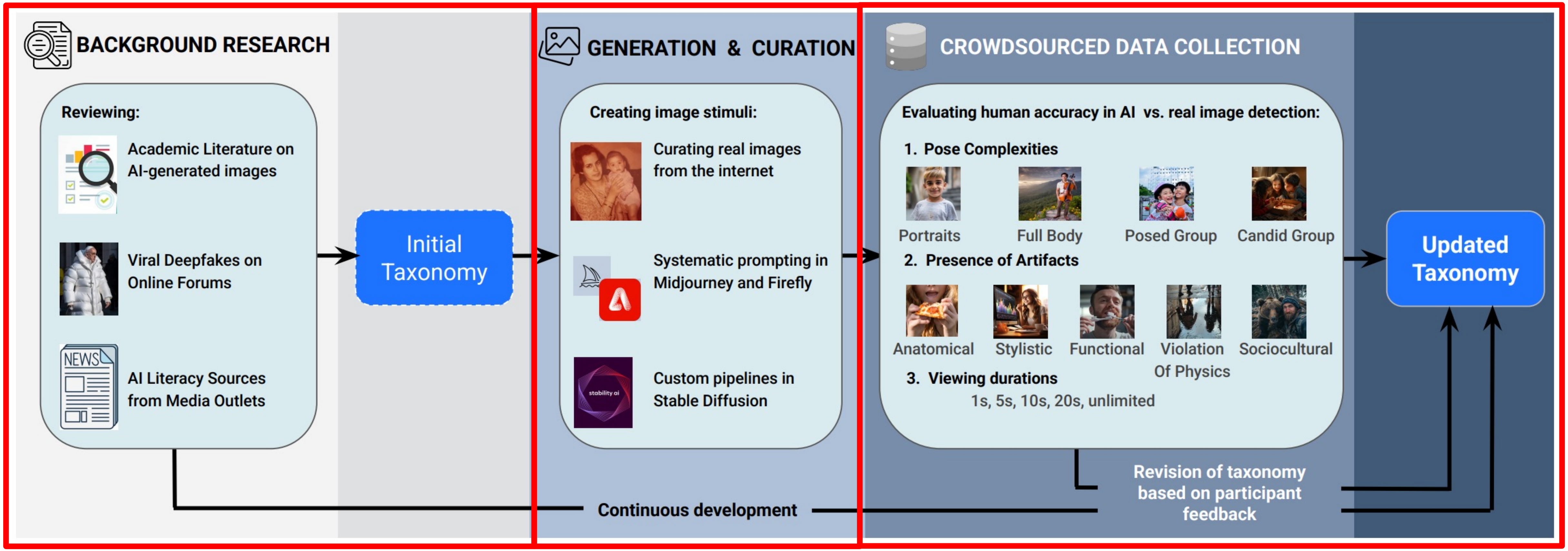
Two images are real. Can you spot them?

N. Kamali, K. Nakamura, **A. Chatzimparmpas**, J. Hullman, M. Groh, How to Distinguish AI-Generated Images from Authentic Photographs, arXiv, 2024



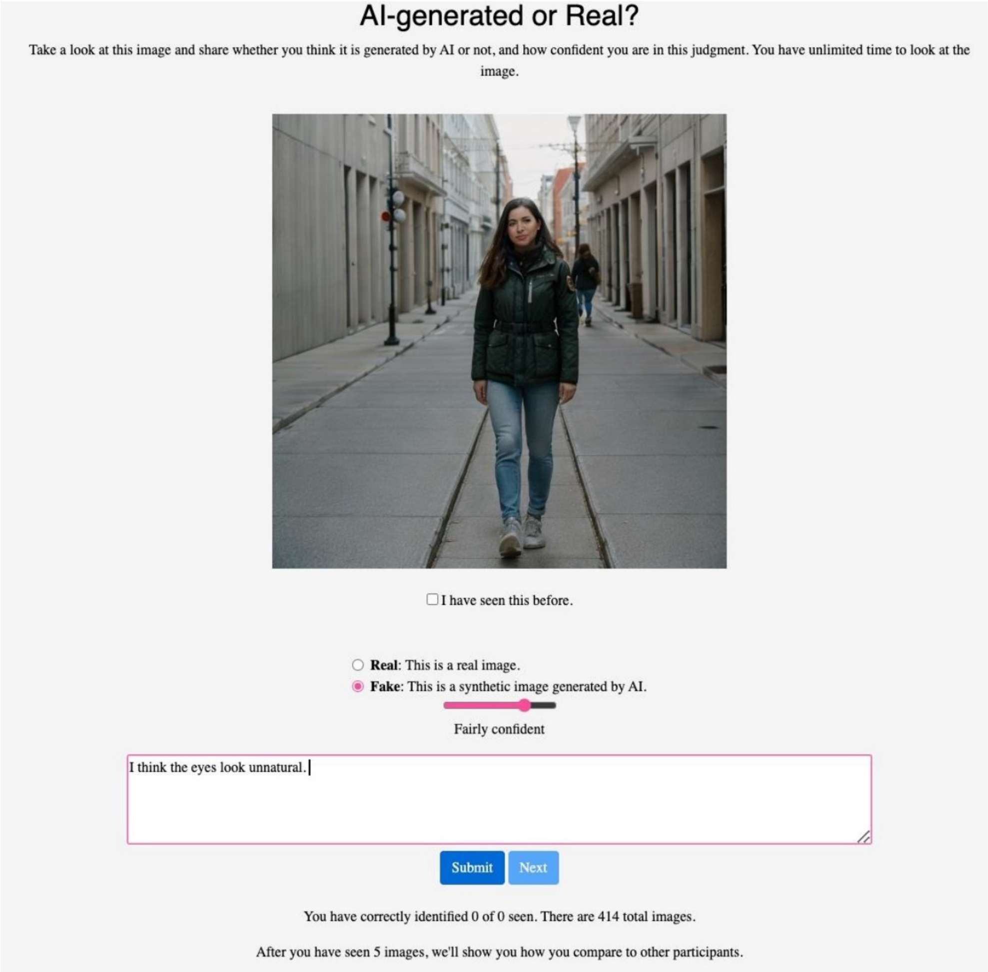
Taxonomy Development Process

Current Research with Northwestern University, USA



N. Kamali, K. Nakamura, A. Kumar, **A. Chatzimparmpas**, J. Hullman, M. Groh, Characterizing Photorealism and Artifacts in Diffusion Model-Generated Images, ACM CHI 2025

A Screenshot of the Experiment Website Interface



https://detectfakes.kellogg.northwestern.edu/

- 450 diffusion-model generated images
- 149 real images

- 749,828 observations
- 35,000 comments

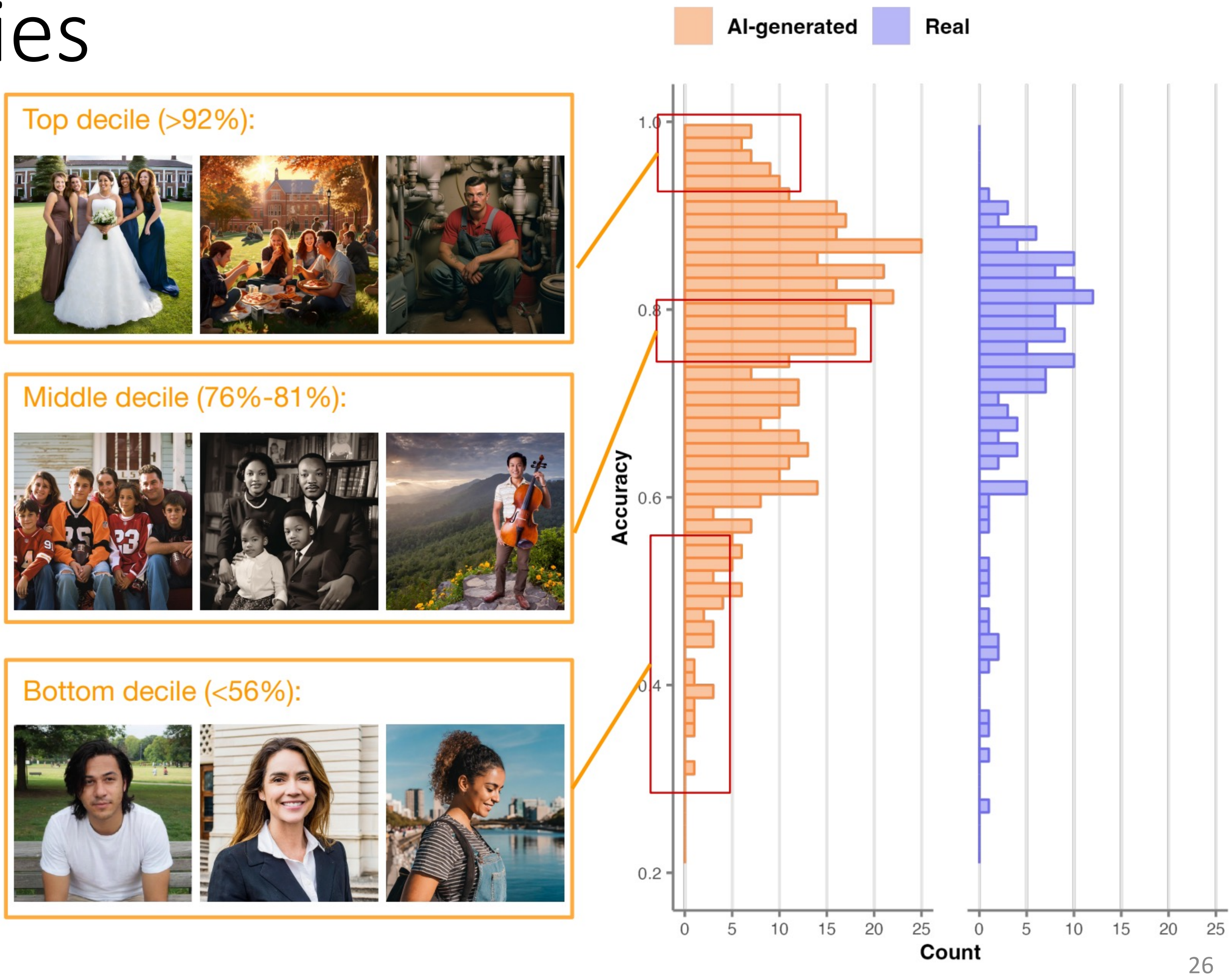
- 50,000+ unique visitors/participants

Pose Complexities

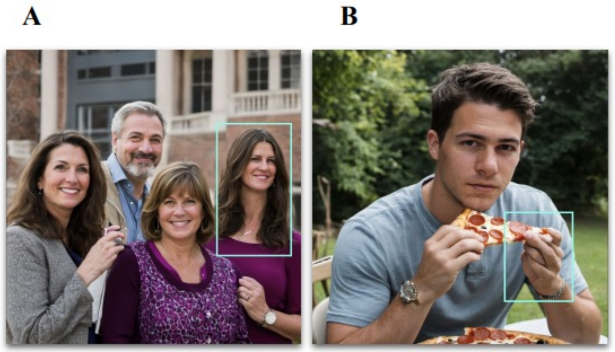
- Posed Group

- Candid & Full Body Groups

- Portraits



Presence of Artifacts



1: Anatomical Implausibilities

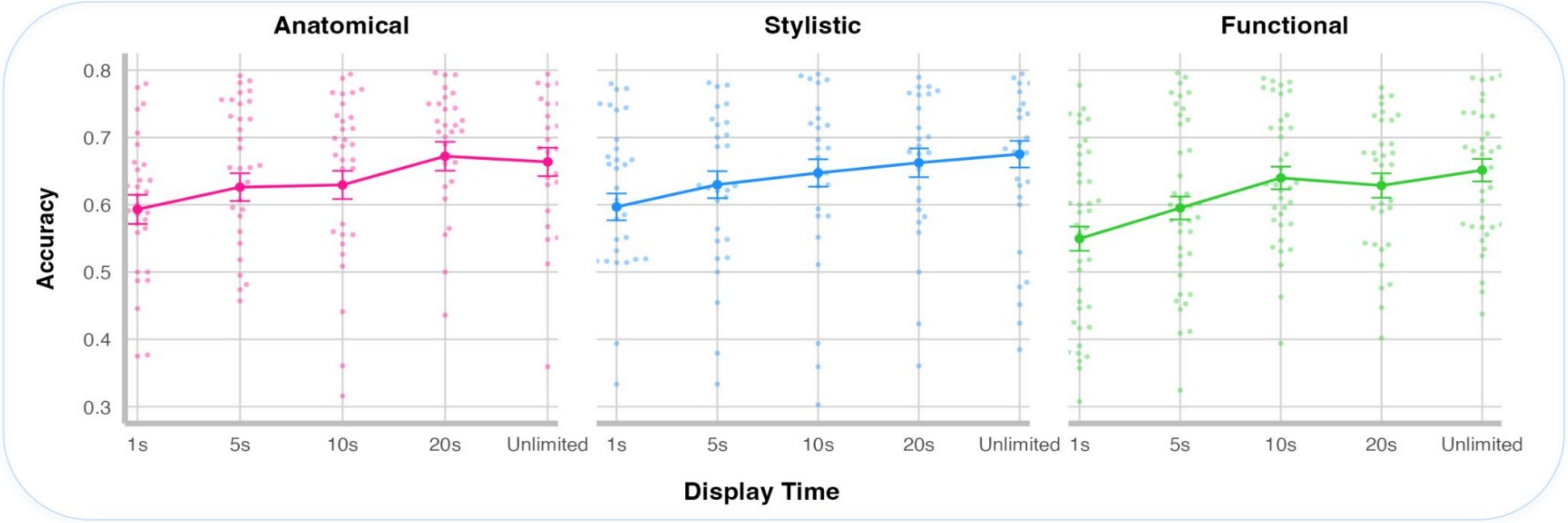
2: Stylistic Artifacts

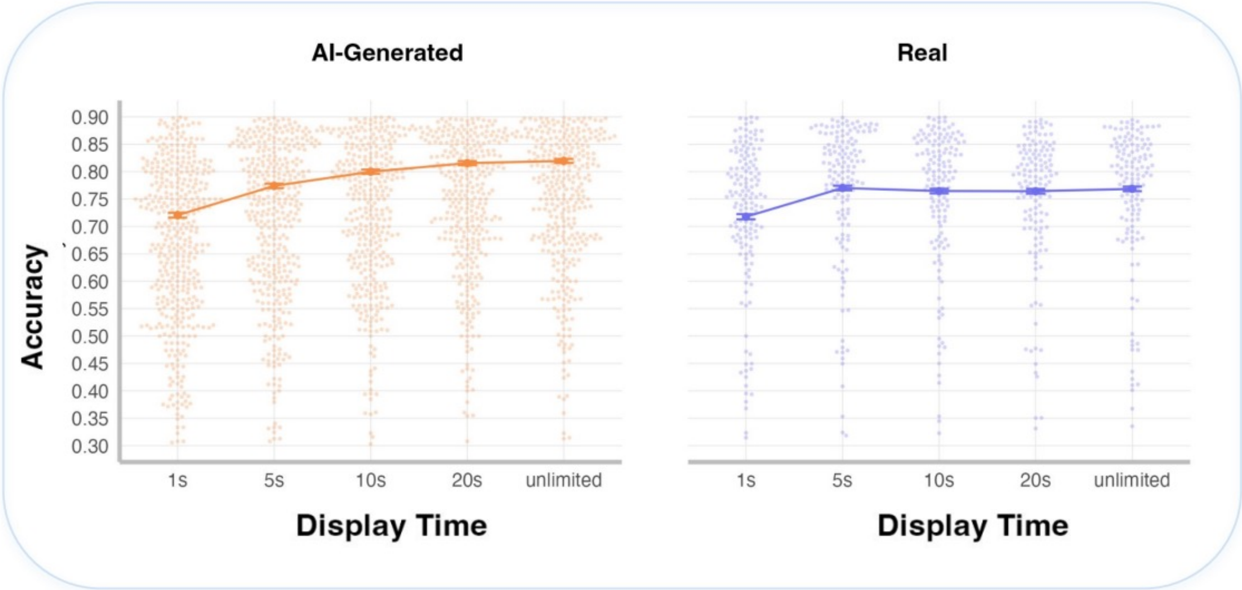
4: Violations of Physics

3: Functional Implausibilities

5: Sociocultural Implausibilities

Accuracy by Artifact Types and Display Times





Future Research

- Use Large Vision Models, Large Language Models, and Large Concept Models to support detection of AI-generated images and redo a user experiment

Modality	Model	Deepfake-Eval-2024 AUC	Prec.	Recall	F1	Original Publication Test Data AUC	Prec.	Recall	F1
Video	GenConViT [8]	0.63	0.60	0.50	0.54	0.96	0.93	0.99	0.96
	FTCN [39]	0.50	0.51	0.67	0.41	0.87	0.91	1.00	0.95
	Styleflow [40]	0.51	0.54	0.43	0.48	0.95	0.96	0.89	0.77
Audio	AASIST [9]	0.43	0.31	0.51	0.39	1.00	1.00	0.95	0.97
	RawNet2 [37]	0.53	0.66	0.39	0.49	0.99	0.60	0.99	0.74
	P3 [38]	0.58	0.36	1.00	0.53	1.00	1.00	0.96	0.98
Image	UFD [35]	0.56	0.63	0.999	0.77	0.94	0.95	0.67	0.75
	DistilDIRE [36]	0.52	0.64	0.87	0.74	0.99	0.99	0.98	0.98
	NPR [10]	0.53	0.69	0.29	0.41	0.98	0.95	0.94	0.94

- Develop **Visual Analytics tools** (like t-viSNE) for this emerging problem to involve the **human-in-the-loop for AI-advised detection** of AI-generated images

Chandra, et al., "Deepfake-Eval-2024: A Multi-Modal In-the-Wild Benchmark of Deepfakes Circulated in 2024", arXiv, 2025

That’s all folks! Thank you!

- t-viSNE Paper: doi.org/10.1109/TVCG.2020.2986996
- Online tool: sheldon.lnu.se/t-viSNE/
- Demo video: vimeo.com/403858988
- Source code: github.com/angeloschatzimparmpas/t-viSNE

Visualization and Graphics Group

Utrecht University