

Machine Learning Meets Visualization - Experiences and Lessons Learned

Quynh Quang Ngo*, Frederik L. Dennig*, Daniel A. Keim, Michael Sedlmair

Abstract: In this article, we discuss how Visualization (VIS) with Machine Learning (ML) could mutually benefit from each other. We do so through the lens of our own experience working at this intersection for the last decade. Particularly we focus on describing how VIS supports explaining ML models and aids ML-based Dimensionality Reduction techniques in solving tasks such as parameter space analysis. In the other direction, we discuss approaches showing how ML helps improve VIS, such as applying ML-based automation to improve visualization design. Based on the examples and our own perspective, we describe a number of open research challenges, as well as limitations combine ML and VIS.

ACM CCS: Human-centered computing → Visual analytics

Keywords: Visual analytics, machine-learning, quality metrics, dimensionality reduction.

1 Introduction

Visualization (VIS) and Machine Learning (ML) are two critical areas for data analysis. On the one hand, ML focuses primarily on learning (predictive) models from large sets of collected data, with the common goal of automatizing a certain task [41]. For example, with many images of animals, we can train a model that enables the computer to tell us, with a certain probability and accuracy, which animals are visible in images that were not among the training data [43]. On the other hand, VIS is mainly concerned with interfaces that present the data in an understandable way and make it accessible for human users [42]. However, often there are ill-defined tasks for which interactive operations can be leveraged to gain insights into the underlying data. Using visualization, a finance expert might, for instance, explore stock data to learn about where to invest next [52]. A biologist might visualize genome data to generate new hypotheses about where a particular disease might stem from [44].

Given that both areas inherently deal with data, there is an intrinsic connection between the two. Often, for instance, data is first visualized to better understand the contained patterns, derive potential hypotheses, and verify that the data collection process has not been faulty. When a task is sufficiently defined, and enough data has been collected, problems can then be modeled with ML and automatized in the next step [56, 35]. However,

beyond this obvious connection, we argue that there are more direct ways of how the two fields of VIS and ML are related and can benefit from each other. Specifically, we argue that VIS can help during the ML model building process. For instance, ML researchers and practitioners need to select hyper-parameters and models, which is often a tedious and lengthy process. Here, visualization can help to provide systematic interfaces to explore and compare different modeling alternatives, deal with multi-objective optimization problems, and learn about uncertainties and sensitivities of models in a rich yet easy-to-access way [55].

On the other hand, VIS can benefit from ML approaches as well. At the moment, for instance, designing good visualizations is still mainly a manual process. The decision of selecting an adjacency matrix-based visualization over a node-link diagram to represent a graph, e.g., is a choice entirely up to the designer who might (or might not) follow existing guidelines [24]. Similarly, it is the designer who needs to select an adequate projection method for multi-dimensional data before it can be shown in a scatterplot [57]. This process necessitates expertise on the designer's side, and wrong decisions can lead to undetected patterns or even misleading representations. Assuming we can collect enough data on these processes, we could instead try to train ML models that help with suggesting good choices and support the designer. Over the last years, such ML-based approaches have become more common in visualization research. A recent survey

*These authors contributed equally to this work.

by Wang and Han [64], for instance, looked at how deep learning can be used for scientific visualization.

Coming primarily from a VIS standpoint, we have worked for a decade on different facets of how VIS and ML might be combined [1, 3, 5, 7, 12, 15, 17, 16, 27, 28, 54, 58, 66, 67]. In this article, we would like to take a step back and reflect on some of the experiences and examples that we studied over the years. We take a broad stand on ML, including widespread supervised learning and unsupervised approaches such as clustering and dimensionality reduction. We also will include other types of optimization approaches that we used that were not necessarily trained from data but allowed us to address structurally similar problems in a new, quantitative way. In the following, we will share a bird’s-eye view of our collective experiences with the hope of providing new inspiration to others working at the intersection of VIS and ML.

2 VIS4ML: Visualization to Improve the Understanding of Machine Learning

Over the recent years, there has been a considerable discussion around explainable artificial intelligence (XAI) and explainable ML. Visual representations can play a key role in XAI as they support communicating complex structures between human and machine. In the keynote for the EuroVis 2017 conference titled “*Visualization: The Secret Weapon of Machine Learning*”[†], Wattenberg and Viegas presented a variety of work demonstrating how VIS could aid in explainability and interpretability for ML. Indeed, this topic has become a booming research trend recently. XAI spans a wide range of topics from supporting debugging models, to deciphering learning processes inside ML models, and fostering education about ML models [2, 26, 29, 72]. Our primary focus has been on two areas, in particular, visual parameter space analysis (VPSA) for ML and XAI through visual interactive learning.

2.1 Visual Parameter Space Analysis for ML

The creation of an ML model often involves setting so-called hyper-parameters such as the number of layers, number of epochs, or the number of dimensions in latent space [31, 65]. To set these parameters, a common approach is to rely on trial and error. Specifically with larger parameter spaces, however, trial and error can become a tedious, unsystematic, and error-prone process; analysts easily forget what exact parameterizations they looked at five minutes ago. A more systematic approach is to instead employ VPSA, i.e., to sample a larger collection of parameter values and visualize the space for the user to explore. If the objectives are well-defined, these steps might also be automated [71]. Yet, often

these types of problems are ill-defined and, as such, necessitate a human-in-the-loop [20, 59]. Multiple different objectives might need to be weighted, objectives might not be even clearly characterized yet, and uncertainties and sensitivities might further influence a decision. To mitigate these problems, we have leveraged VPSA, which primarily has been used for classical simulation models in the past but works similarly well, in our experience, for ML models [55].

As an example, let us assume that we want to find a good dimensionality reduction (DR) model for a given dataset. DR models take multi-dimensional data as input and then output a lower-dimensional projection of the data, either for further usage in the ML pipeline or for the purpose of visualization. In terms of visualization, the output is usually 2D and is typically represented in the form of a scatterplot. Still, before we are able to do that, we need to select among many different DR techniques, such as UMAP [39], t-SNE [63], LLE [48], ISOMAP [62], MDS-based methods [33], PCA [30], etc. [21], and for some of them we also need to set additional parameters. While quantitative error metrics exist, in the end, picking a good visual 2D projection is often still in the eye of the beholder, calling for a human-in-the-loop approach (at least still at the time of writing this article).

Applying VPSA to the space of DR models and parameters, we developed VisCoDeR [15][‡], as shown in Figure 1. The central idea of VisCoDeR is a 2D overview scatterplot, which is called a meta map, that encodes each instance of a parameterized DR model output as a point in the view (see Figure 1 (c)). The meta map is based on a DR model in turn again (here t-SNE, but others are possible). As such, if two points are close in the view, the respective DR instances are similar to each other (see Figure 1 (d)). The distance is computed based on the *visual* similarity of two DR instances (that is, 2D scatterplots), imitating the human perception of the DR output. Like many Visual Analytics tools, VisCoDeR also provides linking and brushing functionality between the different views, most importantly between the meta map (outputs) and the respective input parameter space (see Figure 1 (a)).

Using the tool, analysts can now gain insights into the role of each parameter for the different DR models. For instance, t-SNE [63] requires users to set two parameters, namely perplexity (relates to number of nearest neighbors) and epsilon (learning rate) [68]. The question is now how these parameters affect the visual output of t-SNE and how sensitive they are to changes. To that end, we sampled 1,000 different t-SNE parameterizations, created the respective 2D scatterplots, and visualized them in the VisCoDeR meta map. Using linking and brushing between the meta map and input parameter space, we can now smoothly hover over the two parameterized, see Figure 1(a). This interaction reveals, that

[†]<https://www.youtube.com/watch?v=E701G9-HGEM>

[‡]<https://reencutura.eu/viscoder/discover.html>

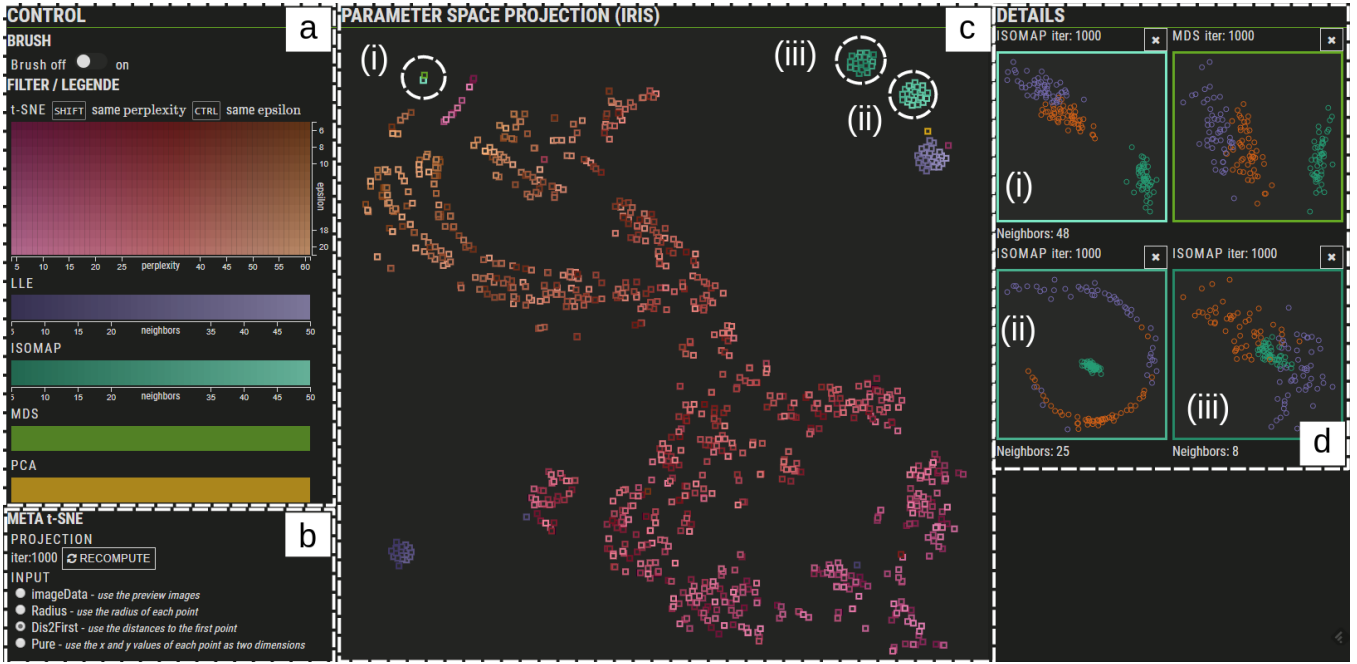


Figure 1: The VisCoDeR system [15] offers an interactive color legend for DR algorithms and their parameterizations (a), meta-parameterization and control over the projection technique (b), meta map of 1004 DR results with activated proximity visualization (c), and a detailed view of the selected DR result (d), allowing for visual parameter space analysis of DR methods.

perplexity has a smooth sensitivity, that is, the visual output changes gradually with the change of the perplexity parameter, see the center of Figure 1(c) where the color encoding for perplexity changes from orange to pink. red In this area, t-SNE scatterplots with the same or similar perplexity are close in the meta map. On the other hand, we can see that epsilon, which is encoded by the brightness channel has seemingly little impact on the t-SNE outcome as t-SNE plots with the same epsilon appear all over the place in the meta map.

In this example, VPSA helped us to conduct a sensitivity analysis of t-SNE on a given dataset. Other tasks that are supported are multi-objective optimization, uncertainty analysis, partitioning, outlier detection and fitting [55]. VPSA is by no means restricted to DR models, but works on all sorts of input-output-based models, which is the case for many ML approaches. We, for instance, also used to make the exploration of hyper-parameters of classification models more systematic [28], and also for deep learning models [27].

2.2 Explainable AI with Visual Interactive Learning

Another important use case for leveraging VIS of ML is Active Learning (AL). In AL, only few labels are available, and a user needs to be prompted to provide more and more labels on the way, following a strategy that maximizes the effectiveness of the labeling process by requesting labels for data items with an uncertain classification result. The new labels are used in subsequent steps to improve the underlying model. The model is trained on sparse data. Thus the model needs to be assessed

for quality and convey its reasoning, especially when the AL classifier is prompting for a label. As such, there is an intrinsic need for human-computer interaction. While classical Active Learning leverages primarily simple labeling interfaces, VIS allows much more sophisticated ways for users to interact with the models, supporting not just active but also inter-active learning [49].

One way to use VIS in the AL pipeline is to support users in understanding the space of (labeled and unlabeled) instances, with the idea that the visualization can help users to select, implicitly or explicitly, further items for labeling. To that end, we developed a system called FDive [17], which learns to distinguish interesting from uninteresting data through an iteratively improving classifier (see Figure 2). To do so, it uses a set of feature descriptors and distance functions to represent the similarity of data items as a distance measure. This allows for interpretable distances, where analysts can derive which data properties are essential for the classification. First, users can express their preference for a data item by labeling a set of items prompted by the system (see Figure 2 (1)). Second, these labels are used to choose a feature descriptor and distance function combination by their ability to represent the user’s preference through distance relations (see Figure 2 (2)).

Finally, the system applies the selected combination to learn a Self-Organizing Map-based classifier (see Figure 2 (3)). This special type of model can be explored and refined by supplying labels at any position. However, uncertain classifications are highlighted to guide users toward items with uncertain labels. This process can be repeated to improve the classification and assess it. We

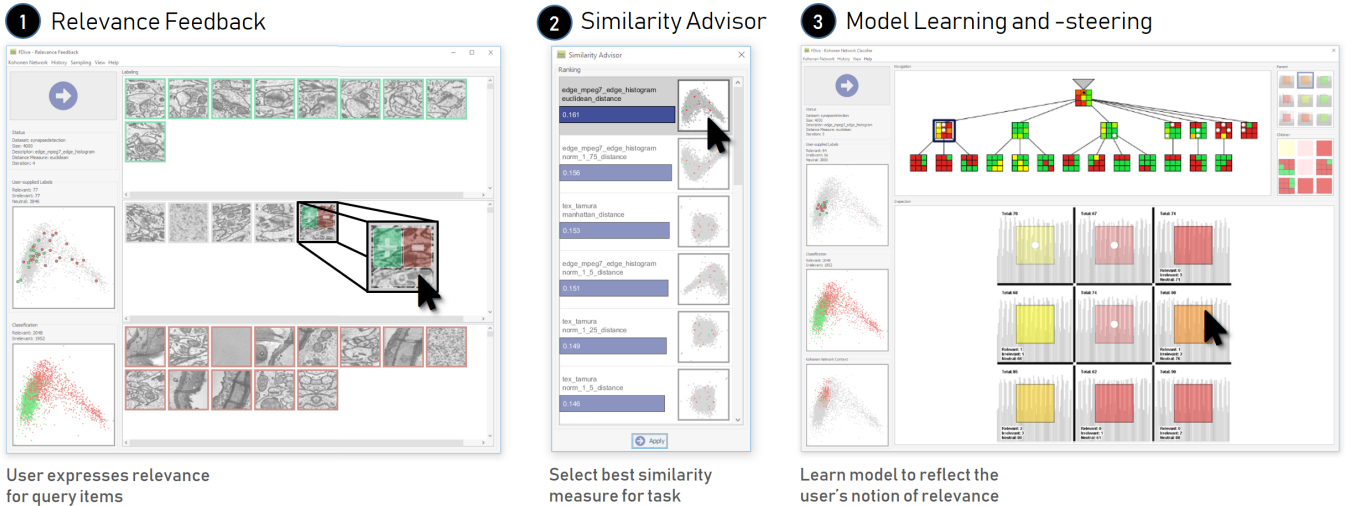


Figure 2: FDive [17] learns to classify interesting from uninteresting data items via an iterative active learning process. It chooses the best-fitting feature descriptor and distance function to reflect the users' notion of interestingness as a distance measure. (1) Users select interesting data items labeling a set of query items. (2) The labels determine similarity measures, i.e., feature descriptor and distance function. (3) The system uses the selected distance measure to learn a classification model. The user can explore and refine the model by supplying more labels for uncertain data items near the decision boundaries. © IEEE 2019

applied our tool to connectomics, a sub-field of neurology where scientists try to map out neuronal connections using electron microscopy images to detect neuronal synapses. There are roughly one billion synapses in 1 mm³ of a brain tissue. Thus, methods of automatically detecting images depicting neuronal synapses are necessary.

With FDive, the analyst can label a small subset of images determined by the system, whether they show neuronal synapses or not (see Figure 2 (1)). The system selects the best separating measure for the labeled images. The analyst can observe the implied data space through an MDS projection of the dataset (see Figure 2 (2)). We found that descriptors focusing on the image texture generally worked the best for this task, since the system immediately converged on a texture-based description of the images. Finally, the analyst can explore the model to determine the classifier's quality.

The classifier has a nested hierarchy, allowing for a more and more fine-grained classification. Clusters with an uncertain classification are highlighted to guide the analyst towards them. The analyst can supply extra labels for the images in those clusters to improve the classifier. We observed that images containing similar cell structures are clustered, including those depicting a neuronal synapse (see Figure 2 (3)). The analyst repeated that process seven times, iteratively supplying more labels until the analyst considered the model adequate while converging on a specific texture descriptor.

It is an interesting question when to use which type of interface. Simple problems such as image classification with a few classes probably are dealt well with a simple labeling interface that is prompted by the algorithm on demand. Complex and ill-defined domain problems, such as connection classification in neurology, on the other hand, might benefit from close integration of labeling

and analytical components. First studies are available that explicitly seek to characterize this space [7, 8] and understand how to visually encode the data in such cases [25].

3 ML4VIS: Improving Visualization with Machine Learning

As described above, VIS plays a central role in making ML explainable and support model building. However, we argue that, vice versa, VIS can similarly benefit from the application of ML.

One of the main application areas of ML in VIS is whether we can use ML to automatize to at least guide several steps in the VIS design process [35]. In fact, "How to design a good visualization?" is one of the grand challenges in visualization. The goal of the visualization design process is to aggregate and encode the data in a way that reveals interesting structures and patterns in the data. To this end, visualizations need to (1) be readable and uncluttered [10], (2) be optimized toward the analysis task, and (3) be tailored to the user's prior knowledge [6].

Automatizing the visualization design process is not a new idea. Already back in 1986, Mackinlay [37] thought to automate the design of graphical presentations through a thorough formalization of the process. As visualization design is intrinsically a perceptual and cognitive process, it stands to reason that modern ML approach could be a good fit for that goal as well. While for a long time, this idea has caught surprisingly little attention, a few years ago, researchers have started investigating the topic more and more [14, 18, 34, 36, 46, 50, 70].

Our work in this area so far has primarily focused on

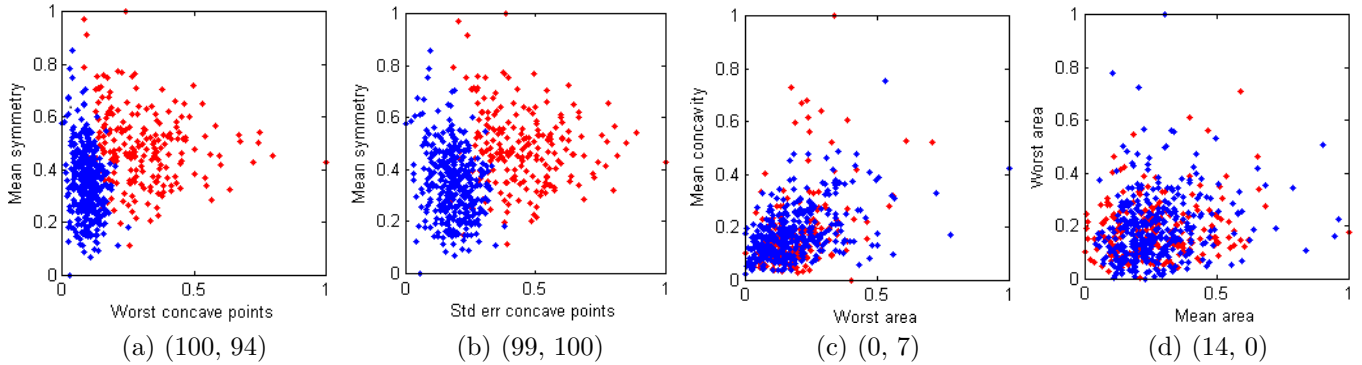


Figure 3: 2D projections of a dataset showing the two classes as red and blue dots. (a, b) show pairs of dimensions with visually good class separation and (c, d) with poor separation. The scores below are two separation measures. The first value shows a γ -Observable Neighbor Graph-based measure and the second value a Distance Consistency measure. They are within 100 for the best separability and 0 for the worst separability [3]. © IEEE 2016

training perceptual models to solve a specific task with a given visualization. So instead of seeking a full end-to-end model automatizing the entire VIS pipeline (and as such likely necessitating huge amounts of training data), we took a bottom-up approach first. To that end, we started our work with learning perceptual models for scatterplots [51], which—according to Ron Rensink—are the “fruit flies of visualization research” [47]. That is, scatterplots are simple enough to control for confounding factors, but rich enough to cover much of the underlying complexity of visual perception and analysis.

In 2015, we proposed a simple framework to implement our idea, consisting of the following three steps [54]: (1) gather an extensive collection of perceptual “ground truth data” from human subject studies, e.g., let participants judge whether a scatterplot shows separated classes or not, (2) predict these judgments with different “models”, e.g., let the model determine the class separation of the plot, (3) evaluate the quality of each “model”, e.g., use the accuracy and generalization to determine the quality of the prediction. With this approach, we essentially set out to train perceptual models of users from empirical data and to use that to automatically select existing models that imitate these users—a typical approach of classical ML. We instantiated this framework with different examples, as described in the following.

3.1 Class Separability for Scatterplots

Following a series of earlier empirical work [58, 57, 11, 53], the first task we were interested in was judging class separability in scatterplots [54, 3]. In Figure 3, (a) and (b) show two separable classes while (c) and (d) show two non-separable classes that are easy for humans to differentiate. The basic idea was then to use ML to train a model that imitates and predicts these human judgments and that rates class separability like humans do. Having such a model would then allow us to automatically spot “interesting” views in large scatterplot matrices [60, 69], guide the selection of DR methods [57], or even search dimensional subspaces [61]—in a way that humans would.

The idea of modeling class separability originated from Sips et al. [60], who proposed hand-crafted measures for that purpose, in a similar vein as the venerable Scagnostics measures [69]. When using these measures for a large collection of 816 scatterplots, we, however found the generalization to other (“unseen”) datasets was poor [58]. As generalizability is a strength of ML, we thus were wondering in how far ML could lead to better models and measures for class separability. With that in mind, we used a carefully collected and cleaned dataset of expert judgments [57] and trained a binary classifier (separable or non-separable) for class separation in scatterplots [54]. Bootstrapping was used to ensure generalizability to other unseen datasets.

Using this approach, we then proposed and automatically evaluated 2002 systematically generated separation measures/models [3]. Using these in the model selection phase of the framework, we indeed found that many of our novel measures substantially outperformed the best state-of-the-art measures. While the best state-of-the-art measure had an accuracy of 82.5% (bootstrapped AUC-ROC), the best new measure had an average accuracy of 92.7%, and overall, 58% of the new measures outperformed the traditional best measure. Of course, our work here is just a starting point. Our proposed model is relatively simple and is still far from nuanced human perception in judging class separability. Still, already with the simple model, we got very good performance, indicating that an ML approach might be a good fit here.

3.2 Cluster Detection in Scatterplots

Another widespread and closely-related task in scatterplots is identifying classes in analysis scenarios where the data is unlabeled [11]. Here the analyst has to deal with monochrome scatterplots. The goal is to identify cluster structures, meaning that groups of data items are separated visually, either by an empty area where no data items are located or by differences in density for overlapping or nested clusters. There are multiple approaches to tackle this problem. Firstly, one could apply clustering

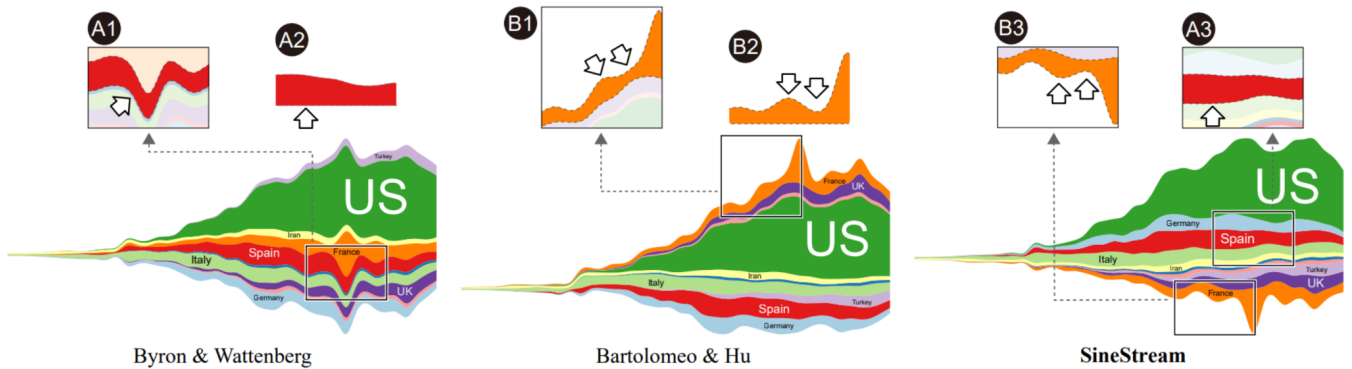


Figure 4: SineStream [12] minimizes the impact of sine illusion effects originating from strong slopes. The arrows highlight parts of the streamgraph, where SineStream [12] represents the thickness of layers more accurately than methodologies by Bryon & Wattenberg [13] and by Bartolomeo & Hu [4]. It improves the readability of streamgraphs by aligning each layer’s orthogonal and vertical orientations. © IEEE 2022

algorithms, such as DBSCAN [32] or CLIQUE [23] to the scatterplot visualization (i.e., image space), trying to detect well-separated clusters. Secondly, this problem also has been tackled with the idea of classical quality measures, such as the “clumpiness” Scagnostics measure [69]. However, we were wondering in how far such heuristic approaches can capture the human understanding of a cluster, which might include the notion of non-globular clusters, clusters that are only separated visually by a small gap, or clusters that are nested, differing only in density. We hypothesized that a more nuanced approach could be useful to capture the notion of what defines a visual cluster from a human perspective and whether ML could provide a better solution for this problem.

We thus proposed ClustMe [1], an ML-based approach to the idea of “clumpiness”. We built ClustMe based on data collected from a study with 34 participants. The participants judged the cluster patterns in 1000 scatterplots of synthetically generated datasets. We generated the datasets by adapting the parameters of a simple Gaussian Mixture Model with two components. To quantify the number of clusters in a scatterplot, the participants counted the number of clusters they could see.

We then created ClustMe by choosing the model that best predicts the human intuition. We performed another study to evaluate ClustMe, in which 31 study participants ranked 435 pairs of scatterplots of real-world and generated data in terms of perceived cluster patterns. We then compared the performance of ClustMe to four other state-of-the-art clustering measures using this data. We also included the “clumpiness” measure of Scagnostics in this comparison. The results showed that ClustMe, out of all measurements, was most consistent with the human rankings. This work again showed evidence that an ML-based approach can mimic classical quality measures and, as such, can be used to improve VIS.

3.3 Example Beyond Scatterplots

While scatterplots are a good starting point, practical visualization, of course, offers many more other types of visual encodings. These can also benefit from a similar modeling approach that helps to automatically optimize parameters of visualizations [40].

We, for instance, developed SineStream [12] to push forward the current state-of-the-art for streamgraph visualizations (see Figure 4). Following a similar process as on class separability above, SineStream was based on the main idea of improving readability by minimizing sine illusion effects in streamgraphs. Such effects reflect the tendency of humans to take the orthogonal rather than the vertical distance between two curves as their distance. In SineStream, we minimize this illusion by optimizing the ordering of the different streams. Quantitative experiments and user studies demonstrated that SineStream improved the readability and aesthetics of streamgraphs compared to state-of-the-art methods.

In comparison to the approaches in Sections 3.1 and 3.2, we, however did not use ML in this example, but classical optimization (simulated annealing in this case). In vain, the idea to automatically optimize visual parameters based on human perception is the same, though. An interesting question for future work is, of course, how far ML might be able to provide further improvements for such approaches as well.

4 Discussion

The previous sections described examples of how machine learning and visualization can benefit each other. There were several challenges that we faced when seeking to combine the two fields, though, which we will discuss in the following. Before concluding, we also want to highlight the limitations of the current state of our work.

4.1 Challenges

Based on our experience working on these topics, we now want to take a step back and reflect on some challenges in combining ML and VIS.

Domain problems with increasing complexity:

Both, ML and VIS can often be seen as providing a “service” to other domain-specific problems. Data-driven approaches leveraging ML and/or VIS have become a standard approach in many application domains now. Tackling the complexities of solving these domain problems, however, often requires a collaborative effort combining expertise from different fields. In our own work, we sought to address this challenge by following a user-centered design process that is fine-tuned to the needs of solving ill-defined problems via data analytics: design study methodology [56]. We found that design studies foster a good way to do interdisciplinary research including VIS researcher, ML researchers (ML), and domain experts, for projects that require to go beyond the scope of “just” combining ML and VIS. There are different ways of how design studies can be initiated. Traditionally, one would start with characterizing the problem through working with domain experts. However, there are also data-first design studies [45] where VIS/ML researchers actually start with visualizing/modeling the data, in order to ideate potential problems that might be solved with the data at hand or identify the respective target groups. Along this line, we still see a potential gap of model-first design studies where one would start with certain types of models and reach out to domain applications afterwards, at least for ill-defined problems.

Interaction: Our work on combining ML and VIS so far has primarily focused on visual encoding. For the entire visual analysis process, however, interaction is equally important. Closely intertwining analytical interactions with ML components bears huge potentials. Recently, Fan and Hauser have, for instance, shown how ML-based approaches can be used to substantially improve linking and brushing interactions, which are a corner stone in many multi-view visual analytics systems. More generally, Endert et al. [19] developed a toolkit which allows direct interactions to a scatterplot and change location of two points based on domain expert knowledge. The model, which is a dimensionality reduction technique, is then updated accordingly. The outcome scatterplot is now a combination of the result from the model and expert knowledge. In the same ground, Inter-active Learning [9] provides interaction to leverage analytical human labeling in the computational modeling process.

A long-standing challenge related to interaction is the question when a human-in-the-loop is actually needed and when we can simply automatize a problem. In our early work [56], we argued that it depends primarily on two factors: information location and task clarity. Information location can be in the head of the analyst or it can be externalized into a computer. The latter is a

critical prerequisite for automation. Task clarity depicts how well- or ill-defined a problem is. Ill-defined problems need to be clarified by a human-in-the-loop. At the moment, automatic ML approaches shine at well-defined tasks. Many important scientific problems, however, are by definition ill-defined [59]. The interesting question is in how far ML-based automatic approaches might become capable to address more ill-defined problems in the future. For sure, the ML advancements over the last decade have been impressive. At the same time, problems are becoming more complex (see above), calling for collaboration between increasingly powerful ML models and increasingly skilled analysts and experts. Characterizing this dynamically changing space is an ongoing and highly interesting challenge at the intersection of VIS and ML.

Data acquisition: One ML-relevant issue, in general, is the acquisition of data. ML is data-intensive, and a lot of labeled data is required. In the case of applying ML for VIS based on perceptual, we need user studies that generate many labeled visualizations describing what humans see in it. ClustMe [1] was built on a study with 1000 scatterplots and 34 participants; for other tasks, the scope of such a study might be much larger. Today, we can use Amazon’s Mechanical Turk to gather label data from a large population. Yet, setting up such studies to collect clean and valuable data can be a challenge, even for seemingly “easy” perceptual tasks such as class separability or cluster separation. In ClustMe [1], participants only labeled scatterplots into two classes: “only one blob” vs. “more than one blob”. When collecting enough data, even these very simple tasks can generate an interesting distribution across different users. For other tasks, the labeling schema might not be that clear cut, though, and participants might disagree more on the presence of a pattern. Thus, it is critical to define tasks accurately and find strategies to resolve disagreements.

Data scale and usage: Another challenge is that VIS and ML operate under different paradigms when it comes to the amount and usage of data. The VIS applications normally use relatively small datasets compared to the ML applications. The main reason is that most visualization needs to be interactive and need to update in milliseconds for users to stay engaged. This limitation does not apply to ML techniques, where a system does not need to be that responsive to queries by the user. One step that is done towards responsive visualization with large datasets is Progressive Analytics [22].

4.2 Limitations

We would like to explicitly remind the reader that these above contributions are only from our own experiences. We would like to encourage further collaborations between the VIS and ML approaches with this article. The presented approaches should not be seen as a comprehensive characterization of all combinations of VIS and ML, but only as some examples. Of course, there are

many other ways to combine ML and VIS.

From the approaches presented in this article, we would like to highlight two concrete limitations. First, we see one current limitation in the area of explaining ML models through effective interactive interfaces [15, 17]. Currently, most systems are based on the linking and brushing technique. While this is a strong technique, for effective integration of VIS and ML, we will also need semantic interaction [19], interactive learning [7], and maybe even in devices other than mouse and keyboard [38].

Second, in using ML to improve VIS, we found that there are limitations to modeling human perception. ClustMe [1] and SepMe [3] are two approaches that heavily rely on the data gathered in studies. With these ML techniques, we are steering into the domain of User Modeling. Such a model can only express a general notion for the task at hand since ML techniques can only aggregate the results of a study and thus can differ from an individual's perception and judgment.

5 Conclusion

We presented a flashback of our own experiences of how machine learning and visualization can benefit from each other. We took a step back, reflecting what we have learned on different ways to bridge the two fields of ML and VIS. From our experiences and own perspectives, we also identified two critical challenges and two limitations faced when combining machine learning and visualization, which we see as research opportunities for future work and hope that others will join us in working on these interesting topics.

Acknowledgement

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the projects A03 and A08 of TRR 161 (Project-ID 251654672).

Literature

- [1] Mostafa M. Abbas, Michaël Aupetit, Michael Sedlmair, and Halima Bensmail. Clustme: A visual quality measure for ranking monochrome scatterplots based on cluster patterns. *Computer Graphics Forum*, 38(3):225–236, 2019.
- [2] Amina Adadi and Mohammed Berrada. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6:52138–52160, 2018.
- [3] Michaël Aupetit and Michael Sedlmair. Sepme: 2002 new visual separation measures. In Chuck Hansen, Ivan Viola, and Xiaoru Yuan, editors, *2016 IEEE Pacific Visualization Symposium*, pages 1–8. IEEE Computer Society, 2016.
- [4] Marco Di Bartolomeo and Yifan Hu. There is more to streamgraphs than movies: Better aesthetics via ordering and lassoing. *Computer Graphics Forum*, 35(3):341–350, 2016.
- [5] Emma Beauxis-Aussalet, Michael Behrisch, Rita Borgo, Duen Horng Chau, Christopher Collins, David S. Ebert, Mennatallah El-Assady, Alex Endert, Daniel A. Keim, Jörn Kohlhammer, Daniela Oelke, Jaakko Peltonen, Maria Riveiro, Tobias Schreck, Hendrik Strobelt, Jarke J. van Wijk, and Theresa-Marie Rhyne. The role of interactive visualization in fostering trust in AI. *IEEE Computer Graphics and Applications*, 41(6):7–12, 2021.
- [6] Michael Behrisch, Michael Blumenschein, Nam Wook Kim, Lin Shao, Mennatallah El-Assady, Johannes Fuchs, Daniel Seebacher, Alexandra Diehl, Ulrik Brandes, Hanspeter Pfister, Tobias Schreck, Daniel Weiskopf, and Daniel A. Keim. Quality metrics for information visualization. *Computer Graphics Forum*, 37(3):625–662, 2018.
- [7] Jürgen Bernard, Marco Hutter, Matthias Zeppelzauer, Dieter W. Fellner, and Michael Sedlmair. Comparing visual-interactive labeling with active learning: An experimental study. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):298–308, 2018.
- [8] Jürgen Bernard, Matthias Zeppelzauer, Michael Sedlmair, and Wolfgang Aigner. VIAL: a unified process for visual interactive labeling. *Visual Computer*, 34(9):1189–1207, 2018.
- [9] Jürgen Bernard, Marco Hutter, Matthias Zeppelzauer, Dieter Fellner, and Michael Sedlmair. Comparing visual-interactive labeling with active learning: An experimental study. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):298–308, 2018.
- [10] Enrico Bertini and Giuseppe Santucci. Visual quality metrics. In Enrico Bertini, Catherine Plaisant, and Giuseppe Santucci, editors, *Proceedings of the 2006 AVI Workshop on Beyond time and errors: novel evaluation methods for information visualization*, pages 1–5. ACM Press, 2006.
- [11] Matthew Brehmer, Michael Sedlmair, Stephen Ingram, and Tamara Munzner. Visualizing dimensionally-reduced data: Interviews with analysts and a characterization of task sequences. In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, pages 1–8, 2014.
- [12] Chuan Bu, Quanjie Zhang, Qianwen Wang, Jian Zhang, Michael Sedlmair, Oliver Deussen, and Yunhai Wang. Sinestream: Improving the readability of streamgraphs by minimizing sine illusion effects. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1634–1643, 2021.
- [13] Lee Byron and Martin Wattenberg. Stacked graphs - geometry & aesthetics. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1245–1252, 2008.
- [14] Hsueh-Chien Cheng, Antonio Cardone, Somay Jain, Eric Krokos, Kedar Narayan, Sriram Subramaniam, and Amitabh Varshney. Deep-learning-assisted volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 25(2):1378–1391, 2019.
- [15] René Cutura, Stefan Holzer, Michaël Aupetit, and Michael Sedlmair. Viscoder: A tool for visually comparing dimensionality reduction algorithms. In *26th European Symposium on Artificial Neural Networks*, 2018.
- [16] Frederik L. Dennig, Maximilian T. Fischer, Michael Blumenschein, Johannes Fuchs, Daniel A. Keim, and Evanthia Dimara. Parsetgnostics: Quality metrics for parallel sets. *Computer Graphics Forum*, 40(3):375–386, 2021.
- [17] Frederik L. Dennig, Tom Polk, Zudi Lin, Tobias Schreck, Hanspeter Pfister, and Michael Behrisch. Fdive: Learning relevance models using pattern-based similarity measures. In Remco Chang, Daniel A. Keim, and Ross Maciejewski, editors, *14th IEEE Conference on Visual Analytics Science and Technology*, pages 69–80. IEEE, 2019.
- [18] Victor Dibia and Çagatay Demiralp. Data2vis: Automatic generation of data visualizations using sequence-to-sequence recurrent neural networks. *IEEE Computer Graphics and Applications*, 39(5):33–46, 2019.
- [19] Alex Endert, Patrick Fiaux, and Chris North. Semantic interaction for visual text analytics. In Joseph A. Konstan, Ed H. Chi, and Kristina Höök, editors, *CHI Conference on Human Factors in Computing Systems*,

- pages 473–482. ACM, 2012.
- [20] Alex Endert, Mahmud Shahriar Hossain, Naren Ramakrishnan, Chris North, Patrick Fiaux, and Christopher Andrews. The human is the loop: new directions for visual analytics. *Journal of Intelligent Information Systems*, 43(3):411–435, 2014.
- [21] Mateus Espadoto, Rafael Messias Martins, Andreas Kerren, Nina S. T. Hirata, and Alexandru C. Telea. Toward a quantitative survey of dimension reduction techniques. *IEEE Transactions on Visualization and Computer Graphics*, 27(3):2153–2173, 2021.
- [22] Jean-Daniel Fekete, Danyel Fisher, Arnab Nandi, and Michael Sedlmair. Progressive data analysis and visualization (dagstuhl seminar 18411), 2018.
- [23] Anna Förster and Amy L. Murphy. CLIQUE: role-free clustering with q-learning for wireless sensor networks. In *29th IEEE International Conference on Distributed Computing Systems*, pages 441–449. IEEE Computer Society, 2009.
- [24] Mohammad Ghoniem, Jean-Daniel Fekete, and Philippe Castagliola. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. *Information Visualization*, 4(2):114–135, 2005.
- [25] Nicolas Grossmann, Jürgen Bernard, Michael Sedlmair, and Manuela Waldner. Does the layout really matter? A study on visual model accuracy estimation. In *2021 IEEE Visualization Conference*, pages 61–65. IEEE, 2021.
- [26] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5):93:1–93:42, 2019.
- [27] Sagad Hamid, Adrian Derstroff, Sören Klemm, Quynh Quang Ngo, Xiaoyi Jiang, and Lars Linsen. Visual ensemble analysis to study the influence of hyperparameters on training deep neural networks. In Daniel Archambault, Ian T. Nabney, and Jaakko Peltonen, editors, *2nd Workshop on Machine Learning Methods in Visualisation for Big Data*, pages 19–23. Eurographics Association, 2019.
- [28] Frank Heyen, Tanja Munz, Michael Neumann, Daniel Ortega, Ngoc Thang Vu, Daniel Weiskopf, and Michael Sedlmair. Clavis: An interactive visual comparison system for classifiers. In Genny Tortora, Giuliana Vitiello, and Marco Winckler, editors, *AVI '20: International Conference on Advanced Visual Interfaces*, pages 9:1–9:9. ACM, 2020.
- [29] Fred Hohman, Minsuk Kahng, Robert S. Pienta, and Duen Horng Chau. Visual analytics in deep learning: An interrogative survey for the next frontiers. *IEEE Transactions on Visualization and Computer Graphics*, 25(8):2674–2693, 2019.
- [30] Ian T. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics. Springer, 1986.
- [31] Minsuk Kahng, Nikhil Thorat, Duen Horng (Polo) Chau, Fernanda B. Viégas, and Martin Wattenberg. GAN lab: Understanding complex deep generative models using interactive visual experimentation. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):310–320, 2019.
- [32] Kamran Khan, Saif ur Rehman, Kamran Aziz, Simon Fong, Sababady Sarasvady, and Amrita Vishwa. DBSCAN: past, present and future. In *The Fifth International Conference on the Applications of Digital Information and Web Technologies*, pages 232–238. IEEE, 2014.
- [33] J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.
- [34] Fritz Lekschas, Brant Peterson, Daniel Haehn, Eric Ma, Nils Gehlenborg, and Hanspeter Pfister. Peax : Interactive visual pattern search in sequential data using unsupervised deep representation learning. *Computer Graphics Forum*, 39(3):167–179, 2020.
- [35] Yuyu Luo, Xuedi Qin, Nan Tang, and Guoliang Li. Deepeye: Towards automatic data visualization. In *34th IEEE International Conference on Data Engineering, ICDE 2018, Paris, France, April 16-19, 2018*, pages 101–112. IEEE Computer Society, 2018.
- [36] Yuxin Ma, Anthony K. H. Tung, Wei Wang, Xiang Gao, Zhigeng Pan, and Wei Chen. Scatternet: A deep subjective similarity model for visual analysis of scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 26(3):1562–1576, 2020.
- [37] Jock D. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141, 1986.
- [38] Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas, editors. *Immersive Analytics*, volume 11190 of *Lecture Notes in Computer Science*. Springer, 2018.
- [39] Leland McInnes and John Healy. UMAP: uniform manifold approximation and projection for dimension reduction. *CoRR*, abs/1802.03426, 2018.
- [40] Luana Micallef, Gregorio Palmas, Antti Oulasvirta, and Tino Weinkauf. Towards perceptual optimization of the visual design of scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1588–1599, 2017.
- [41] Tom M. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.
- [42] T. Munzner. *Visualization Analysis and Design*. AK Peters Visualization Series. CRC Press, 2015.
- [43] Mohammad Sadeh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, and Jeff Clune. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences USA*, 115(25):E5716–E5725, 2018.
- [44] Sabrina Nusrat, Theresa Harbig, and Nils Gehlenborg. Tasks, techniques, and tools for genomic data visualization. *Computer Graphics Forum*, 38(3):781–805, 2019.
- [45] M. Oppermann and T. Munzner. Data-first visualization design studies. In *2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV)*, pages 74–80, Los Alamitos, CA, USA, oct 2020. IEEE Computer Society.
- [46] Jaakko Peltonen and Ziyuan Lin. Information retrieval approach to meta-visualization. *Machine Learning*, 99(2):189–229, 2015.
- [47] Ronald A. Rensink. On the prospects for a science of visualization. In Weidong Huang, editor, *Handbook of Human Centric Visualization*, pages 147–175. Springer, 2014.
- [48] Sam T. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
- [49] Dominik Sacha, Michael Sedlmair, Leishi Zhang, John Aldo Lee, Jaakko Peltonen, Daniel Weiskopf, Stephen C. North, and Daniel A. Keim. What you see is what you can change: Human-centered machine learning by interactive visualization. *Neurocomputing*, 268:164–175, 2017.
- [50] Bahador Saket, Dominik Moritz, Halden Lin, Victor Dibia, Çagatay Demiralp, and Jeffrey Heer. Beyond heuristics: Learning visualization design. *CoRR*, abs/1807.06641, 2018.
- [51] Alper Sarikaya and Michael Gleicher. Scatterplots: Tasks, data, and designs. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):402–412, 2018.
- [52] Tobias Schreck, Tatiana Tekusová, Jörn Kohlhammer, and Dieter W. Fellner. Trajectory-based visual analysis

of large financial time series data. *SIGKDD Explorations*, 9(2):30–37, 2007.

- [53] M Sedlmair, Matt Brehmer, S Ingram, and T Munzner. Dimensionality reduction in the wild: Gaps and guidance. *Dept. Comput. Sci., Univ. British Columbia, Vancouver, BC, Canada, Tech. Rep. TR-2012-03*, 2012.
- [54] Michael Sedlmair and Michaël Aupetit. Data-driven evaluation of visual quality measures. *Computer Graphics Forum*, 34(3):201–210, 2015.
- [55] Michael Sedlmair, Christoph Heinzl, Stefan Bruckner, Harald Piringer, and Torsten Möller. Visual parameter space analysis: A conceptual framework. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2161–2170, 2014.
- [56] Michael Sedlmair, Miriah D. Meyer, and Tamara Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012.
- [57] Michael Sedlmair, Tamara Munzner, and Melanie Tory. Empirical guidance on scatterplot and dimension reduction technique choices. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2634–2643, 2013.
- [58] Michael Sedlmair, A. Tatu, Tamara Munzner, and Melanie Tory. A taxonomy of visual cluster separation factors. *Computer Graphics Forum*, 31(3pt4):1335–1344, 2012.
- [59] Herbert A. Simon. The structure of ill structured problems. *Artificial Intelligence*, 4(3):181–201, 1973.
- [60] Mike Sips, Boris Neubert, John P. Lewis, and Pat Hanrahan. Selecting good views of high-dimensional data using class consistency. *Computer Graphics Forum*, 28(3):831–838, 2009.
- [61] Andrada Tatu, Fabian Maass, Ines Färber, Enrico Bertini, Tobias Schreck, Thomas Seidl, and Daniel A. Keim. Subspace search and visualization to make sense of alternative clusterings in high-dimensional data. In *7th IEEE Conference on Visual Analytics Science and Technology*, pages 63–72. IEEE Computer Society, 2012.
- [62] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.
- [63] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.
- [64] Chaoli Wang and Jun Han. DL4sciviz: A state-of-the-art survey on deep learning for scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, 2022.
- [65] Junpeng Wang, Liang Gou, Han-Wei Shen, and Hao Yang. Dqnviz: A visual analytics approach to understand deep q-networks. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):288–298, 2019.
- [66] Yunhai Wang, Xin Chen, Tong Ge, Chen Bao, Michael Sedlmair, Chi-Wing Fu, Oliver Deussen, and Baoquan Chen. Optimizing color assignment for perception of class separability in multiclass scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):820–829, 2019.
- [67] Yunhai Wang, Kang Feng, Xiaowei Chu, Jian Zhang, Chi-Wing Fu, Michael Sedlmair, Xiaohui Yu, and Baoquan Chen. A perception-driven approach to supervised dimensionality reduction for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(5):1828–1840, 2018.
- [68] Martin Wattenberg, Fernanda Viégas, and Ian Johnson. How to use t-sne effectively. *Distill*, 2016.
- [69] Leland Wilkinson, Anushka Anand, and Robert L. Grossman. Graph-theoretic scagnostics. In John T. Stasko and Matthew O. Ward, editors, *IEEE Symposium on Information Visualization*, pages 157–164. IEEE Computer

Society, 2005.

- [70] Aoyu Wu, Yun Wang, Xinhuan Shu, Dominik Moritz, Weiwei Cui, Haidong Zhang, Dongmei Zhang, and Huamin Qu. Survey on artificial intelligence approaches for visualization data. *CoRR*, abs/2102.01330, 2021.
- [71] Quanming Yao, Mengshuo Wang, Hugo Jair Escalante, Isabelle Guyon, Yi-Qi Hu, Yu-Feng Li, Wei-Wei Tu, Qiang Yang, and Yang Yu. Taking human out of learning applications: A survey on automated machine learning. *CoRR*, abs/1810.13306, 2018.
- [72] Jun Yuan, Changjian Chen, Weikai Yang, Mengchen Liu, Jiashi Xia, and Shixia Liu. A survey of visual analytics techniques for machine learning. *Computational Visual Media*, 7(1):3–36, 2021.



Dr. Quynh Quang Ngo is a post-doctoral fellow at VISUS, University of Stuttgart, Germany. He received his Dr. rer. nat. Degree from University of Münster, Germany, in 2020. His work focuses on conducting research in the field of Visualization. His recent research interests include Dimensionality Reduction, Data Transformation, Graph Layout, and applying Machine Learning to Visualization.

Address: Universität Stuttgart, Visualization Research Center, D-70174 Stuttgart, E-Mail: Quynh.Ngo@visus.uni-stuttgart.de



Frederik L. Dennig received his Master of Science in Computer and Information Science at the University of Konstanz, Germany, in 2019. As a doctoral researcher, he is part of the Data Analysis and Visualization Research Group at the University of Konstanz. His main research interests are visual analytics, pattern detection, and subspace analysis. A specific focus of his research is on the quantification of quality and visual patterns for visualizations and incorporating implicit user feedback into the analysis workflow.

Address: Universität Konstanz, Department of Computer and Information Science, D-78464 Konstanz, E-Mail: frederik.dennig@uni-konstanz.de



Prof. Dr. Daniel A. Keim is head of the Data Analysis and Visualization Research Group in the Computer and Information Science Department at the University of Konstanz, Germany. He has been actively involved in data analysis and information visualization research for more than 25 years. His services to the research community include program chair of InfoVis 1999, InfoVis 2000, KDD 2002, VAST 2006, VAST 2019, and VMV 2022; general chair of InfoVis 2003; and associate editor of IEEE TVCG, IEEE TKDE, and Sage Information Visualization Journal. He received his Ph.D. degree in Computer Science from the University of Munich, Germany. Before joining the University of Konstanz, he was an associate professor at the University of Halle, Germany, and a Technology Consultant at AT&T Shannon Research Labs, NJ, USA.

Address: Universität Konstanz, Department of Computer and Information Science, D-78464 Konstanz, E-Mail: keim@uni-konstanz.de



Prof. Dr. Michael Sedlmair is a professor at the University of Stuttgart and leads the research group for Visualization and Virtual/Augmented Reality there. He received his Ph.D. degree in Computer Science from the University of Munich, Germany, in 2010. Further stops included the Jacobs University Bremen, University of Vienna, University of British Columbia in Vancouver, and the BMW Group Research and Technology, Munich. His research interests focus on visual and interactive machine learning, perceptual modeling for visualization,

immersive analytics and situated visualization, novel interaction technologies, as well as the methodological and theoretical foundations underlying them.

Address: Universität Stuttgart, Visualization Research Center, D-70174 Stuttgart, E-Mail: Michael.Sedlmair@visus.uni-stuttgart.de