Beyond Sight: Empowering Visually Impaired Users with Audible Graphs

Wajdi Aljedaani
University of North Texas

Uday Kiran Chimpiri
University of North Texas

Durgasantosh Gaddam
University of North Texas

Vaseem Ahammed Shaik
University of North Texas

Yaswitha Karasala
University of North Texas

Marcelo M. Eler
University of São Paulo

ABSTRACT

Data visualization is a cornerstone of modern information processing, enabling users to extract insights from complex datasets. However, individuals with visual impairments face significant barriers in accessing and interpreting visual data representations. In response, this paper presents a novel approach to data visualization designed specifically for visually impaired users. Our tool converts Excel data into audible graphs and transforms images into textual or spoken descriptions, enhancing accessibility and inclusivity in data analysis. Leveraging advanced technologies such as natural language processing and speech synthesis. Our tool empowers users to engage with visual data effectively, facilitating greater autonomy and participation in data-driven decision-making processes. Through user-centered design and innovative assistive technologies, this research contributes to bridging the accessibility gap in data visualization, ensuring that individuals of all visual abilities can benefit from the insights hidden within complex datasets.

CCS CONCEPTS

• **Human-centered computing → Statistical Analysis; Visualization tools**.

KEYWORDS

Data Visualization, Accessibility, Voice Assistant, Optical Character Recognition(OCR).

ACM Reference Format:
Wajdi Aljedaani, Uday Kiran Chimpiri, Durgasantosh Gaddam, Vaseem Ahammed Shaik, Yaswitha Karasala, and Marcelo M. Eler. 2024. Beyond Sight: Empowering Visually Impaired Users with Audible Graphs. In *The 21st International Web for All Conference (W4A '24), May 13–14, 2024, Singapore, Singapore.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3677846.3677864

W4A '24, May 13–14, 2024, Singapore, Singapore

ACM ISBN 979-8-4007-1030-8/24/05
https://doi.org/10.1145/3677846.3677864

1 INTRODUCTION

In the era of data-driven decision-making, the power of visual representation through graphs, charts, and images cannot be overstated [7]. These visualizations enable users to comprehend complex datasets, identify patterns, and derive meaningful insights. However, for individuals with visual impairments, accessing and interpreting visual data can pose significant challenges, limiting their participation in data analysis and decision-making processes [3, 3, 6, 8]. Consequently, there is a pressing need for innovative solutions that enable visually impaired individuals to engage effectively with visual data.

In response to this challenge, this paper introduces a pioneering approach to data visualization tailored specifically for individuals with visual impairments. Our tool addresses the limitations of traditional visual representations by converting Excel data into auditory graphs and transforming images into textual or spoken descriptions. By leveraging advancements in natural language processing and speech synthesis technologies, our tool aims to bridge the accessibility gap, empowering visually impaired users to explore, analyze, and interpret diverse datasets. By providing visually impaired users with the tool to engage meaningfully with visual data, we strive to foster greater autonomy and participation in data-driven decision-making processes.

2 TRANSLATING DATA INTO INSIGHTS

The objective of the proposed tool for visually impaired people is to enhance accessibility and inclusivity in data analysis by providing a seamless conversion process from Excel spreadsheets, graphs, and images into formats that are easily perceivable and interpretable through auditory. Specifically, the tool aims to:

- **Convert Excel Data into Auditory Graphs.** Our approach uses algorithms and techniques to translate numerical data and graphical representations commonly found in Excel spreadsheets into auditory graphs that convey the same information effectively through sound.
- **Transform Images into Textual or Spoken Descriptions.** Our approach employs image recognition and natural language processing capabilities to extract relevant information from graphical images and convert it into descriptive textual or spoken formats that can be easily understood by visually impaired users.
- **Ensure Usability and Comprehensibility.** Our approach leverages user interfaces and interaction mechanisms that facilitate intuitive navigation and exploration of visual data, ensuring that visually impaired users can interact with the converted information efficiently and accurately.

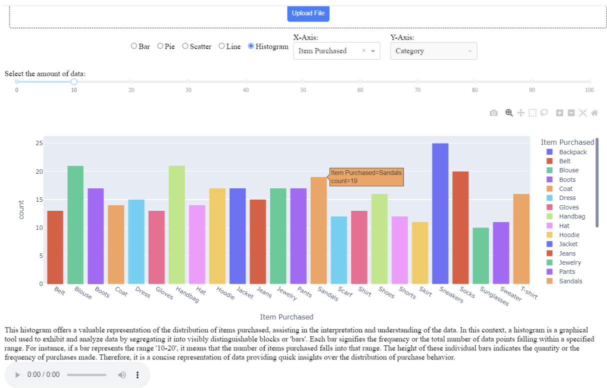

Figure 1: An image of bar graph generated after the sample data taken for data visualization

- **Promote Independence and Engagement.** Our approach empowers visually impaired individuals to independently access, analyze, and interpret visual data, enabling them to participate more fully in data-driven decision-making processes across various domains and disciplines. By achieving these objectives, the data visualization tool aims to break down barriers to information accessibility for visually impaired individuals, fostering greater autonomy, empowerment, and inclusivity in the analysis and interpretation of visual data.

3 VISUALIZATION APPROACH

In this section, we delineate the strategies we adopt to make visualization more accessible, considering different dimensions.

3.1 Extracting Data from Dataset

Our approach is based on a fully automated deep learning-based pipeline that takes a raster dataset as input and outputs an interactive chart with its original data represented in an organized manner. Our tool supports line, pie, and bar charts. We impose the following limits on chart styles because the shapes of charts on web pages vary. The tool can display bar charts in the vertical orientation; stacked and horizontal bar charts are not supported. Legends for pie charts and line charts ought to be colored differently. We train deep neural networks for data extraction, text region identification and recognition, and chart-type classifications.

3.2 Chart Classification

Our tool uses various techniques, depending on the type of graphic, to extract data from it. Initially, input photographs are categorized into four groups: pie charts, histograms, bar charts, and line charts, among others. Images without charts and other chart kinds that our tool does not support are categorized as *others*. We used Convolutional Neural Networks CNN [1, 2, 5] as our classification model since they have demonstrated remarkable results when applied to picture classification applications. In contrast to prior study [4] that employed Alex Net and Google Net for classification, we employ residual networks that produce cutting-edge results in the majority of computer vision applications. To be more precise, we use an already-existing Resnet that has been trained on the Image Net dataset, and we append a global average pooling layer prior to the final fully connected layer. Then, using the Adam optimizer and setting the learning rate to 0.0005, we fine-tune the model on our dataset. The pixels of each image were scaled to 512 x 512.

3.3 Text Extraction

Labels are crucial to effectively communicate the underlying meanings of the data in a chart. The number of categories and the encoding used for each category are shown in the chart legends. The plots' true values are displayed on the x and y axes. A chart's title, x label, and y label provide specific details about the data. In order to accurately extract data from a chart, we have to identify the text values and ascertain the purpose of the text. Text values are extracted, and text sections are localized using human supervision. As a result, we used deep learning to extract text automatically. The process of text extraction is divided into three smaller tasks: text role categorization, which determines if a given text string matches legends, axis labels, tick labels, or titles; textual region detection; and text recognition, referred to as Optical Character Recognition.

3.4 Converting-Text-Speech

Textual descriptions can further be converted into speech using Text-to-Speech technology (TTS). This allows users to listen to the descriptions of the graphs and images using screen reader assistive technologies. Figure 1 presents an example of the tool that generates text from the image and then converts the text to speech.

4 CONCLUSION

Our technique involves automatically reassembling representations that are represented in raster pictures into their original geometric forms. The underlying data may then be extracted using this information. We will use this data to help the visually challenged comprehend chart pictures. Our technique connects to existing screen readers, enabling a vast population to access data that was previously restricted. Any raster chart can be automatically detected, extracted, and presented by our model as a raw data table, a multi-level textual summary, or a thumbnail.

REFERENCES

[1] Hina Afzal, Furqan Rustam, Wajdi Aljedaani, Muhammad Abubakar Siddique, Saleem Ullah, and Imran Ashraf. 2024. Identifying fake job posting using selective features and resampling techniques. *Multimedia Tools and Applications* 83, 6 (2024), 15591–15615.

[2] Wajdi Aljedaani, Furqan Rustam, Mohamed Wiem Mkaouer, Abdullatif Ghallab, Vaibhav Rupapara, Patrick Bernard Washington, Ernesto Lee, and Imran Ashraf. 2022. Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry. *Knowledge-Based Systems* 255 (2022), 109780.

[3] Paulo Sérgio Henrique Dos Santos, Alberto Dumont Alves Oliveira, Thais Bonjorni Nobre De Jesus, Wajdi Aljedaani, and Marcelo Medeiros Eler. 2023. Evolution may come with a price: analyzing user reviews to understand the impact of updates on mobile apps accessibility. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*. 1–11.

[4] Daekyoung Jung, Wonjae Kim, Hyunjoo Song, Jeong-in Hwang, Bongshin Lee, Bohyoung Kim, and Jinwook Seo. 2017. Chartsense: Interactive data extraction from chart images. In *Proceedings of the 2017 chi conference on human factors in computing systems*. 6706–6717.

[5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 6 (2017), 84–90.

[6] Obianuju Okafor, Wajdi Aljedaani, and Stephanie Ludi. 2022. Comparative Analysis of Accessibility Testing Tools and Their Limitations in RIAs. In *International Conference on Human-Computer Interaction*. Springer, 479–500.

[7] Alberto Dumont Alves Oliveira, Paulo Sérgio Henrique Dos Santos, Wilson Estécio Marcílio Júnior, Wajdi M Aljedaani, Danilo Medeiros Eler, and Marcelo Medeiros Eler. 2023. Analyzing Accessibility Reviews Associated with Visual Disabilities or Eye Conditions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.

[8] Rahman Shafique, Wajdi Aljedaani, Furqan Rustam, Ernesto Lee, Arif Mehmood, and Gyu Sang Choi. 2023. Role of Artificial Intelligence in Online Education: A Systematic Mapping Study. *IEEE Access* (2023).