Cluster-Vis: Visualizing Data Clustered in High-Dimensional Spaces

Joshua Kruzan* Luca Fonstad† Cole Foster‡ Alex Hruska§

Brown University

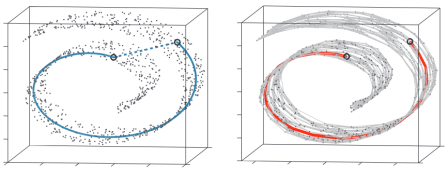


Figure 1: The classic Swiss Roll dataset. Data lies on a manifold of low intrinsic dimensionality compared to its ambient space.

Abstract

High-Dimensional clustering is a relatively uncommon clustering technique because, among other reason, it is difficult to visually analyze the resulting clusters. To aid in this visualization problem, we present Cluster-Vis, an interactive web application to support the visual analysis of clustering that has been performed in a high-dimensional space. To evaluate Cluster-Vis, we present a case study as well as a small user study. Our preliminary studies suggest that an interface with multiple plots and conveniently located documentation in Cluster-Vis increases user execution time and confidence for this analysis task, in comparison to an existing tool.

Keywords: High dimensional, clustering, visual analysis.

1 Introduction

A common practice when clustering is to first reduce the dimensionality of the dataset before generating clsuters. This technique is usually effective and increases the speed at which clusters are calculated. Additionally, by reducing the dataset to two or three dimensions, the clusters can be easily visually analyzed by plotting datapoints in this low-dimensional space and colored by cluster label. However, this introduces the problem of being difficult to visually analyze since there may be no meaningful ways to plot the data in a low-dimensional space.

As a motivating example for why we may want to cluster data in a high-dimensional space, we can consider the classic Swiss Roll dataset seen in Figure 1. In this three dimensional ambient space, the two illustrated points appear relatively close as measured by their Euclidean distance; however, it is clear that the distance along the manifold, shown on the right, is a much more representative distance metric and measures the two points as quite far apart. Note that if the dimensionality were to be reduced to a 2D space, this spiraling relationship containing this distance would be lost. By using clustering algorithms that are compatible with high-dimensional data, this is one type of relationship that can be preserved.

To aid in the visual analysis of clustering that has been performed with such techniques, we designed and implemented Cluster-Vis, a Plotly Dash web application written in Python. Cluster-Vis serves as a dashboard where users can upload clustering results and analyze them via six different interactive plots and figures. To illustrate a usage scenario, we contribute the case study that guided the design and development of Cluster-Vis as well as a small user study comparing Cluster-Vis to an alternative tool.

*e-mail: joshua_kruzan@brown.edu
†e-mail: luca_fonstad@brown.edu
‡e-mail: cole_foster@brown.edu
§e-mail: alex_hruska@brown.edu

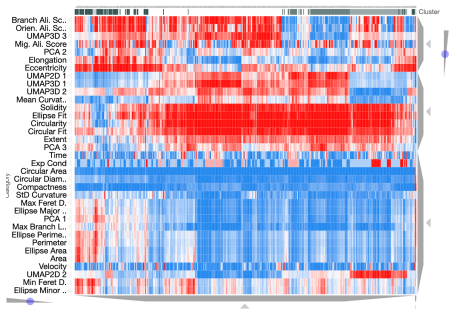


Figure 2: Heatmap from Clustergrammer

2 Related Work

Though none have been cemented as an industry-standard, Cluster-Vis is not the first tool designed to help with this task. Three comparable tools are ClustVis [6], Clustergrammer [2], and Clusterplot [5]. The main feature of both of ClustVis and Clustergrammer is a heatmap which shows normalized feature values for each entry in the dataset. An example heatmap from Clustergrammer can be seen in Figure 2.

Cluster-Vis also features a heatmap that shows a similar representation of the data. Cluster-Vis differs in that it offers multiple other plots and figures. This layout allow users to make observations using a plot of their choosing and verify their observations using other features.

Other related tools exist but are less comparable for reasons including that they don't support high-dimensional clustering [1], are purposed for comparing clustering results [4], or only support time series data [8, 7].

3 Cluster-Vis

Cluster-Vis was implemented in Python using the Plotly Dash framework. Aiming to present users with multiple options for analyzing their data, the main layout consists of a list of plots that users can toggle on and off to render them on the page. The list is broken up into plots that are more helpful for analyzing and exploring the dataset, and others that are more useful for clustering analysis. In addition to the feature expression heatmap present in comparable tools, Cluster-Vis includes a Parallel Coordinates plot (Figure 3) , a scatter plot matrix and two Embedding Plots.

To upload different results, users can choose to place a CSV in the local directory where Cluster-Vis expects data to be stored, or CSV's can be directly uploaded via the Cluster-Vis interface.

4 Evaluation & Results

To evaluate Cluster-Vis, we first contribute a case study to illustrate a usage scenario and then present quantitative and qualitative results from a small user study.

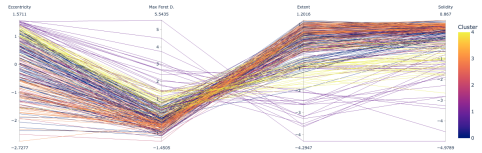


Figure 3: Parallel Coordinates plot from Cluster-Vis

4.1 Case Study

A collaboration with Alex Hrsuka and Cole Foster, from Brown University's Wong Lab and LEMS Lab respectively, guided the development of Cluster-Vis and also illustrates a usage scenario.

Foster was interested in developing a clustering algorithm that computed clusters given the dataset's Relative Neighborhood Graph. Hruska, from the mechanobiology field, had partnered with Foster and was looking to generate clusters for a dataset describing physical and motility data for 100,000 breast cancer cells in various experimental conditions.

When testing novel algorithms and obtaining a potentially promising result, Foster uploaded them to Cluster-Vis where Hruska was able to perform a more in-depth analysis. Using Cluster-Vis's heatmap, Hruska was able to verify that the clusters produced reasonable results; however, using the Parallel Coordinates Plot and Scatter Plot Matrix, Hruska realized the dataset was very imbalanced and would require filtering for desirable results.

Due to time constraints, Foster was unable to finish the development of Relative Neighbor Clustering; however, he was able to produce desirable clustering results for Hruska. To do so, he first reduced the dimensionality of the dataset to 5 dimensions and generated clusters using a common algorithm that supports high-dimensional clustering, DBSCAN. With these results, Hruska was able to use the Cluster-Vis to verify the clustering results made sense for grouping cells by their properties and the experimental conditions they were subjected to. Hruska will include these results in his forthcoming publication: "Programming cellular mechanophenotype and motility in composite 3D nanofiber hydrogels".

4.2 User Study

We conducted a user study using the same clustering results that were produced above. We surveyed 4 individuals, all of which are members of Brown University's Wong Lab and have a basic understanding of the dataset. After familiarizing themselves with both interfaces, we had users complete two similar tasks in both Cluster-Vis and Clustergrammer [2]. After each task, users would rate their perceived effort and demand. At the end of the survey, users provided qualitative responses about their experiences.

4.2.1 Quantitative Results

Varied slightly for each application, the tasks users were asked to complete were as follows:

1. Identify all clusters where you would find an oblong/circular shape cell.
2. Identify experimental conditions that occur in only one/two clusters.

Table 1 summarizes the users' performances on both tasks.

Loosely based on the NASA-TLX Usability Study [3], after completing each task in each application, users measured the usability of the application by rating their perceived Mental Effort, Physical Demand, Temporal Demand, Overall Performance, Overall Effort, and Frustration Level. Results across each task were similar so we present the results averaged by application in Figure 4.

Table 1: User Study Task Performance

	Cluster-Vis		Clustergrammer	
Task	Accuracy	Time (minutes)	Accuracy	Time (minutes)
Cell Shape ID	0.75	2.10	0.75	3.10
Experimental Condition ID	0.5	3.00	0.5	4.37

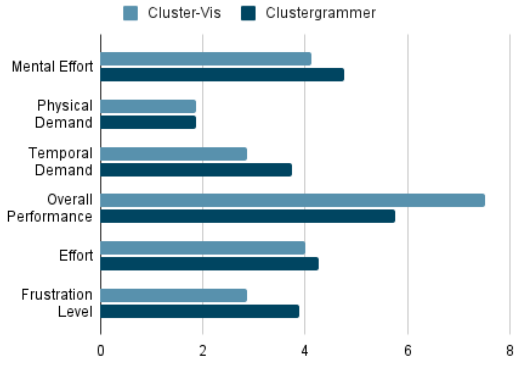


Figure 4: Perceived user demand and effort averaged over each task.

4.2.2 Qualitative Results

In the qualitative portions of the survey, users reported that their experience with Cluster-Vis was more enjoyable and that the multitude of plots made it easy to draw conclusions. Generally, they concluded that Clustergrammer was harder to learn but after some interaction, saw value in the advanced group-by features.

5 Discussion

As seen in Table 1, users were able to perform both tasks faster in Cluster-Vis than in Clustergrammer; however, with equal accuracy. We believe the low accuracy in the first task is due to a misunderstanding of the question. The task asked users to write all valid clusters; however, two of the users only noted 1. For the second task, in both applications, half of the users reported a reasonable, though not the best, answer.

From Figure 4, the results show that users perceived their effort and demand to be slightly less using Cluster-Vis than while using Clustergrammer. Although it is promising that Cluster-Vis was rated as equally or more usable for every question, with only 4 participants, these marginal differences could very likely be due to randomness. The biggest improvement that Cluster-Vis offers is with Overall Performance. This difference shows that users were more confident in their answers using Cluster-Vis than with Clustergrammer. This suggests that by being able to verify observations with other plots, users are more confident in their interpretations.

For future work, it is suggested that the user study is more widely distributed to achieve statistical significance. To achieve the suggested benefits of both Clustergrammer and Cluster-Vis, future work should also look into incorporating Clustergrammer's advanced heatmap features in Cluster-Vis.

6 Conclusion

Based on the evaluation results, it appears that with a multitude of plot options, users are able to more quickly, confidently, and less effortlessly analyze high-dimensional clustering results than with the heatmap in Clustergrammer.

ACKNOWLEDGEMENTS

The authors wish to thank, Alex Hruska and Cole Foster, for their collaboration throughout the semester. The authors wish to thank their professor, Dr. David Laidlaw, for his guiding support and feedback throughout the project.

REFERENCES

[1] Ç. Demiralp. Clustrophile: A tool for visual clustering analysis. *arXiv preprint arXiv:1710.02173*, 2017.

[2] N. F. Fernandez, G. W. Gundersen, A. Rahman, M. L. Grimes, K. Rikova, P. Hornbeck, and A. Ma'ayan. Clustergrammer, a web-based heatmap visualization and analysis tool for high-dimensional biological data. *Scientific data*, 4(1):1–12, 2017.

[3] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[4] S. L'Yi, B. Ko, D. Shin, Y.-J. Cho, J. Lee, B. Kim, and J. Seo. Xclusim: a visual analytics tool for interactively comparing multiple clustering results of bioinformatics data. *BMC bioinformatics*, 16(11):1–15, 2015.

[5] O. Malkai, M. Lu, and D. Cohen-Or. Clusterplot: High-dimensional cluster visualization. *arXiv preprint arXiv:2103.02992*, 2021.

[6] T. Metsalu and J. Vilo. Clustvis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic acids research*, 43(W1):W566–W570, 2015.

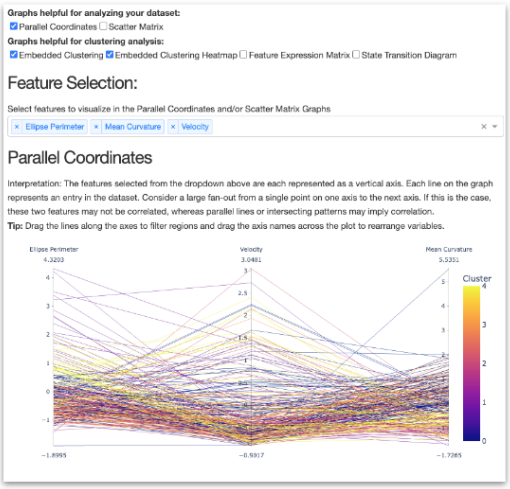
[7] G. S. Michaels, D. B. Carr, M. Askenazi, S. Fuhrman, X. Wen, and R. Somogyi. Cluster analysis and data visualization of large-scale gene expression data. In *Pacific symposium on biocomputing*, volume 3, pages 42–53, 1998.

[8] J. J. Van Wijk and E. R. Van Selow. Cluster and calendar based visualization of time series data. In *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis' 99)*, pages 4–9. IEEE, 1999.

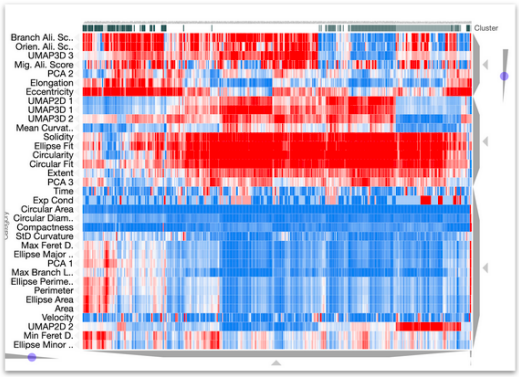
Cluster-Vis: Visualizing Data Clustered in High-Dimensional Spaces

Luca Fonstad* Author
Josh Kruzan† Co-Author
Cole Foster‡ Collaborator
Alex Hruska§ Collaborator

Brown University



= Figure 1: An example of the Cluster-Vis interface



= Figure 2: Clustergrammer's feature expression matrix

Abstract

In this paper we present Cluster-Vis, a novel visualization tool for visualizing data clustered in high-dimensional spaces. Cluster-Vis is a plotly-based dash web app written in python to provide multiple visualizations of high-dimensional clusterings for various scientific applications. We can see an example of the Cluster-Vis interface in Figure 1 In this paper we apply this tool to the following case study: a dataset of twenty-two thousand cells over various timesteps and in various experimental conditions, in order to cluster these cells into relevant groupings based on their expressions within these conditions. After testing various versions of our relevant clustering algorithm we were able to settle on one which provided the highest-quality data in terms of relevant clustering, through an iterative process. Ultimately, our tool was able to provide a high level of value to our collaborators and shows promise as a versatile tool that could be applied to a number of different types of high-dimensional datasets. We prove this quantitatively through a user study in which we test a number of experts with our tool against the most relevant currently existing tool in two different experiments.

1 Related Work

As high-dimensional visualizations are an open problem relevant to various disciplines, there are several comparable tools used for these types of visualizations. The most similar, and the one that we used as a point of comparison for our evaluation is Clustergrammer, as presented in [1], a tool for visualizing high-dimensional data through clustering and interactive visualizations. Clustergrammer offers several potential shortcomings as compared to Cluster-Vis: it only presents one type of visualization, while Cluster-Vis offers several, the feature expression matrix. We can see this feature expression matrix represented in Figure 2 Another comparable tool is ClustVis [2], a tool for visualizing high-dimensional data using PCA clustering. While this tool seems somewhat similar to ours we were not able to get it working, and additionally it only supports PCA clustering rather than custom clustering methods.

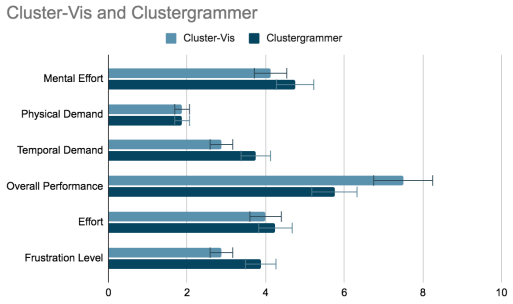
2 Methodology

As mentioned previously, we compared against Clustergrammer [1]. Unfortunately, it was unable to handle our entire dataset—leading us to sample only 500 cells out of twenty-two thousand. However, the tool offers us a chance to investigate whether the number of visualizations and cells represented truly aids users in identifying relevant information about the dataset or not, a useful comparison. In our user study we asked users two different questions—to identify a cluster where they would be most likely to see an oblong/circular cell shape, and to identify an experimental condition that only occurs in one or two cell clusters. We chose both of these tasks because they are both fairly opposite sides of the same coin: attempting to identify the different experimental conditions that define different clusters, and vice versa.

3 Results

The survey results turned out to be nearly identical across both questions, so we average the results across both questions in Figure 3. Here we can see that completely consistently, and promisingly, Cluster-Vis ranks as less effort in all categories. While many of these results are not statistically significant, due to the small sample-size, we focus particularly on the results showing the increased overall performance, which shows more promise given the

*e-mail: luca_fonstad@brown.edu
†e-mail: joshua_kruzan@brown.edu
‡e-mail: cole_foster@brown.edu
§e-mail: alex_hruska@brown.edu

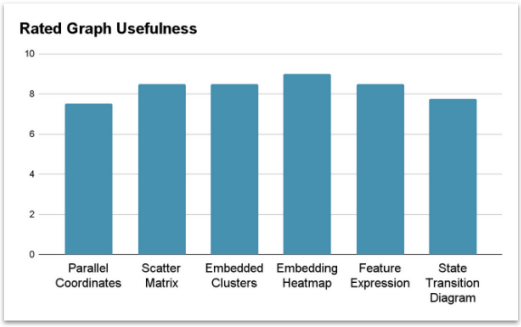


= Figure 3: The average results of our user study with both questions

	Cluster-Vis		Clustergrammer	
Task	Accuracy	Time (minutes)	Accuracy	Time (minutes)
Cell Shape ID	0.75	2.10	0.75	3.10
Experimental Condition ID	0.5	3.00	0.5	4.37

= Figure 4: The time taken by on average for every task in our user survey

extent to which Cluster-Vis was more highly rated than the competition, though ultimately a larger survey would be needed to make conclusive claims about these findings. We also found that users were able to complete tasks significantly faster using Cluster-Vis, as illustrated in Figure 4. We can also see in Figure 4 that users had a somewhat lower accuracy than expected; we can likely attribute this to a miscommunication regarding survey instructions. However, as we see this lowered accuracy reflected consistently across all tests, we do not believe that this impacted our survey results in any meaningful way. Additionally, we asked respondents to tell us how useful they believed each individual graph was, with the question "Please rate the usefulness of the individual graphs and plots". The results of this survey are shown in Figure 5. Here we can see that the graph rated as the most useful was the embedding heatmap, while the graph rated as the least useful was the parallel coordinates graph. However, there was not a huge discrepancy between all the different graphs, as all were considered to be useful on average by all respondents. The feature expression matrix, which was the only graph shown by the Clustergrammer tool, was rated to be the second most useful, which shows the relative usefulness of Clustergrammer, as well as its shortcomings. We also had respondents rate



= Figure 5: The average usefulness of every graph, as rated qualitatively by survey respondents.

the performance of the two tools in a qualitative manner, in which they were able to share their specific feelings on different aspects of the tools. In this survey, respondents informed us that they consistently found Cluster-Vis to be more intuitive and easier to use than the relatively confusing Clustergrammer. They also found that the tips and guidance given for every graph in Cluster-Vis was very helpful as compared to the steeper learning curve and relatively low amounts of guidance provided by Clustergrammer. However, users did find that more advanced features with Clustergrammer's more advanced feature expression matrix were more useful when learned, after their steep learning curve was overcome. In the future it may be useful to incorporate some advanced features from Clustergrammer into Cluster-Vis, to act in concert with the larger suite of visualizations provided. Users did report that they liked being able to view data in different formats in Cluster-Vis. They also reported that the visualizations provided by Cluster-Vis would be more useful in showing to non-expert collaborators, or simply collaborators who are experts in their respective fields but may not be as familiar with computational visualizations of this type.

4 Conclusion

Ultimately we are able to draw a few conclusions from our experiments with Cluster-Vis: our case study involving our collaborators and their dataset of twenty-two thousands cells proved that our tool was capable of giving a thorough analysis of complex and high dimensional data; we were able to find a solution to the complex problem of clustering high-dimensional data through an iterative process, and the results provided by Cluster-Vis will be included in future publications. Additionally, our user study was able to show that the multiple graphs given by Cluster-Vis gave an increased understanding of the data and the different ways that experimental conditions were expressed within and affected clustering, and gave an increased confidence in clustering analytics tasks. We were also able to show that the conveniently located tips and simpler plots were able to shorten the execution time and allowed users to more efficiently complete the various analytics tasks which we assigned them. The results of this study leave several open questions: firstly, an expanded survey that includes other types of applications that were not represented, and other tools to compare to more relevant to those types of tasks, such as Clustvis, would help provide insight into other aspects of our tool. Additionally, it would also be useful to compare setup and time-to-learn between Cluster-Vis and other existing tools; to make things easier we took on the task personally of converting the data given to us by our collaborators to a type which could be accepted by Clustergrammer, which involved a non-trivial amount of conversion. While we thought that this offered a more fair comparison within the context of our user survey, it is very much a factor which would be relevant to any research group attempting to perform an analysis on their own data, and as such may be a relevant thing to consider in future surveys, both within the context of Cluster-Vis and Clustergrammer, as well as other relevant tools.

References

[1] N. F. Fernandez, G. W. Gundersen, A. Rahman, M. L. Grimes, K. Rikova, P. Hornbeck, and A. Ma'ayan. Clustergrammer, a web-based heatmap visualization and analysis tool for high-dimensional biological data. *Scientific Data*, 4(1), 2017.

[2] T. Metsalu and J. Vilo. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res*, 43(W1):W566–570, Jul 2015.

NeuroViz: A Brain and NODDI Metric Visualization Tool

Naomi Lee* Shashidhar Pai† Ryan Cabeen ‡

Brown University, University of Southern California

ABSTRACT

In this paper, we present a novel method called NeuroViz, a web application for visualizing diffusion-weighted magnetic resonance imaging (MRI) data with derived NODDI metrics reflecting gray matter microstructure. This software was developed to provide researchers and clinicians a user-friendly platform for analyzing and interpreting data from a variety of neuroimaging modalities, such as functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI). In addition, we conduct a user study and report its findings validating the efficacy of NeuroViz in comparison to other open-source software and demonstrating that researchers find the tool to be a streamlined pipeline for diffusion MRI and related neurostatistics. The source code for our work can be found at `https://github.com/nlee100/nodejs-heroku-deploy.git`.

Keywords: neuroimaging, neuristatistics, visualization, diffusion-weighted magnetic resonance imaging, NODDI

1 INTRODUCTION

Neuroimaging data analysis and interpretation are essential to the study of the brain and the nervous system. These data, obtained through techniques such as fMRI and DTI, can provide valuable information about brain function. However, the analysis of neuroimaging data can be complex and may require specialized software and statistical skills.

Therefore, to address this challenge, we introduce NeuroViz, a project completed in collaboration with subject matter experts (SMEs) and researchers from the University of Southern California's (USC) Laboratory of Neuro Imaging (LONI) and Brown University's Carney Institute for Brain Science. We design and deploy a web-based application for the 3D visualization and analysis of neuroimaging data, enabling researchers to quickly and easily visualize their patient's whole brain data and view neuro-statistics mapped onto the brain's surface. The application's intuitive interface and suite of powerful tools, including multiple options to display NODDI metric information, enables accessibility to researchers without a programming background. We later discuss the potential impact of NeuroViz on the field of neuroscience, outline the design, and provide examples of its utility for researchers and clinicians working with neuroimaging data.

1.1 Background

Diffusion-weighted MRI generates contrast in MR images by sensitizing MRI measurements to the displacement patterns of water molecules undergoing diffusion. As the cellular structure of tissue directly affects the motion of water particles, diffusion MRI serves as a great tool to study tissue microstructure and aid applications such as tumor characterization or cerebra ischemia [1, 2]. Our interest in the cerebral cortex, the outermost layer of the brain primarily made of grey matter, stems from its association with humans' highest mental capabilities (e.g., movement, language and sensory information processing, attention) [3].

We seek to further the visualization of neurite orientation dispersion and density imaging (NODDI) [4], a diffusion MRI technique that splits the signal arising from three different tissue compartments—intra-neurite water, extra-neurite water, and free water compartments—to estimate neurite microstructure and associations with clinical outcomes. NODDI produces multiple metrics, such as neurite density index (NDI), isotropic volume fraction (IsoVF), and orientation dispersion index (ODI). Using the NODDI estimations of the density and fanning of neurites and the partial volume contamination from cerebrospinal fluid (CSF), we can better capture microstructural tissue abnormalities and clinically relevant phenomena related effects of age and disease.

2 RELATED WORK

With recent computational advancements and their ready applications in the biomedical imaging space, several research efforts have focused on developing tooling to advance the visualization of neuroimaging data. Featuring excellent ease of use and visualization graphics, these tools provide researchers with insights into brain function and disease by facilitating the interpretation of complex volumetric and surface datasets.

Principal among these tools is the Quantitative Imaging Toolkit (QIT) [5] designed by our collaborator, Dr. Ryan Cabeen, to provide researchers the ability to perform 3D visualization and data exploration of neuroimaging datasets. Our project strives to build upon Dr. Cabeen's innovative tool by interpolating QIT's 3D visualization strategy with NODDI metric mapping and an intuitive UI that does not require programming skills, tackling QIT's limitations (i.e., the requirement of technological capabilities of their researchers, the lack of statistical overlays to display NODDI measurements). A secondary leading software to note is BrainBrowser [6], an open-source JavaScript library furnishing web-based 3D visualization tooling based on WebGL. We aim to create an application similarly web-based and using WebGL or similar APIs (e.g., XTK [7], pycortex [8], threeBrain [9]) and address BrainBrowser's lack of generalizability and flexibility for users to adjust the software to their own research needs. Both QIT and BrainBrowser most greatly inspired our tool's design and function and ultimately became NeuroViz's competition in a portion of our user survey.

2.1 Goals

Therefore, the objectives of our web-based tool, NeuroViz, are to construct a streamlined tool to aid neuroscience visualization research at USC and beyond and to enhance understanding of grey matter discrepancies related to different risk factors (e.g., age, disease), accelerating the potential discovery of functional correlations between cortical microstructure variations or identifying biomarkers in regions of interest (ROIs). The major limitations of previous related works include a lack of accessibility to researchers without a programming background and a focus on either improving statistics or visualization, yet not both. Furthermore, no comprehensive data pipeline has previously been engineered to specifically study ODI with a combination of neurostatistics and analytics, and most

*e-mail: naomi_lee1@cs.brown.edu

†e-mail: shashidhar_pai@cs.brown.edu

‡email: ryan.cabeen@loni.usc.edu

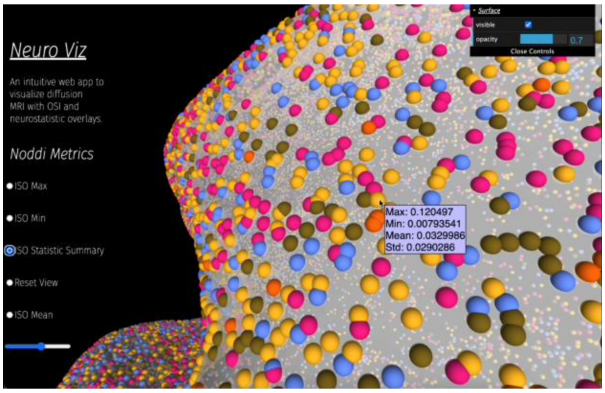


Figure 1: Demo of the NeuroViz.

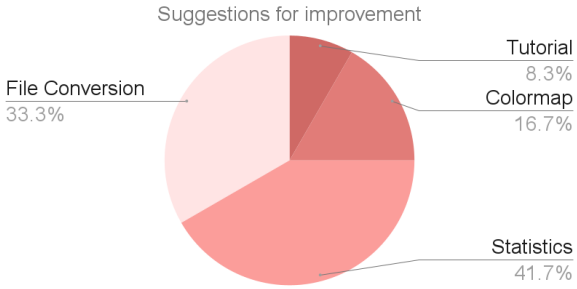


Figure 2: Charted user suggestions from an open-ended survey question.

tools are only commercially available. Those nonproprietary are not compatible with nor adjustable to the file formats used by our collaborators at LONI. Not only does NeuroViz deliver an intutive UI for viewing whole brains and a clear process to inspect and manipulate diffusion MRIs, but it offers an open-source integration of both visualization and statistics, specifically NODDI, into a singular pipeline and compatibility with the several types of the file formats provided by LONI.

3 Methodology

3.1 Data

The data used in this study is from the Human Connectome Project (HCP) and preprocessed by Dr. Ryan Cabeen into surface and NODDI metric data. It contains volumetric, surface, ROI, and NODDI metric information (i.e., intra-cellular volume fraction, intra-restricted volume fraction, isotropic diffusion component, NDI, ODI) of a single patient. While our demoed web application exhibits the ISO NODDI metric of our patient, this measurement can be easily substituted to display and utilize other metrics as well.

3.2 Implementation

Our tool was built using XTK, an open-source library and the first JavaScript framework to visualize and interact with medical imaging data using WebGL. Its impressive features were harnessed to create NeuroViz's dynamic 3D visualizations and statistical overlays upon the vertices of diffusion MRI surface data. Pycortex, a Python open-source library by the University of California, Berkeley's Gallant Lab was also considered but remained unimplemented due to its software inflexibility and documentation deficiency.

As seen in Figure 1, Neuroviz's features enable the uploading and operation of whole brain surface data with superimposed NODDI metrics, which are leveraged using a panel on the left-hand side. For example, one feature performs a k-means clustering of the data based on its NODDI metric values and colors the cluster's member vertices based on a pre-defined colormap. Users are subsequently able to hover over any vertex on the surface to view the vertex's cluster label and/or respective NODDI metrics (e.g., min, max, mean, and standard deviation of the chosen metric). Another feature enables users to accentuate vertices based on a specified threshold value. Further controls in the upper right-hand corner permit adjustment of the visibility and opacity of surface figure.

4 User Study

With a total of eight participants, consisting of six SMEs and two non-experts (e.g., peer computer scientists), we coordinated a user study to determine the effectiveness of NeuroViz. Employing a composition of a quantitative (e.g., provide a discrete rating for the ease of selecting a NODDI metric on a Likert scale) and qualitative (e.g., answer an open-ended question concerning the most useful research features of the tool) survey, we asked our participants to complete a variety of tasks in NeuroViz and provide feedback through the aforementioned survey. The results of this user-survey feedback were utilized to condition and fine-tune the tool.

4.1 Results

All participants in our user study communicated and evidenced great appreciation for our development of NeuroViz. When compared with the competition, QIT and BrainBrowser, our tool was preferred by our participants due to a variety of factors, such as cost-effectiveness, ease of use, web UI, and the interposed visualization of neuroimaging and neurostatistics. Similar feedback was extremely helpful, and at our participants' suggestion that was expressed $\approx 42\%$ in the evaluation survey's responses, we implemented a hovering feature for statistics (Figure 2). Other suggestions, such as adding a tutorial for new users, have become our subsequent courses of action to increase utility for our research tool.

Based on user answers from our evaluation survey, our leading priorities are to make this tool more convenient for researchers include adding support for converting file formats (e.g., .vtk to .stl, binary to ASCII) and the ability to create colormaps for surface data or interposed statistics. Given our struggle to access and utilize reliable open-source software, we are open to collaboration or new directions for this tool in the application of neuroimaging, neurostatistics, and beyond.

5 Conclusion

The NeuroViz web application is found to be a useful resource for presenting and analyzing neuroimaging data. It empowers researchers to more effortlessly and speedily examine and interpret complex brain data in heterogeneous formats (e.g., volume, surface, ROI), assisting a deeper understanding of brain function and disease. Its user-friendly interface and diverse visualization options make it suitable for researchers regardless of experience and programming ability, and our user study has shown NeuroViz to greatly supplement the faculty to study the brain and advance of neuroscience research.

Acknowledgements

We wish to thank our collaborators Dr. Ryan Cabeen and Dr. Kirsten Lynch at the USC's LONI for their guidance and feedback, without which this project would not be possible. We would also like to thank our collaborators at Brown Iniversity's Carney Institute for Brain Science for their mentorship and evaluations. Finally, we wish to thank our professor, Dr. David Laidlaw, and fellow classmates for their support and constructive criticism over the course of the semester.

REFERENCES

[1] Ryan P. Cabeen, Arthur W. Toga, and John M. Allman. Frontoinsular cortical microstructure is linked to life satisfaction in young adulthood. *Brain Imaging and Behavior*, 15(6):2775–2789, 2021.

[2] Ryan P Cabeen, Arthur W Toga, and John M Allman. Mapping frontoinsular cortex from diffusion microstructure. *Cereb. Cortex*, June 2022.

[3] Hikaru Fukutomi, Matthew F. Glasser, Hui Zhang, Joonas A. Autio, Timothy S. Coalson, Tomohisa Okada, Kaori Togashi, David C. Van Essen, and Takuya Hayashi. Neurite imaging reveals microstructural variations in human cerebral cortical gray matter. *NeuroImage*, 182:488–499, 2018.

[4] Hui Zhang, Torben Schneider, Claudia A. Wheeler-Kingshott, and Daniel C. Alexander. Noddi: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61(4):1000–1016, 2012.

[5] Ryan P Cabeen, David H Laidlaw, and Arthur W Toga. Quantitative imaging toolkit: software for interactive 3d visualization, data exploration, and computational analysis of neuroimaging datasets. *ISMRM-ESMRMB Abstracts*, pages 12–14, 2018.

[6] Tarek Sherif, Nicolas Kassis, Marc-Ã‰tienne Rousseau, Reza Adalat, and Alan C. Evans. Brainbrowser: Distributed, web-based neurological data visualization. *Frontiers in Neuroinformatics*, 8, 2015.

[7] Haehn Daniel, Rannou Nicolas, Ahtam Banu, Grant Ellen, and Pienaar Rudolph. Neuroimaging in the browser using the x toolkit. *Frontiers in Neuroinformatics*, 8, 2014.

[8] James S. Gao, Alexander G. Huth, Mark D. Lescroart, and Jack L. Gallant. Pycortex: An interactive surface visualizer for fmri. *Frontiers in Neuroinformatics*, 9, 2015.

[9] John F. Magnotti, Zhengjia Wang, and Michael S. Beauchamp. Rave: Comprehensive open-source software for reproducible analysis and visualization of intracranial eeg data. *NeuroImage*, 223:117341, 2020.

Automated Visual Weighting for Multidimensional Time Series Data: Applications to E-Nose Sensors

Patrick Maynard *

Brown University

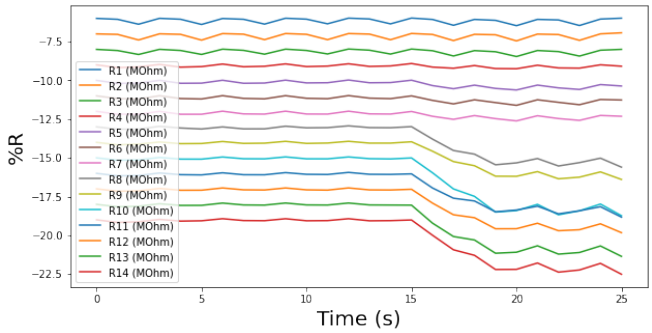


Figure 1: Stacked timeseries representation commonly found in chemical sensor literature.

Abstract

Multi-Dimensional timeseries data is abundant in scientific applications. However, current visualization techniques struggle to capture global comparisons, while also maintaining local resolution at scale. This work develops a prototype visualization platform that aims to resolve this trade-off. Using a new algorithm auto-α, the system represses redundant information by assigning fixed budgets for visual weighting. The method is applied to the Electronic Nose (E-Nose) domain and shows improvement over previous domain visualizations in an encouraging pilot study.

Keywords: Multi-Dimensional timeseries, visualization, resolution, auto-α.

1 Introduction and Related Work

An Electronic Nose (E-Nose) consists of an array of chemical sensors that respond with changes in resistance when in the presence of different chemical compounds. Each sensor in the array is designed to respond with different levels of sensitivities, dependent on the species of gas encountered. This produces a multi- dimensional time series that can be used to detect the type of gas species. These sensors provide a wealth of information, however current analysis is not leveraging the use of advanced visualization techniques, resulting in inefficiencies of analysis.

Most E-Nose visualizations currently use a simple stacked view of the time series with an offset (Figure 1), resulting in inefficiencies at scale [4]. This allows the user to compare changes between sensors with respect to the time axis. However, this format significantly limits the capacity of the user to perceive valuable local information. Stacking sensor data requires the scale relative to the screen size to be greatly reduced. Humans have perceptual limitations in perceiving multi-dimensional time series data in this format, these issues cannot be resolved by increasing screen resolution in the high-dimensional case [5]. Additionally, some compounds produce similar response to the same gas species, figure 1 shows sensors R1-R3 and R12-R14 all produce similar response.

Current multi-dimensional time series visualizations are centered around creating alternate representations of the original data. Li et al. applied Hidden Markov Models (HMMs) to transform multi-dimensional time series into uni-variate time series [2]. This reduces the complexity of the visualization, but removes the ability to compare inter-sensor differences, critical in E-Nose applications. Tsmap3d uses Mulit-Dimension Scaling (MDS) and Principle Component Analysis (PCA) to create an efficient projected space [1]. This provides strong visualization ability of the general feature space, but would have limited capacity to resolve trends for individual sensors. Other works like Mtsad have explored the integration of radar and box plots as alternative means for representing high-dimensional time series data [3]. These methods provide interesting insight into anomalies, but inhibits the ability to compare nuances of individual time series. This work demonstrates an alternative method that can eliminate redundant information and maintain the ability to make local comparisons, not solved by previous methods.

2 Methodology

To improve on current methods, this work provides two views of of the original data: (1) a local view of the timeseries, and (2) a global view of projected similarity. To resolve clutter in (1) this method assigns a fixed budget of visual weighting to a selected groups of sensors. The sensors are grouped together by extracting features, and applying K-Means clustering to group sensors exhibiting similar behavior. This work finds that coeficients from the Discrete Wavelet Transformation (DWT) and AUC provide a descriptive basis for feature spaces (see apendix for details), though we encourage users to implement their own features to match the intended application.

After feature extraction and clustering, this method employs auto-α, a novel algorithm for visual weighting:

1. Extract features from each time series
2. Apply K-means clustering
3. Loop through K clusters
   (a) Find time series mean of cluster
   (b) Loop through n time series
      i. Calculate MSE w.r.t mean
   (c) Normalize all MSE values on (0,1) to determine alpha parameter
4. Project features using PCA to create global view

By restricting a fixed transparency budget for each cluster, this method ensures that redundant information will not be given too much weight in the resulting visualization.

*e-mail: patrick_maynard@brown.edu

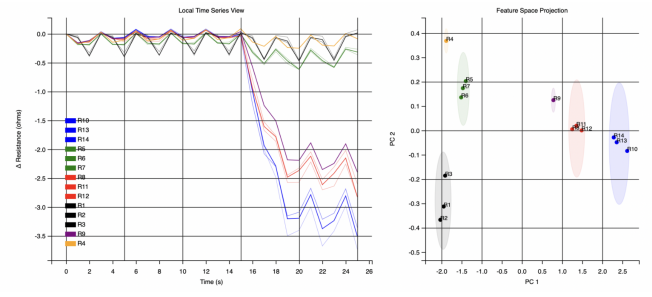


Figure 2: Previous dataset visualized with this work. Transparency of local view sensors is weighted using auto-α.

Figure 3: Comparison of performance for this work vs. stacked timeseries representation. Mean accuracy is reported as bar height with standard error of the mean in black.

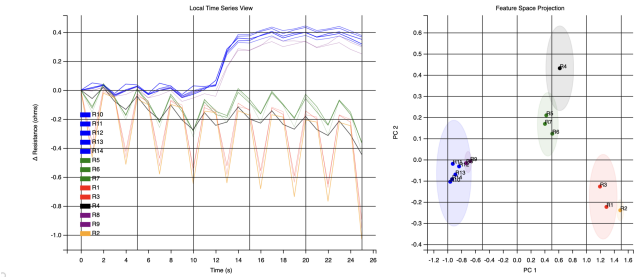


Figure 4: Question 3 contains sensors with very faint response such as those appearing in the green cluster.

3 Results

Figure 2 shows the results of the system using the same data as figure 1. In this case, the previously responding sensors are clearly visible, as well as sensors R5-R7 (green), which were previously difficult to resolve.

3.1 User Study

To analyze the performance of this system on E-Nose data, a pilot user study was created. The study asks users to determine which sensors show response to a given gas exposure. The users were given 6 examples visualized with this work and the stacked timeseries representation shown in figure 1. Latin Squares Sampling was used to determine the order in which the examples were viewed. A time limit of one minute was set to fix the analysis to accuracy. In order to see if this system would enable non-expert users to perform analysis on this data, the user study was broken down into expert ($N_e = 3$) vs. non-expert ($N_{ne} = 7$) users.

3.1.1 Overall Results

Overall, the results of this pilot study are encouraging. We find moderate statistical significance ($p \approx 0.01$) for both the expert and non-expert users when comparing the stacked timeseries view to this work (see appendix for details). While this work does significantly improve the ability of non-expert users to perform the task, the users were still not able to match expert performance on the stacked timeseries views. It is likely that with a larger sample size and more difficult examples for the experts, the results will converge with lower p-values. Note that the smaller standard error of the mean present in this work (figure 3) could be due to users being more likely to select the entirety of a cluster when selecting answers.

3.1.2 Results on Individual Questions

Looking at individual user accuracy for each question, we find that in some cases the task might be too easy for expert users (see figure 5 appendix: Q1, Q2, Q6). For some examples there seems to be a gradient showing better performance with this method. One instructive example where expert users showed better performance from the use of this work was question 3 (figure 4). In this case there are several sensors with a notably small response in the green cluster. Checking the individual responses, it is clear that experts in this category missed the responding sensors when analyzing the stacked view. This highlights the importance of scale when comparing multiple sensors for response and provides evidence for the strength of this method.

3.1.3 Global Feature View

When asked, "Did you find the global feature view helpful?" all experts agreed that this was useful when trying to perform the task (see appendix figure 6). Only one non-expert users claimed this had been helpful. The lack of positive response from non-expert users can likely be attributed to inadequate explanation being provided in the task description or lack of prior experience with analysing projected spaces.

4 Discussion and Conclusions

While the pilot study shows promise, there is still a lot left to explore to show the potential value of this system. Upon reflection, we found that a timed accuracy-based task may not serve as the best metric. A more extensive study may benefit from asking users more subjective questions about the system. At application time, this data is likely to be explored with no fixed length of time, asking user what insights they were able to draw may demonstrate more favorable results.

In addition to testing this system on publicly available datasets, this work was tested on private NASA datasets used for the ongoing development of a new E-Nose device. Dr. Sultana found the clustering particularly motivating for this use case. When developing the array of sensors, it is important to understand if certain manufacturing differences result in specific behavioral similarities. By clustering these sensors together, this can be observed w.r.t specific target gases.

We have demonstrated a preliminary improvement on current E-Nose visualization techniques and promising potential for the auto-α method. As a baseline, these results show that this system may improve analysis on E-Nose data that could impact a wide variety of scientific domains. Going forward, we plan to further develop the functionality of this visualization platform and package this for use by the team. We believe that continued feedback from high impact usage would provide critical information needed to release this to a wider audience and justify the strength of auto-α further.

REFERENCES

[1] S. Kamburugamuve, P. Wickramasinghe, S. Ekanayake, C. Wimalasena, M. Pathirage, and G. Fox. Tsmap3d: Browser visualization of high dimensional time series data. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 3583–3592. IEEE, 2016.

[2] J. Li, W. Pedrycz, and I. Jamal. Multivariate time series anomaly detection: A framework of hidden markov models. *Applied Soft Computing*, 60:229–240, 2017.

[3] V. Pham, N. Nguyen, J. Li, J. Hass, Y. Chen, and T. Dang. Mtsad: Multivariate time series abnormality detection and visualization. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 3267–3276. IEEE, 2019.

[4] M. Yakob, D. Mustika, R. N. Ida, and A. P. Rachmad. Design of e-nose as an instrument identification of diseases through the respiratory tract. In *Journal of Physics: Conference Series*, volume 1428, page 012062. IOP Publishing, 2020.

[5] B. Yost and C. North. The perceptual scalability of visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):837–844, 2006.

5 APENDIX

5.1 Individual User Accuracy

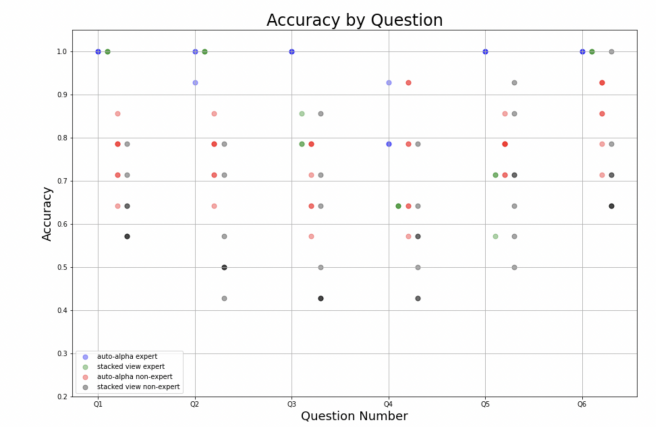


Figure 5: Accuracy by question with each point corresponding to specific user performance.

5.2 Global View Preference

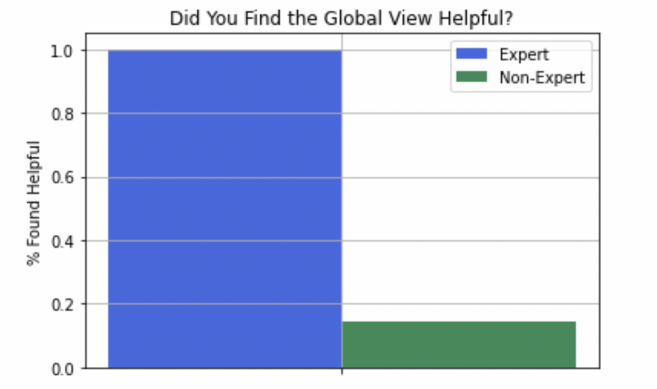


Figure 6: Survey results from asking users if they found the global view helpful.

5.3 Discussion on Statistical Significance

In order to resolve statistical significance, this work treats each user example as an independent example and measures the accuracy obtained by that user. We use a t-test of significance to obtain p-values and compare each of the categories, with any $p \leq 0.001$ highlighted in red.

	This Work (expert)	Stacked Timeseries (expert)	This Work (non-expert)	Stacked Timeseries (non-expert)
This Work (expert)	1	0.009589	4.27226e-11	2.37127e-13
Stacked Timeseries (expert)	0.009589	1	0.012772	2.103512e-06
This Work (non-expert)	4.27226e-11	0.012772	1	1.951218e-06
Stacked Timeseries (non-expert)	2.37127e-13	2.103512e-06	1.951218e-06	1

Auto-Diagnosis Platform For Dry Eye Disease

Shixuan Li* Yuxuan Zhao† Xiuli Wang‡

Brown University, CYL Hospital

ABSTRACT

In order to relieve the pressure of frequent visiting for Dry Eye Disease (DED) patients and the burden for doctors, we developed a web-based auto-diagnosis platform that enables remote consulting for patients and hospitals. The platform consists of two sections. One is a diagnosis model that takes in patients' self-taken eyelid images, segment out (deep learning model), analyzes the morphology information (vision computing) for each meibomian gland, and classifies the atrophy severity for each gland. The other is a web-based platform connected with hospitals' health record systems, allowing doctors to track patients' records and give proper treatment advice. Our model achieved significantly higher DED diagnosis accuracy, and our platform received very favorable feedback from patients and doctors.

Keywords: Dry Eye Disease, Auto Diagnosis, Remote Consulting

1 INTRODUCTION

Dry Eye Disease (DED) is one of the most widespread diseases in Ophthalmology, influencing 20% to 30% of the global population (around 66 million people in the U.S). However, DED is usually not a life-threatening disease since over 90% of DED patients only have minor symptoms, such as disability of tear secretion or pain in the eyeball. According to [CITE], 85% of the DED patients only need eye drops instead of surgeries or further inspections.

Although DED is not dangerous, it brings too many people to the hospital and significantly increases doctors' pressure and workload. According to CYL Hospital's records, DED-related patients take 10% to 15% of doctors' outpatient time. Thus, DED not only takes hospital too much time and labor but also only bring little profit —- most DED patients only need fundamental inspections, which on average cost $20 per person.

2 RELATED WORK

The relative size of areas of meibomian gland atrophy, or gland loss area, is an essential clinical measure for assessing meibomian gland dysfunction severity. Meiboscore, a metric first proposed by Arita et al. 2008 [1], which measures the ratio of meibomian glands' area and the total analysis area, is one of the most widely acknowledged evaluation standards. Currently, clinicians estimate the degree of meibomian gland atrophy subjectively by comparing the area of glandular loss with the total eyelid area [9] [8]. Although commonly used, the method only evaluates the overall severity of gland atrophy, not detailed individual meibomian gland morphological features. Recent studies have shown that morphological features of meibomian glands (such as length, curvature, or tortuosity, and local contrast [1]) may also be indicative of meibomian gland dysfunction severity and related to ocular surface disease [3] [6] [13].

Figure 1: Instance Segmentation & Morphology Analysis

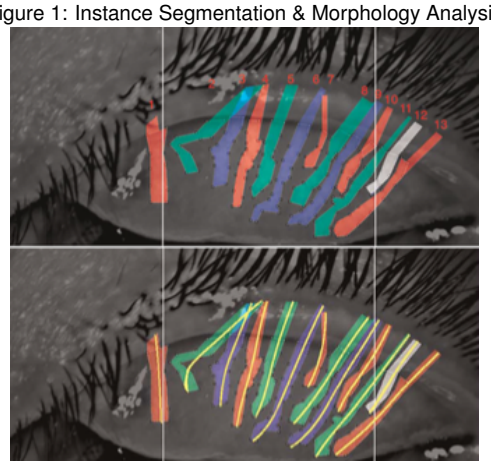


Table 1: Morphology Analysis Feature Extraction Example

Index	Contrast	Length	Width	Tortuosity	#Nodes	Direction °
1	12.1	4.27	0.608	15.9	2	81
2	11.8	9.74	0.3	49	5	42
3	20.5	6.24	0.427	25.4	2	77
4	17.6	6.97	0.378	20.9	3	89
5	18.4	6.71	0.397	22.1	2	82
...						
13	17.2	9.96	0.322	32.5	2	59

Because of meibomian glands' dense distribution and non-uniform shapes, traditional instance segmentation models perform poorly. Thus, on the one hand, previous studies use either image processing software, such as ImageJ [11] [2], or semantic segmentation models [15] [14] [16] to extract the semantic gland areas. On the other hand, some studies implement hypothesis-test-based atrophy severity prediction models by manually extracting individual gland's morphological information and designing new diagnosis metrics [12] [4]. However, the above studies are either unable to catch individual gland information or unable to automate the diagnosis process.

In this project, we not only designed an instance segmentation model (with a series of novel data processing procedures, network design, loss functions, and training details) that provide a satisfying accuracy but also proposes several morphological analysis metrics. As a result, we are able to automate the diagnosis process while capturing all morphological features.

3 METHODOLOGY

Our project comprises a vision computing solution for image-based auto diagnosis and a web application that serves doctors and patients.

*e-mail: shixuan_li@brown.edu

†e-mail: yuxuan_zhao@brown.edu

‡e-mail: wangxla@sina.com

[1]Local Contrast: Average gland region intensity normalized by its surrounding intensity.

Figure 2: Our Diagnosis Accuracy vs. Previous Studies

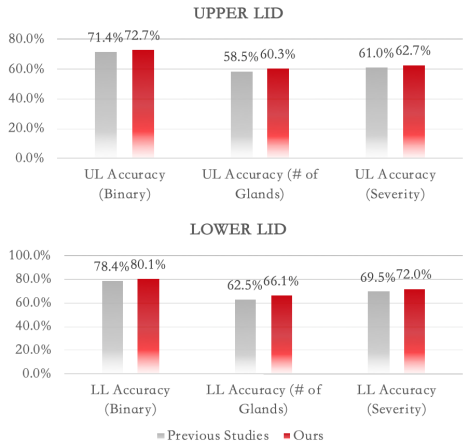
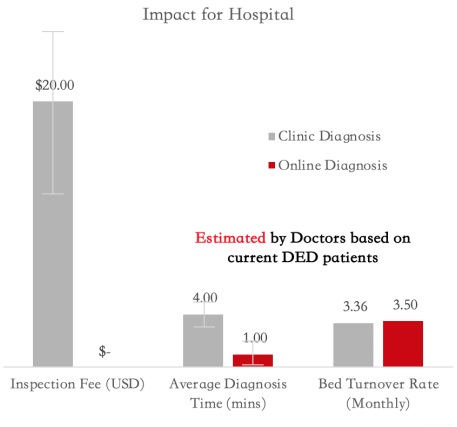


Figure 3: Time & Cost Saving For Patients, Effort Saving For Doctors



3.1 Auto-Diagnosis Model

We propose a three-step solution for the DED diagnosis. Firstly, based on the real-life meibomian gland image data from Dr.Wang from CYL Hospital, we trained an instance segmentation model [5] with many original adjustments. The model follows a two-step process. It firstly takes the input images (patient's self-taken eyelid images) and predicts the pixels of gland area. Then, we implemented a GAN [10] model to refine the raw output of segmented glands.

Secondly, we proposed a series of morphology analysis methods that takes the segmented glands as inputs and outputs several shape features for each gland, such as length, width, tortuosity, and curvature degree. This procedure also conforms to doctors' real-life diagnosis regulations and thus offers significantly helpful references that shorten doctors' diagnosis time.

Lastly, based on all the extracted morphology analyses, we built a statistical model to predict the atrophy of each gland. The advantage of a statistical model is that it can provide good reasoning for the diagnosis. For example, if the model thinks a gland has atrophied, we can tell from sensitivity tests and fitting parameters that it is judged by the gland's abnormal contrast and tortuosity.

3.2 Auto-Diagnosis Platform

We also build a web application to build the connection between the hospitals and patients. When patients feel uncomfortable, they can use a special lens (which costs about $8) to take some photos of their eyelids. The images will be uploaded and stored in the hospital's database. Notice that patients must register on the hospital's website and verify their personal information before using our web application since everything will be kept as a personal health record. Then, our auto-diagnosis model will predict and display all intermediate analysis processes to the doctors who double-check the diagnosis results. Doctors will validate all records and give treatment suggestions to patients, which will be displayed on the patient's panel and in their emails.

The platform was built upon the Django framework, with pure H5+CSS3 as the front end and MySQL as the database.

4 Contribution & Evaluation

Our project not only outperforms previous studies in diagnosis accuracy but also brings convenience and improved efficiency to patients and hospitals.

4.1 Technical Breakthrough

In this project, we built better models for segmentation and classification and proposed a series of new metrics that significantly helped the diagnosis performance. Firstly, our deep-learning-based instance segmentation[5] model achieved higher semantic mIOU. Secondly, because of the coverage of either their eyelashes or fingers, most patients' uploaded images do not expose the entire gland, which infers the direction glands are secreting the oil. Thus, we proposed a novel method using online PCA and modified Principle Curve [7] to reveal the entire gland. Lastly, as shown in Figure 2, our statistical model achieved higher accuracy in severity evaluation and gland identification.

4.2 User Study

We interviewed 26 professional ophthalmology doctors in CYL hospital and Hebei 2nd Hospital (1 department head, 3 attending physicians , 2 fellows, 4 chief residents, 9 residents and 7 interns) and 20 patients for their feedback.

All doctors expressed their concern about the large patient population for DED and looked forward to taking care of the minor-symptom patients online. After trying out our platform with some real-life cases, 100% of the professionals agree on the convenience of the workflow and think that the displayed morphology analysis information is beneficial. They also expect a much lower average diagnosis time for each patient and thus save time for surgeries and other training. Although remote diagnosis makes hospitals lose some inspection fees, the chief residents expect a higher general profit for hospitals since doctors have more time on urgent and complex cases, improving the monthly bed turnover rate[2].

Among the interviewed patients, three elder patients expressed concern with AI diagnosis. The rest 17 interviewees all think that the platform can be beneficial for the diagnosis of Dry Eye Disease, especially in saving travel and wait time. The function they love the most is the precise morphology analysis and diagnosis reason, as well as the record system that allows them to keep track of their health condition.

5 Conclusion

As society gets more fast-paced, there are growing calls for remote medical consulting. Our auto-diagnosis platform broke the barrier of automating meibomian gland segmentation and morphological feature extraction, achieving better atrophy evaluation accuracy and valuable open-box reasoning for DED diagnosis. The platform not only brings convenience to patients but also better running efficiency and profit for hospitals.

[2]Monthly Bed Turnover Rate: Average time a patient stay in ophthalmology beds is 12.84 days. The unit here is the # of patients get out of hospital beds / # of beds in ophthalmology within 1 month.

REFERENCES

[1] Reiko Arita et al. "Noncontact Infrared Meibography to Document Age-Related Changes of the Meibomian Glands in a Normal Population". In: *Ophthalmology* 115 (May 2008), pp. 911–5. DOI: 10.1016/j.ophtha.2007.06.031.

[2] Reiko Arita et al. "Objective image analysis of the meibomian gland area". In: *The British journal of ophthalmology* 98 (June 2013). DOI: 10.1136/bjophthalmol-2012-303014.

[3] Nien Chyong Jy et al. "Effects of age and dysfunction on human meibomian glands". In: *Arch Ophthalmol* 129 (2011), pp. 462–469. DOI: 10.1001/archophthalmol.

[4] Qi Dai et al. "A Novel Meibomian Gland Morphology Analytic System Based on a Convolutional Neural Network". In: *IEEE Access* PP (Feb. 2021), pp. 1–1. DOI: 10.1109/ACCESS.2021.3056234.

[5] Bert De Brabandere, Davy Neven, and Luc Van Gool. *Semantic Instance Segmentation with a Discriminative Loss Function*. 2017. DOI: 10.48550/ARXIV.1708.02551. URL: https://arxiv.org/abs/1708.02551.

[6] "Grading and baseline characteristics of meibomian glands in meibography images and their clinical associations in the Dry Eye Assessment and Management (DREAM) study". In: *The Ocular Surface* 17.3 (2019), pp. 491–501. ISSN: 1542-0124. DOI: https://doi.org/10.1016/j.jtos.2019.04.003.

[7] Trevor Hastie and Werner Stuetzle. "Principal Curves". In: *Journal of the American Statistical Association* (1989). DOI: 10.2307/2289936.

[8] Pult Heiko and Riede-Pulta Britta. "Non-contact meibography in diagnosis and treatment of non-obvious meibomian gland dysfunction". In: *Journal of Optometry* 5 (2012), pp. 2–5. DOI: 10.1016/j.optom.2012.02.003.

[9] Pult Heiko and Nichols Jason. "A review of meibography". In: *Optometry and Vision Science* 89 (2012), pp. 760–769. DOI: 10.1097/OPX.0b013e3182512ac1.

[10] Phillip Isola et al. *Image-to-Image Translation with Conditional Adversarial Networks*. 2016. DOI: 10.48550/ARXIV.1611.07004. URL: https://arxiv.org/abs/1611.07004.

[11] Xiaolei Lin et al. "Characterization of Meibomian Gland Atrophy and the Potential Risk Factors for Middle Aged to Elderly Patients With Cataracts". In: *Translational Vision Science Technology* 9 (June 2020), p. 48. DOI: 10.1167/tvst.9.7.48.

[12] Xiaoming Liu et al. "Scribble-Supervised Meibomian Glands Segmentation in Infrared Images". In: *ACM Transactions on Multimedia Computing, Communications, and Applications* 18 (Aug. 2022), pp. 1–23. DOI: 10.1145/3497747.

[13] Callinor S. "How to Treat Miscros Meibomian Gland Dysfunction". In: *Optician Online* (2020). URL: https://www.opticianonline.net/features/picturing-%20meibomian-gland-dysfunction.

[14] Ripon Saha et al. "AI-based automated Meibomian gland segmentation, classification and reflection correction in infrared Meibography". In: (June 2022).

[15] Ye-Ye Zhang et al. "Artificial Intelligence to Detect Meibomian Gland Dysfunction From in-vivo Laser Confocal Microscopy". In: *Frontiers in medicine* 8.10 (2021), pp. 774–334. DOI: 10.3389/fmed.2021.774344.

[16] Zuhui Zhang et al. "Meibomian Gland Density: An Effective Evaluation Index of Meibomian Gland Dysfunction Based on Deep Learning and Transfer Learning". In: *Journal of Clinical Medicine* 11 (Apr. 2022), p. 2396. DOI: 10.3390/jcm11092396.

ACKNOWLEDGEMENTS

The authors wish to thank, Jiayu Ren and Jian Hao, for their collaboration throughout the semester. The authors wish to thank their professor, Dr. David Laidlaw, for his guiding support and feedback throughout the project.

Application for Interactive Brain Visualization in Stroke Diagnosis

James Li * Ryan Cabeen †

Brown University

Abstract

We propose a novel method for interactive visualization of brain MRI that correlates 3D models to 2D cross-sectional images. We report findings by conducting user studies and a paired-sample t-test that assesses if our software is more efficient and user-friendly for lesion identification.

Keywords: interactive visualization, human-computer interaction, neuroimaging.

1 Introduction

MRI visualization is now a widely adopted tool and crucial for lesion identification in stroke diagnosis. However, neurologists and radiologists are used to viewing brain scans in 2D image stacks. Hence many of the benefits of 3D volumes have yet to be fully exploited.

This extended abstract presents a novel approach to interactive brain visualization–a prototype web application developed to take advantage of 3D models and to address the problems of information loss and navigation difficulty that come with conventional image stacks.

In addition, this extended abstract also presents and discusses the results of a user study on the features of the web application and its performance compared to conventional methods.

2 Background And Related Work

The benefit of visualizing MRI data for stroke registries over conventional alphanumeric data was discussed extensively in previous research. Image_QNA developed a software package that processes quantitative MRI data in 2D and performs experiments to assess its feasibility and utility [4]. In our research, we intend to carry out a similar development and experiment protocol to testify our hypothesis and address Image_QNA's limitation of lack of support for multi-dimension visualization.

More recently, the Stroke Preclinical Assessment Network(SPAN) tested an image-based stroke assessment tool incorporating machine learning techniques into an automated pipeline for stroke diagnosis [2]. Cabeen, one of the principal researchers at SPAN, also published a software package named Quantitative Imaging Toolkit (QIT) [1], that provides 3D visualization and computational analysis. Inspired by their work, this project attempts to build upon them by incorporating the 3D visualization component from QIT into the stroke assessment process following the protocol of SPAN's image-based stroke assessment tool, addressing QIT's limitation on lack of functional usage in the diagnosis process.

Furthermore, combining 2D and 3D imaging to facilitate data exploration has already been proven effective in other medical applications. Namely, an article from Neuroradiology conducted an in-depth user study that demonstrated the potential of correlating 2D and 3D images of the middle ear and adjacent structures for surgical planning [5].

*e-mail: james_li@brown.edu

†e-mail: cabeen@gmail.com

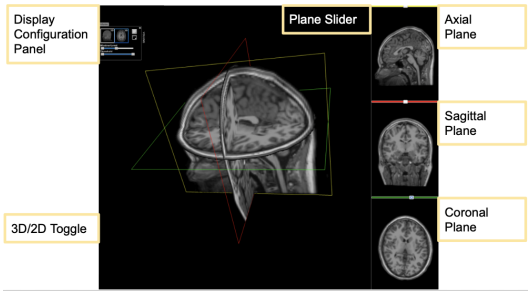


Figure 1: Screenshot of The Web Application

3 Methodology

Based on interviews with Ryan Cabeen, we carry out the following web application and user study.

3.1 Web Application

As indicated in Figure 1, the web application consists of several components.

First, in the center exists the 3D model of the cross-sections of a brain for navigation and location purposes.

Second, in the top left corner is a configuration panel to toggle 3D and 2D views and to adjust opacity and threshold settings.

Third, on the right of the interface are three plane viewers and sliders for axial, sagittal, and coronal views, respectively. When the user moves any slider, all three planes will render the image of the new location, and so will the cross-sections in the 3D model. Lastly, users can drop and render any MRI data in nifti file format.

3.2 User Study

The user study consists of 3 sections–instructions, tasks, and evaluation.

Firstly, participants follow question-guided instructions to learn the tasks and how to use the web application.

Secondly, participants are assigned two tasks to perform, one using conventional image stacks as the benchmark and the other using the web application. In each task, participants view data from 2 subjects and determine which subject is in its late stage of ischemic stroke. We then assess their performance by the accuracy of the results and the time taken to complete them.

Thirdly, participants report their preferences and rate the potential drawbacks and how valuable each web application feature is. Questions regarding participants' subjective opinions are answered in multiple choice or Likert Scale from 1 to 5.Thirdly, participants report their preferences and rate the value of each web application feature. Questions regarding participants' subjective opinions are answered in multiple choice or Likert Scale from 1 to 5.

One thing to bear in mind is that all example data are rodent brain scans from SPAN's database, and all four subjects each user views are unique, hence no redundant subjects.

4 Results

In the end, we collected results from 4 participants, two being professionals in neuroimaging or related research fields and the other two being medical students who reported having some prior experience with stroke and MRI scans. For better consistency, we inverted the time taken to complete the task, renamed it to "Efficiency", and normalized all metrics in Figure 2 to percentages.

To testify if the web application is more efficient than 2D image stacks for lesion identification, we compare the accuracy and efficiency as the objective metric. In addition, we used the Likert scale and multiple-choice answers to provide more subjective context for the performance of the two approaches.

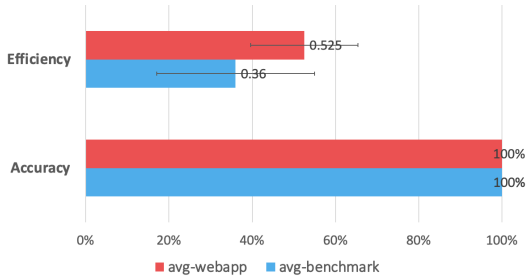


Figure 2: Performance Comparison between Benchmark and WebApp

As indicated in Figure 2, while accuracy remains intact, we observed a 16.3% improvement in efficiency when using the web application. However, due to the small sample size, the change was not statistically significant enough, as the p-value is about 0.159 from the paired sample t-test.

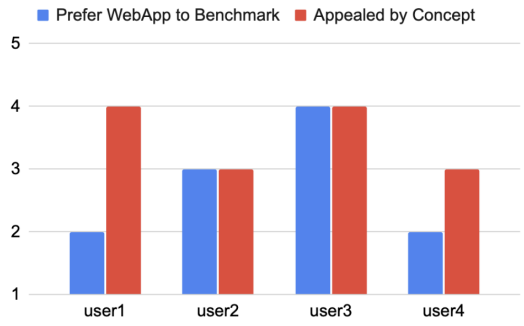


Figure 3: Preference for WebApp and Concept

For preferences over approaches and interests in concepts, we asked Likert scale questions. For preferences, we asked if participants preferred the web application over the benchmark, with 1 being benchmark infinitely better and 5 being web application infinitely better. The results were neutral to negative, hinting at a lack of functionality, usability, or both. Regarding interests, we asked participants to what extent the concept of combining 2D and 3D visualization in a single web app appealed to them, with one being not at all and five being highly interested. Most participants reported that they were appealed by combining 2D and 3D in an all-in-one application, as indicated in Figure 3.

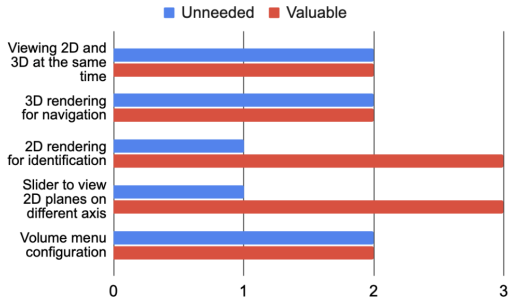


Figure 4: Feature Evaluation

To pinpoint the drawbacks and limitations of the web application, we asked which features were valuable in assisting them in completing their tasks in a multiple-choice format. While three features received ambiguous ratings, 2D rendering and plane sliders were favored by most participants, as shown in Figure 4.

5 Discussion

Due to the limited number of participants involved in this user study, there was no significant difference in efficiency between the web application and the benchmark. However, it does provide some insight that could shed light on future research.

Firstly, one participant with professional background reported that image resolution is much more critical than dimensionalities. Hence, rendering tools that can achieve higher resolution in less time could be valuable in developing brain visualization for stroke diagnosis.

Secondly, participants with less experience feel more positively towards the web application, hence experienced professionals may be biased against learning the new web application when they provide subjective feedback. This phenomenon should be verified and controlled when conducting future user studies.

Lastly, another participant with professional background also reported that using 3D navigation features, such as rotation, zooming, and sliding, is more challenging to complete on standard trackpads than using mice. This observation aligns with existing research on the effect of different input methods on performance, posture, and comfort [3].

6 Conclusion

In short, there was no significant difference in efficiency between using the web application and the benchmark.

However, 3D resembling characteristics, such as dimensionality association and plane sliders, are favored by participants, indicating the possibility for more in-depth research that provides statistical significance to our findings.

Acknowledgements

The author wishes to thank Ryan Cabeen for advising this research and providing the data needed to validate the web application.

References

[1] R. P. Cabeen, D. H. Laidlaw, and A. W. Toga. Quantitative imaging toolkit: software for interactive 3d visualization, data exploration, and computational analysis of neuroimaging datasets. *ISMRM-ESMRMB Abstracts*, pages 12–14, 2018.

[2] R. P. Cabeen, J. Mandeville, F. Hyder, B. G. Sanganahalli, D. R. Thedens, A. Arbab, S. Huang, A. Bibic, E. Tarakci, J. Mihailovic, et al. Image-based stroke assessment for multi-site preclinical evaluation of cerebroprotectants. *arXiv preprint arXiv:2203.05714*, 2022.

[3] G. Kar, A. Vu, B. Juliá Nehme, and A. Hedge. Effects of mouse, trackpad and 3d motion and gesture control on performance, posture, and comfort. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 59, pages 327–331. SAGE Publications Sage CA: Los Angeles, CA, 2015.

[4] D.-E. Kim, K.-J. Park, D. Schellingerhout, S.-W. Jeong, M.-G. Ji, W. J. Choi, Y.-O. Tak, G.-H. Kwan, E. A. Koh, S.-M. Noh, et al. A new image-based stroke registry containing quantitative magnetic resonance imaging data. *Cerebrovascular Diseases*, 32(6):567–576, 2011.

[5] T. Rodt, P. Ratiu, H. Becker, S. Bartling, D. Kacher, M. Anderson, F. A. Jolesz, and R. Kikinis. 3d visualisation of the middle ear and adjacent structures using reconstructed multi-slice ct datasets, correlating 3d images and virtual endoscopy to the 2d cross-sectional images. *Neuroradiology*, 44(9):783–790, 2002.

NeuroViz: A Brain and NODDI Metric Visualization Tool

Shashidhar Pai* Naomi Lee† Rayan Cabeen ‡

Brown University
University of Southern California

ABSTRACT

We present NeuroViz, a web-based application for the visualization and analysis of neuroimaging statistical data. This tool was developed to address the need for a user-friendly platform for researchers and clinicians to analyze and interpret data from functional magnetic resonance imaging (fMRI), diffusion tensor imaging (DTI), and other neuroimaging modalities. In this research paper, we discuss the design and implementation of NeuroViz and its potential impact on the field of neuroimaging. We also outline the evaluation of the tool and provide examples of its utility for researchers and clinicians working with neuroimaging data. The source code for our work can be found at `https://github.com/nlee100/nodejs-herokudeploy.git`

Keywords: Neuro-statistics, Neuroimaging visualization, NODDI metrics, Functional magnetic resonance imaging (fMRI),

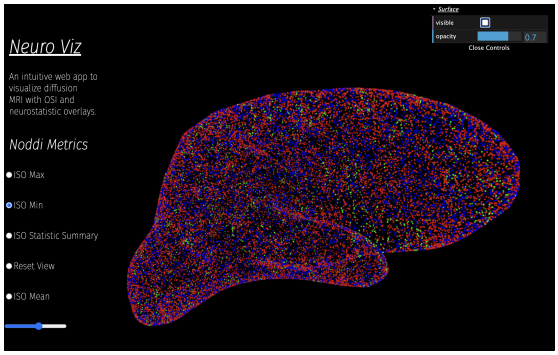


Figure 1: Overview of NeuroViz

1 INTRODUCTION

The analysis and interpretation of neuroimaging data are critical aspects of research in the field of neuroscience. Techniques such as functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI) provide valuable insights into brain function[1]. Diffusion-weighted MRI is a technique that captures the movement of water molecules to create contrast in MR images. It can be used for applications such as tumor characterization or analyzing the effects of cerebra ischemia[3, 2]. One specific technique called neurite orientation dispersion and density imaging (NODDI), a diffusion MRI technique that splits the water orientation signal arising from three different tissue compartments, namely intra-neurite water, extra-neurite water, and free water compartments, can be used to study the microstructure of neurites and their relationship to clinical outcomes. The analysis of this data can be complex and require specialized software and statistical expertise.

To address this challenge, we have developed NeuroViz(Fig 1) in collaboration with researchers from LONI – Laboratory of Neuro Imaging, University of Southern California, and Carney Institute for Brain Science, Brown University. We designed a web-based application for the visualization and analysis of neuroimaging data. Using NeuroViz, researchers can quickly and easily visualize their data and view neuro-statistics plotted onto the brain's surface in 3D. The application features an intuitive interface and a suite of powerful tools, including multiple options to visualize the Neurite Orientation Dispersion and Density Imaging(NODDI)[4] information, making it accessible to users without a programming background.

Finally, we discuss future directions for the development of NeuroViz and its potential to advance the field of neuroscience. In our evaluation, most researchers and neuroscientists reported that the streamlined pipeline of NeuroViz significantly reduced the effort required to analyze and visualize the neuroimaging data compared to traditional methods or other software like QIT and Brain Browser. Furthermore, the users identified the user-friendly interface, accessibility, and statistical visualization capabilities of the application as the key reasons for preferring it over other available options.

2 RELATED WORK

In recent years, several research efforts have focused on developing tools for visualizing and analyzing neuroimaging data. These tools have aimed to facilitate the interpretation of complex volumetric and surface data sets and provide researchers with insights into brain function and disease. One notable example is the Quantitative Imaging Toolkit (QIT)[1] developed by our collaborator, Rayan Cabeen, at the University of Southern California. It was specifically developed for tractography and microstructure analysis of diffusion magnetic resonance imaging datasets[5], but it has capabilities that are generally useful for other imaging modalities as well. Some of the other example software suites that provide similar utility are AFQ-Browser[6], and SlicerDMRI[8]. Another notable tool is the Brain Browser[7] which is an open-source JavaScript library exposing a set of web-based 3D visualization tools primarily targeting neuroimaging.

Our web app, NeuroViz, overcomes the major limitations in the above tools and techniques of not incorporating neuro-statistics in their tools and the accessibility to individuals without a programming background. The standard for visualizing volumes of any type of MRI data requires building upon WebGL, threeBrain, and other graphical libraries but does not include a comprehensive data pipeline for research ease or specifically studying ODI using our presented combination of imaging.

One challenge we encountered was that many of the available software tools were not compatible with the file formats used by our collaborator, LONI. Furthermore, most of these tools were proprietary and not easily customizable or open source. In contrast, the NeuroViz web application we developed is designed to support LONI's file formats while also being adaptable to other file formats and open source to allow for easy modification and updates.

*e-mail: shashidhar_pai@brown.edu
†e-mail: naomi_lee1@brown.edu
‡e-mail: ryan.cabeen@loni.usc.edu

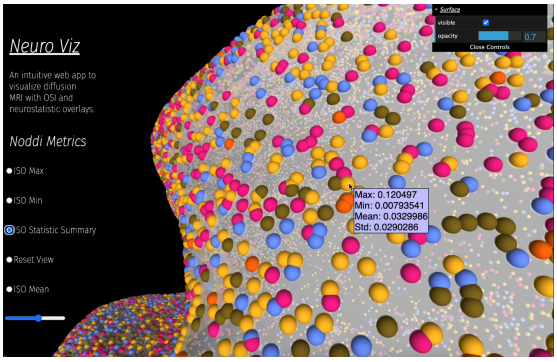


Figure 2: Based on User Survey: Hover over vertex for statistics

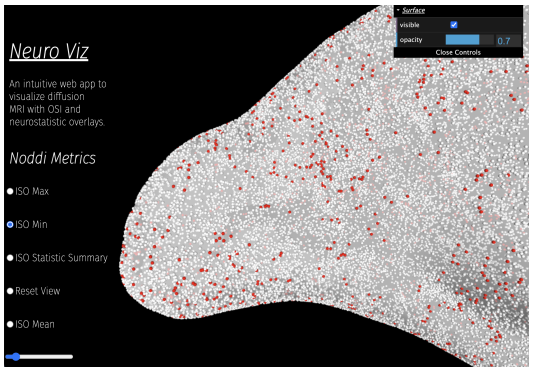


Figure 3: Functionality to select vertices based on interactive controls

3 Methodology

3.1 Data

Our team encountered a significant obstacle in the form of a lack of data and the steep curve to learn the domain expertise to understand the fMRI scan data from the Human Connectome Project. The data included both volumetric and surface information from a single subject, as well as various NODDI metric files, which capture various aspects of brain tissue structure and function. We developed a web application to visualize the ISO NODDI metric of the subject, but it is easily adapted to display any of the other metric values as well.

3.2 Web Application

The NeuroViz web application was built from the ground up using the XTK JavaScript library[4], which is an open-source library for rendering and manipulating volumetric data. The XTK library was used to create interactive 3D visualizations of neuroimaging functional MRI data and overlay the NODDI metrics on the surface/volumetric data.

The web application allows users to load and manipulate brain surface data overlayed with the NODDI ISO metric. It offers various options for analyzing and visualizing the data, including the ability to cluster the metric and view the resulting clusters on the brain surface, represented with different colors. Users can hover over any vertex on the brain surface to see its cluster identification or NODDI metric as seen in Fig 2. There is also a feature that allows users to highlight vertices based on a specified threshold as seen in Fig 3. The application also provides controls on the right side for adjusting the visibility and opacity of interpolated surface data.

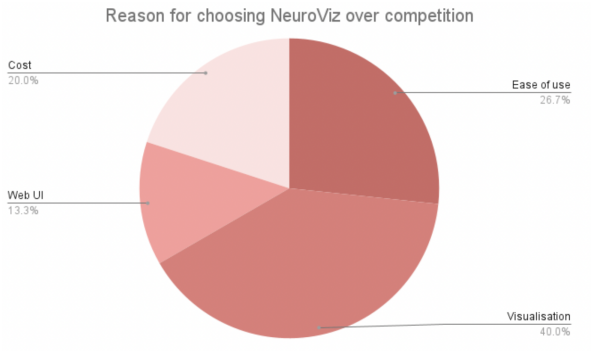


Figure 4: Charted user survey for selecting NeuroViz.

3.3 User Evaluation

To evaluate the effectiveness of the NeuroViz, we conducted a combination of observational studies and a qualitative survey. We recruited eight participants, including six experts and two non-experts, to complete a range of tasks using the NeuroViz web application. After completing the collaborative tasks, the participants completed a questionnaire in which they provided feedback on some open-ended questions and rated various statements about the software on a Likert scale. The results of this user-survey feedback were used to refine and improve the application.

4 Results and Discussion

All the users we surveyed expressed immense appreciation for the web application. We had very positive results from the user survey, with all the users choosing our NeuroViz tool over QIT and Brain Browser. The reasons stated by the users for choosing our tool were mainly the power of visualizing with neuro-stats, ease of use of the tool, cost-effectiveness, and the web UI, as seen in Fig 4. We also received excellent feedback for the provision of selecting the NODDI metric within the web application. From the survey, we received one user who suggested the color map for visualizing the metric provided fair usefulness. In contrast, the others rated the ability to visualize the metric as a good score. One possible future functionality that be added is the provision to select a color-scale for the a particular metric.

5 Conclusion

In conclusion, the NeuroViz web application has proven to be a valuable tool for analyzing and visualizing neuroimaging data. It allows researchers in the field of neuroscience to explore and interpret complex volumetric and surface brain data, enabling them to gain a deeper understanding of brain function and disease easily and quickly. The intuitive interface and wide range of visualization options make it a useful resource for both novice and experienced researchers. Overall, the NeuroViz web application has been shown to greatly enhance the ability to study the brain and has the potential to contribute significantly to the advancement of neuroscience research.

Acknowledgements

The authors wish to thank collaborators Ryan P. Cabeen and Kirsten Lynch at the University of Southern California for their advice and feedback throughout the project. The authors would also like to thank Professor David H. Laidlaw for his guidance and support throughout the semester. We also thank our classmates for their feedback throughout the project.

REFERENCES

[1] R. P. Cabeen, D. H. Laidlaw, and A. W. Toga. Quantitative imaging toolkit: software for interactive 3d visualization, data exploration, and computational analysis of neuroimaging datasets. *ISMRM-ESMRMB Abstracts*, pages 12–14, 2018.

[2] R. P. Cabeen, A. W. Toga, and J. M. Allman. Frontoinsular cortical microstructure is linked to life satisfaction in young adulthood. *Brain Imaging and Behavior*, 15(6):2775–2789, 2021.

[3] R. P. Cabeen, A. W. Toga, and J. M. Allman. Mapping frontoinsular cortex from diffusion microstructure. *Cereb. Cortex*, June 2022.

[4] H. Daniel, R. Nicolas, A. Banu, G. Ellen, and P. Rudolph. Neuroimaging in the browser using the x toolkit. *Frontiers in Neuroinformatics*, 8, 2014.

[5] H. Fukutomi, M. F. Glasser, H. Zhang, J. A. Autio, T. S. Coalson, T. Okada, K. Togashi, D. C. Van Essen, and T. Hayashi. Neurite imaging reveals microstructural variations in human cerebral cortical gray matter. *NeuroImage*, 182:488–499, 2018.

[6] J. S. Gao, A. G. Huth, M. D. Lescroart, and J. L. Gallant. Pycortex: An interactive surface visualizer for fmri. *Frontiers in Neuroinformatics*, 9, 2015.

[7] T. Sherif, N. Kassis, M.- Rousseau, R. Adalat, and A. C. Evans. Brainbrowser: Distributed, web-based neurological data visualization. *Frontiers in Neuroinformatics*, 8, 2015.

[8] F. Zhang, T. Noh, P. Juvekar, S. F. Frisken, L. Rigolo, I. Norton, T. Kapur, S. Pujol, W. Wells, A. Yarmarkovich, and et al. Slicerdmri: Diffusion mri and tractography research software for brain cancer surgery planning and visualization. *JCO Clinical Cancer Informatics*, (4):299–309, 2020.

Morphology Quantification and Auto-Diagnosis for Dry Eye Disease

Yuxuan Zhao* Shixuan Li† Xiuli Wang‡

Brown University, Beijing CYL Hospital

ABSTRACT

We developed a novel and accurate model and an online platform to diagnose dry eye disease, saving time for patients and doctors. Our diagnosis model took in patients' eyelid images, segmented Meibomian Glands by deep learning model, generated morphology information, and classified atrophy for every gland. Then we built an online platform on top of this diagnosis model, and patients can send their eyelid images to this platform and receive treatment advice from doctors. Doctors can review and update patients' health records, auto-annotated Meibomian Glands, and dry eye disease diagnosis. In our study, we delivered significantly higher accuracies on dry eye disease (up to 80.1%), more specific morphological details for every gland, and a very favorable online platform based on feedback from doctors and patients.

Keywords: Deep Learning, Dry Eye Disease, Auto Diagnosis

1 INTRODUCTION

1.1 Background

Dry eye disease affects one in five adults, and most patients have mild symptoms and only need eye drops treatment[6]. Consequently, doctors spent much time diagnosing mild dry eye patients by annotating and analyzing Meibomian Glands atrophy from their eyelid images. To help doctors diagnose accurately and quickly, we developed a model to analyze the Meibomian Glands of these eyelid images and provide severity and treatment suggestions.

We created an online platform based on this model and deployed it for the Beijing CYL hospital. The ratio of doctors to patients in China is inadequate in big cities, so doctors are usually occupied. According to Beijing CYL hospital's record, dry eye disease patients take up to 15% of doctors' outpatient time. At the same time, most of them are mild, so they only need a fundamental inspection. Our automated diagnosis platform can save doctors three-quarters of diagnosis time based on their observations and estimation. Thus doctors can have more time to treat and manage those severe patients and increase the bed turnover rate.

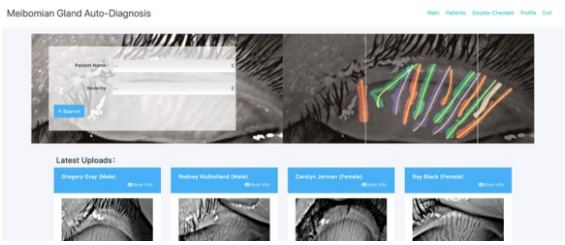


Figure 1: Patient platform

*e-mail: yuxuan_zhao@brown.edu
†e-mail: shixuan_li@brown.edu
‡e-mail: wangxla@sina.com

1.2 Domain Goals

We have three scientific aims: the first is a deep learning solution for Meibomian Glands segmentation, identifying every individual gland without any overlap. The second is a new set of criteria for analyzing Meibomian Glands, which is preferable for doctors and medical regulations. The third one is a new workflow to use fewer doctors' and patients' time.

2 RELATED WORKS

Arita et al. proposed the meiboscore stages as a standard evaluation metric to represent the severity of Meibomian Gland dysfunction. Arita et al. divided the lost Meibomian Gland area by the total eyelid area and made a threshold of its fractional result as meiboscore. Many later studies [2, 4, 10] found a correlation between morphological features such as length, curvature, tortuosity, and local contrast with Meibomian Gland dysfunction severity.

Lin et al. [8] segmented Meibomian Gland via a polygon selections function from the ImageJ software. Besides ImageJ, Saha et al. [11] segmented the glands with deep learning and got a consistent mean ratio of 25.12% (manual annotation is 26.23%). Liu et al. [9] had the first instance segmentation with point-level annotation for each Meibomian Glands. Nevertheless, it only had the location information and spline of each gland without the actual shape of the annotation. Dai et al. [3] added more morphological features such as ocular surface disease index questionnaire (OSDI), tear-meniscus height (TMH), tear break-up time (TBUT), and corneal fluorescein staining (CFS). They had a solution with 91% IoU and 100%. Zhang et al. [12] built models based on transfer learning (Mask-RCNN and U-Net) for meibomian gland segmentation and evaluated the efficacy of MG density in the diagnosis of MG dysfunction. The study resulted in 93% IoU and 100% repeatability for tarsus segmentation.

These studies about Meibomian Glands segmentation either could not provide the individual gland segmentation and morphological metric or could not provide the diagnosis of dry eye disease severity. However, our approach can provide both of them. Moreover, Mask-RCNN has the highest gland segmentation accuracy, and our solution outperforms it.

3 APPROACH

3.1 Auto-Diagnosis Model

We divided the diagnosis of dry eye disease into three steps. In the first step, we adopted Brabandere et al.'s framework [5] and added some changes to train a gland instance segmentation model with real-life eyelid images from the OCULUS system in CYL Hospital. The model took a grayscale image of the eyelid taken by the patient as input and outputted pixel-level gland regions. Traditional instance segmentation cannot handle the intersection and overlap between glands well, so they can only output semantic results. However, we had better instance segmentation by using an end-to-end model, projecting the information of each pixel onto a 1024-dimensional space, separating each gland by a clustering algorithm, and refining by a GAN model [7].

In the second step, we used the morphological analysis method to take these segmented glands as input and output shape characteristics of each gland, such as length, width, tortuosity, and curvature

degree. As we mentioned in the introduction, doctors need to observe each gland's shape and atrophy degree to propose the severity of dry eye disease. Therefore, these parameters of the gland returned at this stage can help doctors measure the condition more quantitatively and accurately, so they can improve the efficiency of diagnosis.

Lastly, based on the morphology feature of the Meibomian Gland that we exported earlier, we built a statistical model to predict the degree of atrophy of each gland. Doctors can see the overall diagnosis result and review each specific gland's prediction results to reason their own opinion.

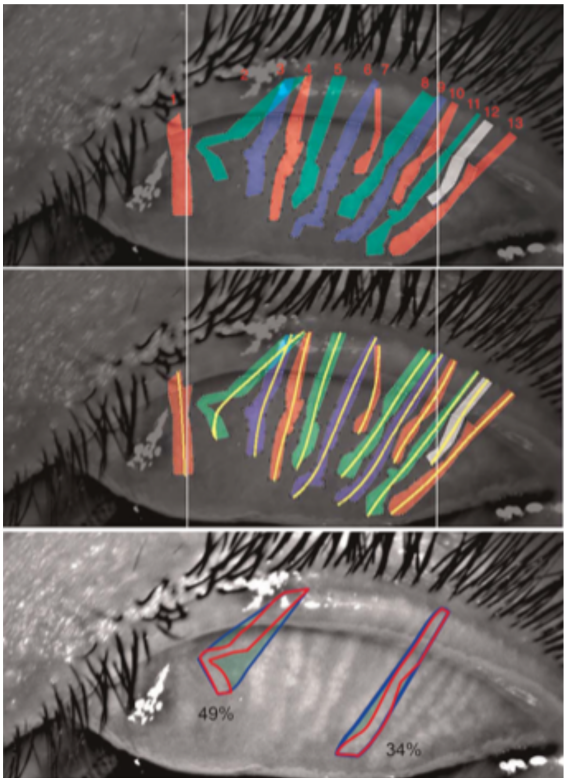


Figure 2: Gland Segmentation and Morphology Analysis

3.2 Online Platform

We developed a web application with the Django framework based on the dry eye disease automatic diagnosis model we created earlier. The front end of the website was implemented with HTML5, and the back end managed queries with MySQL. Patients can roll up their eyelids and take a grayscale photo. Then they need to register personal accounts and set passwords. After creating an account, patients logged in to the website and uploaded their eyelid photos to the hospital's database. The automatic diagnosis model would segment the image and predict the severity of each Meibomian Gland's atrophy. Doctors must also log in to this website and review the diagnosis results. After validation and confirmation by the doctor, the results would be sent to the patient's mailbox and displayed on the website's homepage.

3.3 User Study

We gave out questionnaires to 46 participants. To specify, there are 7 interns, 9 residents, 4 chief residents, 2 fellows, 3 attending physicians, 1 department head, and 20 patients.

4 Results and Discussion

There is no standard public database of eyelid images, and other scientific research is based on its private database, which prevents us from directly comparing our results with the results in other papers. Therefore, we have reproduced the most mainstream and accurate Mask-RCNN model and conducted fine-tuning based on our eyelid image data. We used its results as a benchmark, and our model outperformed it in binary classification, number of glands, and severity (shown in Figure 3).

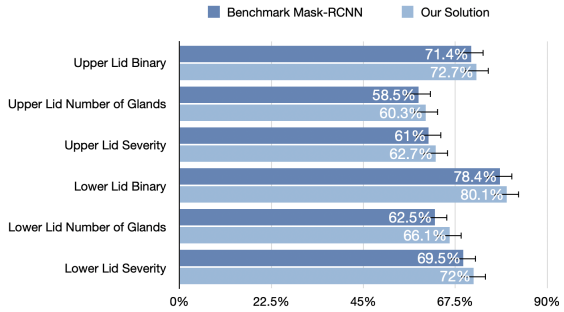


Figure 3: Accuracy Comparison

Also, the doctors from Beijing CYL hospital estimated that the hospital would save three-quarters of diagnosis time and improve the monthly bed turnover rate. Bed Turnover Rate (BTR) is defined as the number of patients treated per bed in a month [1]. In terms of the patients, they can save a reasonable inspection fee ($20).

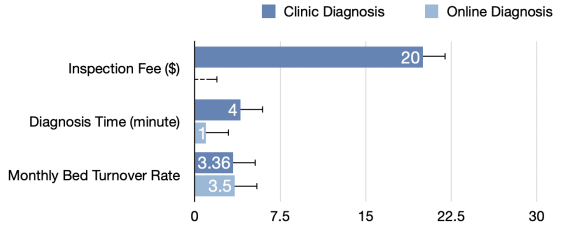


Figure 4: Positive Impact on Patients and Doctors

Regarding user studies, doctors were tired of treating many mild dry eye patients. All of them preferred the online diagnosis workflow and considered the morphology information of Meibomian Glands very useful for diagnosis.

For the patients, 17 of 20 participants (except three elderly patients) thought this platform saved much time in traveling and waiting. The timely diagnosis results with their glands annotation were favorable to the participants.

5 Conclusion

Our dry eye disease diagnosis model is the most accurate (up to 80.1%) and state-of-the-art, and we built a fully-featured favorable online diagnosis platform with it. Our platform leveraged outpatient pressure for dry eye disease for hospitals, saving three-quarters of diagnosis time and labor. In terms of the patients, our platform helped them to save travel and waiting time and the inspection fee (average $20).

Acknowledgements

Special thanks to our professor, Dr. David Laidlaw, for his guidance, support, and feedback throughout the project. Thank Jiayu Ren and Jian Hao for their cooperation throughout the semester.

REFERENCES

[1] H. E. Aloh, O. E. Onwujekwe, O. G. Aloh, and C. J. Nweke. Is bed turnover rate a good metric for hospital scale efficiency? a measure of resource utilization rate for hospitals in southeast nigeria. *Cost Effectiveness and Resource Allocation*, 18(1):1–8, 2020.

[2] S. Challinor. How to treat miscros meibomian gland dysfunction.

[3] Q. Dai, X. Liu, X. Lin, Y. Fu, C. Chen, X. Yu, Z. Zhang, T. Li, M. Liu, W. Yang, et al. A novel meibomian gland morphology analytic system based on a convolutional neural network. *IEEE Access*, 9:23083–23094, 2021.

[4] E. Daniel, M. G. Maguire, M. Pistilli, V. Y. Bunya, G. M. Massaro-Giordano, E. Smith, P. A. Kadakia, P. A. Asbell, et al. Grading and baseline characteristics of meibomian glands in meibography images and their clinical associations in the dry eye assessment and management (dream) study. *The ocular surface*, 17(3):491–501, 2019.

[5] B. De Brabandere, D. Neven, and L. Van Gool. Semantic instance segmentation with a discriminative loss function. *arXiv preprint arXiv:1708.02551*, 2017.

[6] Q. Findlay and K. Reid. Dry eye disease: when to treat and when to refer. *Australian prescriber*, 41(5):160, 2018.

[7] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[8] X. Lin, Y. Wu, Y. Chen, Y. Zhao, L. Xiang, Q. Dai, Y. Fu, Y. Zhao, and Y.-e. Zhao. Characterization of meibomian gland atrophy and the potential risk factors for middle aged to elderly patients with cataracts. *Translational vision science & technology*, 9(7):48–48, 2020.

[9] X. Liu, S. Wang, Y. Zhang, and Q. Yuan. Scribble-supervised meibomian glands segmentation in infrared images. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 18(3):1–23, 2022.

[10] C. J. Nien, S. Massei, G. Lin, C. Nabavi, J. Tao, D. J. Brown, J. R. Paugh, and J. V. Jester. Effects of age and dysfunction on human meibomian glands. *Archives of ophthalmology*, 129(4):462–469, 2011.

[11] R. K. Saha, A. Chowdhury, K.-S. Na, G. D. Hwang, Y. Eom, J. Kim, H.-G. Jeon, H. S. Hwang, and E. Chung. Ai-based automated meibomian gland segmentation, classification and reflection correction in infrared meibography. *arXiv preprint arXiv:2205.15543*, 2022.

[12] Z. Zhang, X. Lin, X. Yu, Y. Fu, X. Chen, W. Yang, and Q. Dai. Meibomian gland density: An effective evaluation index of meibomian gland dysfunction based on deep learning and transfer learning. *Journal of Clinical Medicine*, 11(9):2396, 2022.