



Article

# Designing a Tactile Document UI for 2D Refreshable Tactile Displays: Towards Accessible Document Layouts for Blind People

Sara Alzalabny <sup>1,\*</sup> , Omar Moured <sup>2,3</sup> , Karin Müller <sup>2</sup>, Thorsten Schwarz <sup>2</sup> , Bastian Rapp <sup>4</sup> and Rainer Stiefelhagen <sup>2,3</sup>

<sup>1</sup> NeptunLab, University of Freiburg, 79110 Freiburg, Germany

<sup>2</sup> ACCESS@KIT, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany; omar.moured@kit.edu (O.M.); thorsten.schwarz@kit.edu (T.S.); rainer.stiefelhagen@kit.edu (R.S.)

<sup>3</sup> CVHCI@KIT, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

<sup>4</sup> Freiburg Center of Interactive Materials and Bioinspired Technologies, University of Freiburg, 79110 Freiburg, Germany

\* Correspondence: sara.zalabny@neptunlab.org

**Abstract:** Understanding document layouts is vital for enhancing document exploration and information retrieval for sighted individuals. However, for blind and visually impaired people, it becomes challenging to have access to layout information using typical assistive technologies such as screen readers. In this paper, we examine the potential benefits of presenting documents on two-dimensional (2D) refreshable tactile displays. These displays enable the tactile perception of 2D data, offering the advantage of dynamic and interactive functionality. Despite their potential, the development of user interfaces (UIs) for such displays has not advanced significantly. Thus, we propose a design of an intelligent tactile user interface (TUI), incorporating touch and audio feedback to represent documents in a tactile format. Our exploratory study for evaluating this approach revealed satisfaction from participants with the experience of directly viewing documents in their true form, rather than relying on screen-reading interpretations. Additionally, participants offered recommendations for incorporating additional features and refining the approach in future iterations. To facilitate further research and development, we have made our dataset and models publicly available.

**Keywords:** tactile user interface; document layout; 2D refreshable tactile display; accessible layout



**Citation:** Alzalabny, S.; Moured, O.; Müller, K.; Schwarz, T.; Rapp, B.; Stiefelhagen, R. Designing a Tactile Document UI for 2D Refreshable Tactile Displays: Towards Accessible Document Layouts for Blind People. *Multimodal Technol. Interact.* **2024**, *8*, 102. <https://doi.org/10.3390/mti8110102>

Academic Editors: Wei Liu, Jan Auernhammer and Takumi Ohashi

Received: 8 October 2024

Revised: 30 October 2024

Accepted: 4 November 2024

Published: 8 November 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The graphical layout of a document encompasses visual semantics, including the size, shape, and location of the content matter [1–3]. It is not merely about organizing visual and textual elements. It also significantly contributes to faster comprehension, easy reading, and better understanding. Effective layout design enhances information retrieval and aids in maintaining the reader's attention [4]. For instance, when reading a scientific paper, one can easily navigate between the sections to find specific table or equation references. Reading order is also inherited from the layout presentation of the document. However, for blind and visually impaired people, the lack of accessibility in document layout poses a barrier to achieving similar reading advantages as their sighted peers [5]. A study of 11,000 scientific PDFs showed that the majority lack basic accessibility features, with only 2.4% satisfying all accessibility requirements [6]. Further investigations demonstrated that layout knowledge could enable blind and visually impaired people to collaborate and co-author effectively with sighted individuals [2].

Significant research efforts have been made in the field of Document Layout Analysis (DLA) [7–10], focusing on extracting key features from documents by segmenting and identifying different regions within them. Despite substantial advancements in DLA, there

remains a gap in research on effectively extracting and presenting layout information in formats accessible to blind individuals. Recent studies in computer vision [11,12] have concentrated on editing and extracting visual layouts, yet they have not addressed how to adapt these layouts for users with diverse sensory needs. Moreover, many of these approaches fall short of the accessibility standards required for blind individuals, particularly those who depend on tactile feedback.

The common tools used by blind and visually impaired individuals to access digital documents generally fall into two categories: those that interpret data linearly and those that preserve the original document's format, thereby maintaining its spatial layout. Assistive technologies like screen readers and one-line Braille displays present information in a linear sequence, either through audio or Braille feedback, without conveying spatial structure or layout organization [13]. In contrast, printed-embossed documents and 2D tactile displays retain spatial information by rendering the original layout, while printed-embossed solutions are often more affordable but time-consuming and prone to wear and tear. On the other hand, 2D refreshable tactile displays offer interactive features and support multimodal exploration, making them an increasingly popular option despite their higher cost [14].

There has been limited progress in the accessibility community regarding how blind users can interact with document content through tactile means, possibly due to the absence of standardized design principles [15]. Therefore, in this work, we aim to analyze how UIs for tactile displays can be designed using more standardized approaches and assess their impact on enhancing document comprehension for blind and visually impaired users compared to traditional methods.

In this paper, we propose a method for automatically extracting layout information from various types of documents and presenting it in an accessible format for blind individuals using 2D refreshable tactile displays. To achieve this, we employ a deep learning approach to automatically retrieve document structures, using a benchmark dataset specifically developed for this task. Additionally, we present a design of an intelligent TUI to convey document structure to blind users through 2D tactile displays. Our system further supports multimodal interactions, allowing users to explore documents interactively through both tactile feedback and audio. The key contributions of this work are as follows:

- We design a method to automatically extract document layouts and a new dataset comprising common complex layouts, including slides and newspapers.
- We construct an interactive TUI for displaying documents on 2D refreshable tactile displays to enable blind users to navigate and explore complex document layouts through haptic and button-based interactions and auditory feedback.

## 2. Related Work

In this section, we discuss research efforts focused on DLA and converting documents into accessible formats. Moreover, we discuss the existing literature on tools and methods developed to assist blind individuals in reading documents.

### 2.1. Document Layout Analysis Methods

Traditionally, the field of DLA focused on enterprise applications, where businesses require efficient processing of large volumes of documents to perform tasks such as information retrieval and key-value extraction [10]. Recent advancements in deep learning have redefined DLA as a visual object detection or segmentation problem, utilizing convolutional neural networks (CNNs), such as the YOLO family models [16]. In contrast to CNN-based methods, models like DiT [17] and RoDLA [18] leverage image transformers designed explicitly for DLA, yielding promising results. Moving beyond single-modal approaches, LayoutLMv3 [19] integrates text, vision, and layout information into a unified architecture, pushing the boundaries of multi-modal document understanding.

Not only have the methods and architectures in the DLA field advanced but benchmark datasets have also seen significant progress. PubLayNet [20] was one of the first large-scale

DLA benchmark datasets, followed by DocLayNet [21], which offers 80,000 documents from various categories. Both datasets incorporate only two modalities—visual and layout. Other more specialized datasets, such as FUNSD [22] and ReadingBank [23], expand on these by including textual, visual, layout, and reading order modalities.

However, in daily life, especially for blind and visually impaired individuals, documents beyond traditional structures are frequently encountered. These can include magazines, fliers, or other designs where layouts deviate from common structures, often incorporating creative design elements. Existing datasets in the field, such as DocLayNet and PubLayNet, focus solely on structured documents, which limits model reliability and robustness in real-world scenarios where blind users might capture or scan documents to access information.

On the other hand, current accessibility approaches focus primarily on element-wise accessibility. For instance, they focus on the utilization of large vision–language models to summarize and understand documents [24]. Omar et al. [25] proposed a method to create tactile materials from document visuals like charts using an intelligent interface. In [26], Sechayk et al. proposed screen-reading applications to make slide decks more accessible. Wang and Cachola et al. [27] applied specialized models to convert PDF metadata into HTML, demonstrating strong performance in extracting and presenting document structures. However, the approach is computationally intensive due to the reliance on multiple models and is limited to PDFs with a text layer, restricting its broader applicability. While these models are valuable for reading documents and improving accessibility, they do not fully support blind and visually impaired users in interacting with and exploring different document sections independently.

Despite these advancements, challenges remain in achieving robustness across diverse document types and ensuring accessibility of the various layout structures commonly used in documents. In our approach, we aim to develop a more abstract module that enables interactions with the document as a whole, allowing users to explore various sections and gain a holistic understanding of the layout and content.

## 2.2. Document Access by Screen Readers

When it comes to accessing documents, screen readers are the most commonly used assistive technology by blind and visually impaired individuals. Screen readers can offer a sequential or a linear representation of digital documents. Therefore, the documents must be provided in an accessible format, with all relevant metadata included. This sequential representation, while informative, sacrifices alignment with the original spatial layout of the document [28].

Several studies explored how documents can be converted into an alternative one-dimensional representation. This representation aims to convey the same content with reduced complexity for better access through screen readers. A popular approach involves converting the document into an HTML format [29,30]. Wang et al. [31] proposed an approach for scientific documents. They first identified the reading order of different elements and then mapped it into a one-dimensional hierarchical HTML structure. Similarly, Peng et al. [32] presented a method to enhance slide document accessibility by generating a hierarchical structure from a slide deck. This transformation allows readers to navigate the slide deck from higher to lower-level descriptions. However, this transition from a 2D to a 1D mapping of document layouts comes at the expense of losing spatial information.

Previous works have also addressed how screen readers can be improved by exploring ways to enhance the interactive aspects of screen readers. For instance, Khurana et al. introduced SPRITES, a system that utilizes various keyboard surface interactions to represent visual-spatial elements such as page menus, lists, and maps [33]. Similarly, Gadde et al. presented DASX, a system that allows users to quickly navigate to the most relevant section within a page using a single shortcut [34]. Another innovative approach, proposed by Vtyurina et al. [35], integrates smartwatch/smartphone applications with screen readers. This approach incorporates hand gestures and voice commands to navigate the document,

enhancing the reading experience for blind individuals [35]. Ahmed et al. [36] introduced another method to allow blind users to skim through documents without relying on visual cues. In their approach, users are presented with keywords collected from the page text, organized based on topic similarity. Subsequently, a summary is generated from these keywords. Their research findings suggest that this technique can be valuable for providing screen reader users with a high-level overview of a page.

Despite the efforts to use screen readers to represent document layouts, these tools have typically been tested with accessible documents, simple graphical elements, and predefined document categories such as webpages. This assumption may be challenging to uphold in today's documents, which often feature variable or non-standard layouts and incorporate more visuals like charts and equations.

### 2.3. Haptic Representation of Documents

Alternatively, some approaches in the literature focused on conveying document layouts while preserving their 2D structure, often by incorporating haptic feedback. Safi et al. [37] introduced a method using vibration sensors to represent document layouts. In this research, they developed a finger-sized embedded system with a micro-vibrator. The intensity and the frequency of the vibration are adjusted based on the structure of the touched document on the tablet. Similarly, Maurel et al. [38] proposed a strategy for allowing blind users to access vibrotactile documents using hand-mounted actuators that deliver localized vibrations with varying intensities and frequencies. By correlating light intensity with vibration frequency, users could distinguish between different elements and borders within the layout. Moreover, while their research focused on webpage layouts, the evaluation did not include blind users.

Other similar methods utilized embossed papers [2,39,40] for representing documents. In this approach, variant heights and textures are assigned to document elements. Another study followed a similar approach but used wooden stacked laser-cut pieces to convey the document [41]. A significant advantage of these printed tactile materials is that they are affordable; however, they only convey static layouts. Once printed, they cannot be altered and provide no interactions. In comparison, 2D refreshable tactile displays can surpass these conventional layout representation methods by providing blind and visually impaired individuals with a dynamic and more user-friendly experience. This is achieved through the utilization of interactive features, including page toggling, panning, and zooming [42].

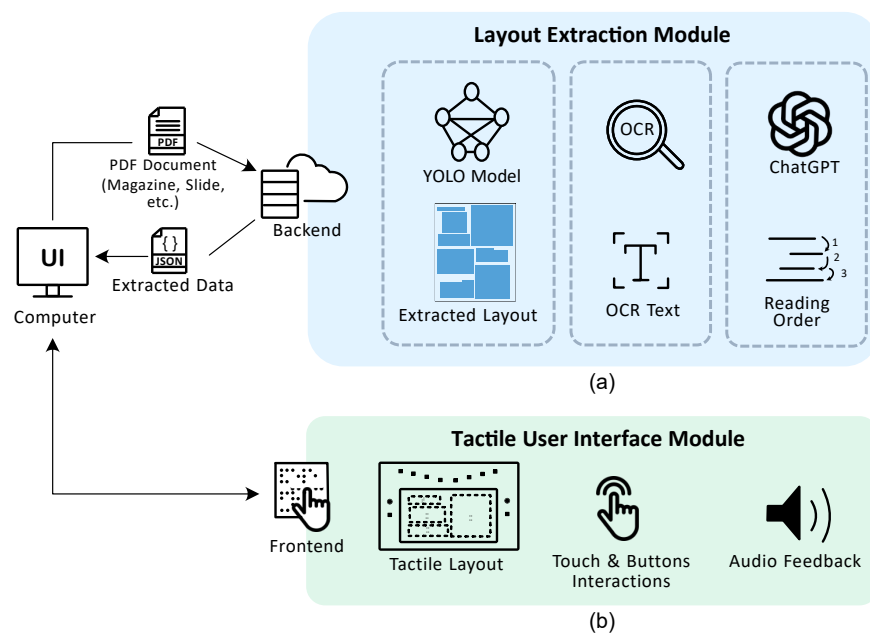
In the following section, we present our approach for automatically extracting layout information from complex documents and a new TUI for depicting the extracted information on 2D refreshable tactile displays.

## 3. Materials and Methods

In this section, we introduce our proposed system for making document layout accessible using 2D tactile refreshable displays. The system consists of two primary modules, as depicted in Figure 1: Figure 1a, layout extraction module; Figure 1b, TUI module. The layout extraction module runs on a server and consists of a trained detection model to extract metadata from documents automatically. The metadata contain the spatial layout information and the textual data. The metadata are then passed to the tactile representation module, which converts the extracted metadata into a tactile format that can be displayed on 2D refreshable tactile displays. Additionally, this module handles user interactions in the form of touch and button clicks, and provides users with audio feedback.

The interface was tested on the Metec Hyperbraille 2D refreshable tactile display [43]. The device features 60 by 104 actuators (pins) and a resolution of 10 dpi. It incorporates touch capabilities on its tactile surface. The device's hardware features a total of 19 buttons. These include two sets of navigation buttons on both sides of the tactile surface and a scroll button at the bottom. The device operates through an application programming interface (API) that enables the user to control the actuators and program the functionality of the

buttons. In the following section, we present a detailed examination of the two modules in our proposed system.



**Figure 1.** The pipeline of our tactile document system consisting of (a) the layout extraction module, which utilizes the YOLOv10 detection model and an OCR model, as well as ChatGPT to extract metadata from each predicted bounding box, and (b) the tactile representation module, which responsible for representing the document’s metadata in a tactile format. This module handles touch and button interactions, and provides audio feedback for the auditory representation of text elements.

### 3.1. Layout Extraction Module

**Dataset:** To enable effective document layout extraction, a model requires training on a diverse and representative document layout dataset. Among the primary datasets in this field is DocLayNet [21], a comprehensive collection covering categories such as financial reports, scientific articles, laws and regulations, government tenders, manuals, and patents. These categories share a standardized layout structure across sources; for instance, scientific articles from various publishers tend to exhibit similar sectioning and design. However, in real-world scenarios, layouts are often much more varied. We refer to these as artistic layouts—a category that includes visually complex documents like magazines—which are typically crafted by designers using multiple layers and artistic elements.

To address this gap, we propose the Artistic Document Layout (ArtDocLay) dataset in this paper [44]. The dataset was initially curated by crawling 1017 documents from Commoncrawl using a keyword-based filtering approach to capture eight unique artistic categories (refer to Table 1). We then manually selected 37 documents from each category, ensuring that each document showcased unique and varied layouts, even within individual pages. Following DocLayNet’s annotation guidelines, we labeled a total of 324 images and 3526 bounding boxes. Table 1 provides statistical insights into our dataset. We also provide training, validation, and test splits with a 70:10:20 percent ratio, respectively.

**Model:** To extract the layout of documents, we utilize object detection models. A crucial requirement for our model selection is the ability to operate in real-time and perform efficiently on low-power devices. For this purpose, we determined the YOLO family to be suitable, given its balance of speed and resource efficiency.

We employed the latest version, YOLOv10, as provided by Ultralytics. YOLOv10 is a single-shot detector, which makes it distinct from prior versions by incorporating NMS-free dual-label assignments and a holistic efficiency- and accuracy-driven model design. This design choice reduces computational redundancy, enhancing both speed and accuracy [16].

**Table 1.** Artistic documents included in our dataset and their corresponding book and image counts.

| Category             | Books | Image |
|----------------------|-------|-------|
| brochure             | 7     | 44    |
| newspaper            | 5     | 48    |
| books                | 3     | 84    |
| magazine             | 4     | 52    |
| slide                | 4     | 79    |
| poster               | 2     | 2     |
| flier                | 6     | 11    |
| infographic          | 4     | 4     |
| Total (8 categories) | 35    | 324   |

For comprehensive experiments, we trained all YOLOv10 variants, from YOLOv10-x (X-Large, with 29.5 Mparameters) to YOLOv10-n (Nano, with 2.3 M parameters). In the first stage, all models were trained on the DocLayNet dataset for 100 epochs, starting from COCO-pretrained weights made available by Ultralytics. In the second stage, we fine-tuned ArtDocLay’s training split for 10 epochs, using zero warmup for a smoother transition and a higher learning rate of 0.01 to accelerate learning on the new dataset. Table 2 summarizes the evaluation of the trained models on artistic document layouts.

**Table 2.** Performance of different YOLOv10 models trained on DocLayNet and fine-tuned on ArtDocLay in terms of mAP50 and mAP50:95, with improvement margins shown in blue.

| Model    | Train Dataset           | mAP50        | mAP50:95     |
|----------|-------------------------|--------------|--------------|
| YOLOv10x | DocLayNet               | 31.9         | 19.5         |
| YOLOv10b |                         | 29.5         | 17.3         |
| YOLOv10l |                         | 36.7         | 19.3         |
| YOLOv10m |                         | 33.0         | 17.7         |
| YOLOv10n |                         | 30.4         | 18.6         |
| YOLOv10s |                         | 31.8         | 18.4         |
| YOLOv10x | Fine-tuned on ArtDocLay | 56.6 (+24.7) | 31.5 (+12.0) |
| YOLOv10b |                         | 55.7 (+26.2) | 34.4 (+17.1) |
| YOLOv10l |                         | 60.5 (+23.8) | 33.7 (+14.4) |
| YOLOv10m |                         | 53.7 (+20.7) | 30.6 (+12.9) |
| YOLOv10n |                         | 64.8 (+34.4) | 29.0 (+10.4) |
| YOLOv10s |                         | 52.8 (+21.0) | 29.6 (+11.2) |

For a fair evaluation, we ensured consistent image resolution across all models. Table 2 presents the test performance on the artistic dataset, highlighting the limited performance of models trained solely on the DocLayNet dataset, which is primarily suited to structured documents. This limitation is particularly relevant for blind users, as higher error rates can lead to frustration—such as when an element labeled as an image caption turns out to be a table or when captions go undetected entirely.

Our dataset, ArtDocLay, demonstrates a meaningful improvement, with YOLOv10-X showing a gain of +24.7 mAP50 and +12.0 mAP50:95 over DocLayNet alone. Qualitatively and quantitatively, this improvement means that bounding boxes are more accurately aligned with objects in the page, both in location and class. This alignment facilitates a better reading order and element identification, reducing hallucinations and enhancing

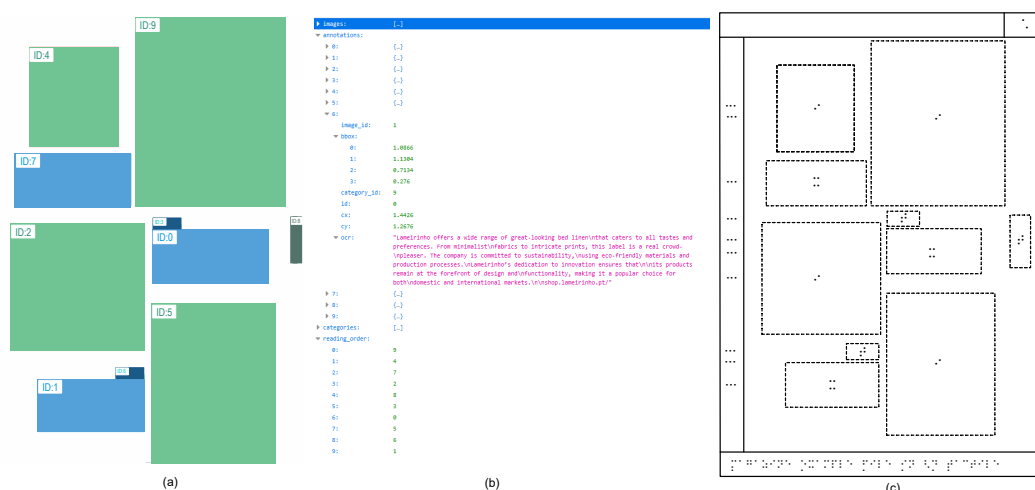
accessibility for blind users. For inference speed, the YOLOv10-X variant achieved a runtime of approximately 6 ms on an NVIDIA A40 GPU.

In our framework, we utilize the X variant with a confidence threshold of 0.5. These predictions are then processed in subsequent steps to establish reading order and generate tactile layout representations, as detailed in later sections.

**Reading Order:** For blind readers, a logical reading order is essential for understanding document content as intended. To establish this, we first annotate the page image with bounding boxes, each labeled with an ID in the top-left corner, as shown in Figure 2a. Then, use OCR to extract text from each bounding box. Non-textual elements, such as figures and charts, are paired with their captions to retain context. We then prompt GPT-4 with the annotated layout and OCR data, asking it to arrange elements into a coherent reading sequence based on spatial and textual cues. To accomplish this, we provided the model with a prompt that summarized the task as follows:

"Your task is to reorder the bounding box IDs to reflect the natural reading flow of the document page based on the content and the spatial location of the text. You will be provided with an image of the document where bounding boxes are drawn, along with metadata describing each bounding box, including coordinates, class name (e.g., 'Title', 'Text', 'Section-header'), and OCR-extracted text. Please return the bounding box IDs in the correct reading order as a JSON array."

The final reading order, along with spatial and textual content, is structured in a JSON file, as shown in Figure 2b.



**Figure 2.** An example of a document at the various stages of the system pipeline: (a) Following document segmentation, bounding boxes are generated for each element, with ids assigned to each bounding box. (b) A JSON file is created containing the bounding box coordinates, reading order, and OCR-generated text. (c) The tactile representation of the document using our interface. In the first view mode, the interface displays bounding boxes, while the second view mode shows only Braille letters and element identifiers to represent the document elements.

### 3.2. TUI Module

To design our TUI for 2D tactile displays, we followed a participatory approach by collaborating with a blind user. This collaboration aimed to define the criteria for how the interface should present documents and how the interactions should be designed. For this purpose, we conducted semi-structured interviews with our collaborator during the design phase. Through an iterative design process, we developed the interface design in two primary cycles.

#### 3.3. First Prototype: Static TUI

The first prototype we created to represent the metadata we extracted from documents using our trained model, maintaining the original document layout. The concept involved

placing a Braille character at the center of each document element. For example, a paragraph was marked by the Braille letter “x”, while a title was represented by the letter “t”, as detailed in Table 3.

**Table 3.** The detected elements using our trained model and their tactile representation.

| Detected Sub Classes                                  | Mapped Classes | Tactile Representation |
|-------------------------------------------------------|----------------|------------------------|
| Document title,<br>section title, header              | Title (t)      | ⠠                      |
| Paragraph, footer, caption,<br>page number, list-item | Text (x)       | ⠠                      |
| Table                                                 | Table (b)      | ⠠                      |
| Equation, code                                        | Math (m)       | ⠠                      |
| Figure                                                | Image (i)      | ⠠                      |

We presented the prototype to our collaborator using the Metec Hyperbraille tactile display and sought feedback on improving the representation and defining fundamental interactions. The collaborator recommended displaying the bounding boxes around document elements for the representation. Regarding the interactions, the collaborator suggested the following points:

- Implementing all interaction buttons symmetrically on both sides of the device to allow the users to explore the document using both hands and avoid losing the orientation at any time when using the buttons. This would also enhance the accessibility for left and right handed users.
- Integrating audio feedback to provide the text of each element when needed.
- Providing the user with an acoustic feedback while using the navigation buttons to signal the occurrence of a change.
- Including a feature to offer additional details, such as the number of columns, element types, and document format (e.g., paper, slide, receipt), upon request.

### 3.4. Second Prototype: Interactive TUI

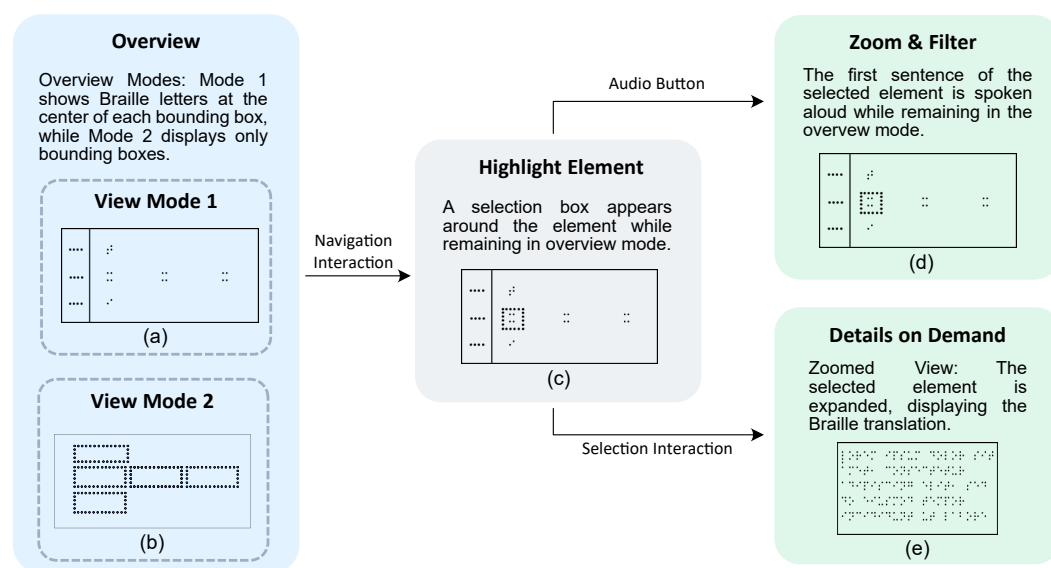
To implement the interactions suggested by our collaborator, we developed a second prototype of the interface. The design of this new interface was guided by Shneiderman’s Visual Information Seeking Mantra (VISM) [45], which provides a framework for creating interactive visualizations that facilitate efficient data exploration. Although VISM is initially intended for visual interfaces, prior research has successfully adapted its principles for designing audio interfaces for blind users [46]. Building on this, we explored how VISM’s guidelines could be applied to the design of TUIs that can run on 2D tactile displays.

#### 3.4.1. VISM Principles

The design of our second prototype involved implementing the three VISM principles: (a) Offer a summary or overview of the document on the tactile display. (b) Allow users to zoom in or out of different levels of detail by navigating between sections or levels of information. (c) Provide detailed information upon demand. When users select a specific item or area of interest, additional information about it can be displayed in a format suitable for tactile display. The following section describes the implementation of the three principles in detail.

**Overview Principle:** The overview principle was implemented in the interface by depicting three main components on the tactile display, as shown in Figure 3a. The initial component is the “Element Guide”, which is a vertical rectangular region with horizontal lines. These lines indicate the existence of a document element, such as text, title, table, math, and image, at that specific level. These lines align with the y-values of the bounding box

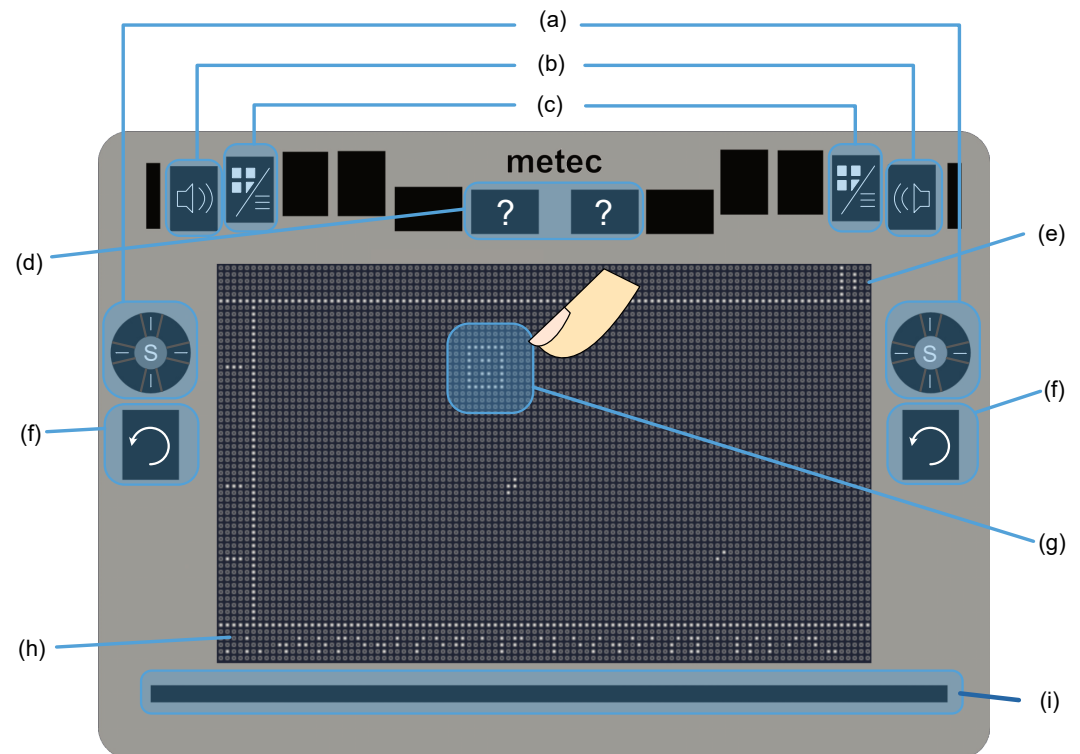
centers that hold each element. The second component is the “Class Identifier” area, where a Braille character is placed at the center of the bounding box. This character represents the layout class, for example, “x”, for a text element. The last part is the “Divider Line”, which separates the previously mentioned sections, aiding users in distinguishing between the first two components. Furthermore, there is an alternative overview known as the “bounding boxes mode”, as shown in Figure 3b. In this mode, the interface displays bounding boxes around each element, representing them as rectangles with dimensions corresponding to the width and height of the respective element. Additionally, in each view mode, two sections are always present on the screen, the header section and the footer section, as depicted in Figure 4e and Figure 4h, respectively. The header section shows the current page number, while the footer section shows the file name. The goal of the two view modes is to provide users with a quick overview of the document, enabling them to easily visualize the different elements present.



**Figure 3.** The different views available in the tactile document interface, based on the VISM principles. (a) Element identifier overview mode. (b) Bounding boxes overview mode. (c) Selection of an element to explore through navigation buttons or touch. (d) Zoom and filter view. (e) Details-on-demand view.

**Filtering Principle:** The zoom and filter principle is defined in the VISM as a method for reducing the complexity of the data representation by removing unnecessary information from the user’s view, enabling focused exploration of relevant details. In our interface, this is implemented by allowing users to select specific elements within a document while in overview mode, as demonstrated in Figure 3d. Upon selection, users receive brief acoustic feedback, consisting of the first sentence from the chosen element, providing concise information without needing to switch to a detailed view. This approach helps users grasp key content quickly while maintaining their overall orientation in the document.

**Details-on-demand Principle:** The details-on-demand principle is realized by allowing the user to select any element in the overview mode, as shown in Figure 3c, and move to another detailed view. The detailed view, as shown in Figure 3e, displays the selected element of interest on the display. The Braille translation will be displayed if the selected element is a text. If the element is an image, the alternative text will be presented. If the alternative text is unavailable, a standard text message informs the user that this is an image without alternative text.



**Figure 4.** The interactions available in the tactile document interface and the corresponding buttons used on the HyperBraille display. (a) Navigation controls, (b) audio feedback, (c) view mode button, (d) help buttons, (e) page number, (f) back button, (g) document element with a selection box, (h) file name footer, and (i) page navigation button.

### 3.4.2. Interaction Concepts

We designed the interface with fundamental interactions that facilitate document exploration and transitions between the various available view modes. Given blind users' familiarity with haptic interactions, touch interactions were prioritized. However, we also implemented the same interaction concepts using buttons to ensure accessibility for any 2D tactile display that lacks touch capabilities. The tactile and audio feedback in our design is intended to help users develop a comprehensive understanding of document structure. For instance, in overview mode with Braille letters, tactile feedback conveys spatial relationships, supporting users in forming a mental map of the document's layout. Additionally, audio signals complement specific interactions, such as distinct sounds when navigating between elements and a crash sound that indicates the document boundary. Together, these cues provide clear structural information, helping users maintain orientation within the document. In total, we integrated seven interaction concepts into our interface, which are detailed in the following sections.

**Selection and Navigation Using Touch:** The first interaction possibilities are the selection and navigation by touch. By placing one finger on the Braille letter of a document element and pressing the select button "S" in the navigation buttons, as shown in Figure 4a, the user can navigate to the details on demand view, where the complete text is displayed in Braille.

**Selection and Navigation Using Buttons:** The same interactions are executed using a set of four directional buttons and a selection button labeled "S," as depicted in Figure 4a. When the selection button is pressed, a rectangle is drawn around the first element in the document, determined by the reading order of the elements, as illustrated in Figure 4g. An auditory signal is also played to notify the user that an element has been selected. If the user attempts to navigate beyond the end of a row or the end of the document, a distinct audio signal resembling a crash is emitted to indicate that further navigation in this specific

direction is not possible. Pressing the selection button again expands the highlighted element, displaying the full text in Braille.

**Audio Feedback:** In overview mode, when an element is highlighted by placing one finger on it or using the selection button, the user can listen to the first sentence within that element by pressing the audio buttons, as shown in Figure 4b. This functionality is designed to assist blind users in skimming through a document, providing brief information about the text within the element without losing their sense of the overall layout. Similarly, if the same button is pressed while in the “Details on Demand” view mode (after selecting an element and displaying its content), the user can listen to the full text that is displayed.

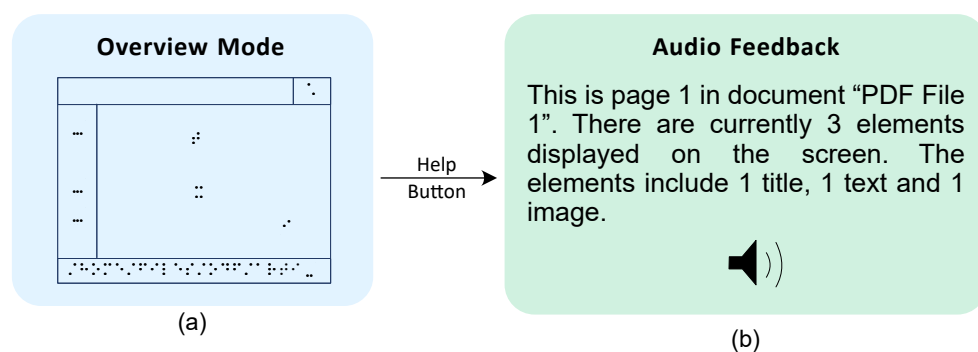
**Return to Overview Mode:** Switching back from the “Details on Demand” view, where the text element is expanded, to the overview is carried out by clicking the back buttons, as shown in Figure 4f.

**Bounding Boxes Overview Mode:** The user can switch to the bounding boxes overview mode by pressing the overview mode button, as shown in Figure 4c. However, this mode is static, meaning that the user cannot select or navigate between different elements.

**Scroll Pages:** By pressing the scroll bar in Figure 4i once, the user can scroll up and down the page. If the user holds the scroll bar longer, they can navigate to the next or previous page.

**Extra Information:** If the user presses the help buttons, as illustrated in Figure 4d, the user can obtain extra acoustic information about the file’s name, the current page, and the number and types of the elements currently visible on the screen. For instance, when the user clicks the help button, the document shown in Figure 5a would generate the following spoken text:

“This is page 1 in document PDF File 1. There are currently three elements displayed on the screen. The elements include one title, one text, and one image”.



**Figure 5.** (a) A document represented in the overview mode using the proposed interface. (b) The corresponding auditory information given to the user after clicking the help button.

#### 4. Preliminary Study

We conducted a preliminary study with three blind users to evaluate the second prototype of the interface. We were interested in observing the participants while using the interface to explore complex documents, such as magazine articles with multiple columns, and analyze how intuitive the interactions are, as well as whether this 2D representation of documents is helpful for blind users. Additionally, the user study was required to assess the effectiveness of the VISM approach for presenting documents on 2D refreshable tactile displays. We had the following three goals for our study: First, to determine how successful the users are in using the interface for reading and skimming documents with complex layouts. Second, to assess the effect of using the interface on the efficiency of reading and skimming compared to conventional aids such as screen readers. Third, to analyze the users’ satisfaction with the interface and its interactions.

#### 4.1. Participants

We recruited participants through university mailing lists. Table 4 shows the demographic data of the participants. The participants had prior experience with 2D tactile displays and depended on screen readers for reading documents. All the participants could read Braille and had a high level of education, at least a bachelor's degree or higher.

**Table 4.** Demographic information of the user study participants.

| Participant | Age/Gender | Reading Assistive Tools  |
|-------------|------------|--------------------------|
| P1          | 48-57/M    | Screen reader/1D Braille |
| P2          | 23-32/M    | Screen reader/1D Braille |
| P3          | 23-32/F    | Screen reader            |

#### 4.2. Procedure

Each study lasted around 60 min and included one participant. The participants provided their informed consent at the beginning of this study, in accordance with our university's ethical guidelines. Each study consisted of four sessions: the first session was an introductory session, during which each participant received a training document containing step-by-step instructions for 15–20 min, guiding them on how to use the interface. In the second session, the participants were provided a four-column magazine article [47] to read on the tactile display using our TUI. In the third session, the participants were given another magazine article [48] to read using the NVDA [49] screen reader on a computer.

Following each sessions two and three, the participants were asked to perform the following tasks:

- Task 1: Skim the document and give a quick summary of the topic and the main key points.
- Task 2: Answer a question about certain information in the document. (Document (a): How does the writer stay updated on the rapidly occurring changes? Document (b): Why do some individuals prefer to be anonymous online.)
- Task 3: Explain the structure of the document.

In the last session, the participants were asked to provide feedback on the ease of using the new interface and their interactions with it through open-ended questions.

#### 4.3. Results

Overall, we observed that all participants had no difficulties understanding the concept of presenting documents in a 2D tactile format. They could navigate the provided document using the interface, extract information, and answer the questions effectively. In the following section, we present our findings.

Regarding the different view modes available in the interface, all participants demonstrated an ability to figure out the document type by interpreting the structure using the bounding boxes view mode. However, when the participants were asked about their preferred view mode, all participants favored the bounding boxes view over the centered Braille letter view, mentioning that it was easier to comprehend spatial arrangement. P1 and P3 suggested combining both view modes, allowing simultaneous access to bounding boxes and the central letter views, as the central letter adds extra information about the type of element represented by the bounding box. P2 answered when asked why he preferred the bounding boxes overview more:

“It made me personally quite happy because I have just seen a document again as it was there, so not just in 1D as on the Braille 1D display, but as it really is”.

Participant P3 also confirmed that the bounding box view mode was useful in understanding the structure of the document and added that the markings provided at the left of

the display in the case of the centered Braille letter is a very useful mechanism and should also be present in the case of the bounding box view mode.

Regarding the “Details on Demand” concept, it was noted that both participants found the mechanism of switching from overview to detailed view to be logical. However, P1 was observed to have some difficulties while searching for specific information in the document. The participant later highlighted that knowing whether the scroll bar scrolled within the page or the whole page was sometimes confusing. Therefore, the participant suggested integrating feedback mechanisms when scrolling within the same page, as when moving between pages, such as playing two different audio signals in each case.

Regarding the “Zoom and Filter” concept, Participant 3 (P3) appreciated how the audio feedback is implemented in the interface, noting that hearing only the first sentence in overview mode was particularly useful. This feature helped her quickly navigate the document and skip less interesting sections.

All participants preferred touch interactions over purely button-based navigation when accessing the document. Additionally, it was noted that participants did not utilize the help function while reading. When asked about this, all participants indicated that they found it unnecessary and did not feel the need for assistance.

Regarding new features, P1 suggested integrating an additional function into our trained model that would automatically generate a summary of the document, allowing users to listen to this summary with the press of a button. Additionally, P3 expressed interest in incorporating editing options within the interface, such as highlighting specific text elements and adding comments. P3 also proposed a feature enabling users to place one finger on a word, triggering audio feedback that reads only the corresponding sentence.

In the last session, participants were tasked with skimming the second document using a screen reader. The screen reader was unable to interpret the document’s structure correctly, resulting in an incorrect reading order. This made it difficult for users to understand the topic discussed in the document, navigate through the document, and access specific information.

## 5. Discussion

Based on our findings, presenting document layouts on 2D tactile displays shows promising potential for offering blind users access to digital documents while preserving the original layout of documents. Participants found the data representation and navigation between modes, guided by the VISM approach that we adopted for the TUI, logical and intuitive. However, there is room for improvement to enhance the user experience and address usability limitations identified in our preliminary study.

### 5.1. Optimizing View Modes for Enhanced User Experience

It was observed that all participants favored the bounding boxes view mode for understanding the document layout and structure. However, two participants noted that the alternate view mode, which featured a Braille letter at the center of each bounding box and horizontal lines indicating the presence of elements, was also beneficial for identifying and detecting the document’s different elements. Based on this feedback and a suggestion from one participant, it may be valuable to combine both modes into a single view. This unified mode would display bounding boxes with Braille letters at the center while enabling users to navigate and interact with elements. Such a design could offer a more efficient and user-friendly experience.

In general, the view modes in our TUI were found to effectively provide participants with information about the structure of the document, including the position and size of various elements. In contrast, participants in our study were unable to access the same information when using a screen reader. This observation aligns with the findings of Li et al. [2], which highlighted that the most significant challenge faced by blind participants in their study was the inability of screen readers to convey spatial relationships between layout elements in documents.

### 5.2. Importance of Interaction Feedback for Improved Navigation and Orientation

In our TUI design, the interactions for navigating between document elements and switching between different view modes were generally sufficient to support user exploration. However, we observed that providing immediate feedback for any interaction that alters the content displayed is crucial for maintaining the orientation of a document. For instance, one participant experienced difficulties with the page orientation, particularly during scrolling between pages of the document. In our interface design, we found it essential to include audio feedback, such as a boundary alert or a subtle sound cue, to notify the user when they reached the end of a page. Without this, users would need to keep one hand on the display at all times to detect changes.

This aligns with the findings of Chase et al. [39], who emphasized the need for tightly coordinated haptic and audio feedback in user interfaces. Their research indicated that haptic guidance cues effectively complemented audio feedback. Our co-designer, during the initial phases of designing our TUI, also underscored the importance of this coordination. This suggests that integrating both feedback types in a harmonized manner is important for enhancing the user experience for blind users.

### 5.3. Additional Features

Participants in our exploratory study provided suggestions for new features in the next versions of the interface. One participant (P3) requested editing capabilities, such as text highlighting and the ability to add comments. Given the positive feedback on the touch interaction concepts, exploring how these editing features could be implemented through touch interactions would be a valuable direction for future development. Additionally, the same participant proposed adding a feature to listen to text word-by-word by placing a finger on the text in the “Details on Demand” view mode. This could further improve interaction and accessibility.

In our interface, we included a help feature that provides users with extra information about the document layout. However, we noticed that the participants did not utilize it during the study. Some explained that they were able to complete tasks without requiring additional assistance. This suggests the need to revise the help functionality’s content and presentation. For instance, incorporating a document summary, as one participant suggested, could enhance this feature’s relevance and usefulness.

While our interface presented graphical information with alternative text when available, a brief placeholder indicated the presence of an image when such text was missing. Although this was not reported as a significant issue during our preliminary study, it may be because blind users are accustomed to similar limitations in screen readers and one-dimensional tactile displays. However, given the capabilities of 2D tactile displays, exploring methods for presenting the original image itself would be worthwhile. This approach could provide users additional context and a richer understanding of the content.

## 6. Limitations

While our approach for presenting documents on 2D tactile displays showed promise, shortcomings and open questions identified in the previous section should be further investigated. Future works should focus on refining our methods and exploring their potential contributions to developing emerging TUIs for blind users. A formal evaluation with a more extensive and diverse group of participants would be necessary to rigorously assess the system’s effectiveness. Moreover, as all participants in our study were Braille readers, it is essential to explore how non-Braille readers interact with the interface.

Additionally, comparing our approach for presenting documents with recent advancements in screen readers would provide valuable insights into the benefits of offering spatial layout information. One key finding from our study was that a participant who had previously been sighted and familiar with document structures showed a strong preference for receiving layout information. This suggests the importance of involving a more comprehensive range of participants, particularly those with varying degrees of familiarity

with document layouts from their past. Investigating how this familiarity influences user preferences could deepen our understanding of how spatial layout information supports individuals with different visual experiences.

## 7. Conclusions

In this paper, we introduced an intelligent TUI that presents documents in a 2D tactile format, allowing blind individuals to explore and interact with complex document layouts using 2D refreshable tactile displays. The data representation concepts were inspired by the VISM by Ben Shneiderman, which provides guidelines for designing effective information visualization systems. The core interaction methods were developed with a blind contributor to ensure accessibility and relevance. Additionally, the interface included a module that automatically extracted layout metadata from complex documents using a YOLOv10 object detection model trained on an augmented dataset with diverse layouts.

Participants in our exploratory study found this document representation approach beneficial for understanding spatial information and quickly skimming through content. They could navigate a complex magazine article structure effectively, which was not feasible when using a traditional screen reader. In our future work, we will aim to address the limitations identified in this study, such as refining the overview modes and enhancing feedback mechanisms for the interaction concepts in our interface. These improvements could help advance future research on designing 2D tactile interfaces and contribute to developing more standardized methods for presenting spatial information to blind users.

**Author Contributions:** Conceptualization, S.A. and O.M.; methodology, S.A. and O.M.; software, S.A. and O.M.; validation, S.A., O.M., K.M. and T.S.; formal analysis, S.A. and O.M.; investigation, S.A. and O.M.; resources, S.A. and O.M.; data curation, S.A. and O.M.; writing—original draft preparation, S.A., O.M., T.S. and K.M.; writing—review and editing, K.M. and T.S.; visualization, S.A.; supervision, B.R. and R.S.; project administration, B.R., K.M., T.S. and R.S.; funding acquisition, B.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the European Research Council (ERC), grant agreement no. 816006, and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie, grant agreement no. 861166.

**Institutional Review Board Statement:** This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee at the Karlsruhe Institute of Technology.

**Informed Consent Statement:** Informed consent for participation was obtained from all subjects involved in this study.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors upon request.

**Acknowledgments:** Special thanks to the ACCESS@KIT Institute for providing the spaces where the user study was conducted.

**Conflicts of Interest:** All authors declare no conflicts of interest. The funding sponsors had no role in the design of this study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

## References

1. Kearney-Volpe, C.; Hurst, A. Accessible web development: Opportunities to improve the education and practice of web development with a screen reader. *TACCESS* **2021**, *14*, 1–32. [\[CrossRef\]](#)
2. Li, J.; Kim, S.; Miele, J.A.; Agrawala, M.; Follmer, S. Editing spatial layouts through tactile templates for people with visual impairments. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–11.
3. Potluri, V.; Grindeland, T.E.; Froehlich, J.E.; Mankoff, J. Examining visual semantic understanding in blind and low-vision technology users. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–14.
4. Bishop, A.P. Document structure and digital libraries: How researchers mobilize information in journal articles. *Inf. Process. Manag.* **1999**, *35*, 255–279. [\[CrossRef\]](#)

5. Dorigo, M.; Harriehausen-Mühlbauer, B.; Stengel, I.; Dowland, P.S. Survey: Improving document accessibility from the blind and visually impaired user's point of view. In Proceedings of the Universal Access in Human-Computer Interaction. Applications and Services: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011, Orlando, FL, USA, 9–14 July 2011; Springer: Berlin/Heidelberg, Germany, 2011; Proceedings, Part IV 6, pp. 129–135.
6. Wang, L.L.; Cachola, I.; Bragg, J.; Cheng, E.Y.Y.; Haupt, C.H.; Latzke, M.; Kuehl, B.; van Zuylen, M.; Wagner, L.M.; Weld, D.S. Improving the Accessibility of Scientific Documents: Current State, User Needs, and a System Solution to Enhance Scientific PDF Accessibility for Blind and Low Vision Users. *arXiv* **2021**, arXiv:2105.00076.
7. Borges Oliveira, D.A.; Viana, M.P. Fast CNN-Based Document Layout Analysis. In Proceedings of the 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy, 22–29 October 2017; IEEE: Piscataway, NJ, USA, 2017. [[CrossRef](#)]
8. Breuel, T.M. Two Geometric Algorithms for Layout Analysis. In *Document Analysis Systems V*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 188–199. [[CrossRef](#)]
9. Li, M.; Xu, Y.; Cui, L.; Huang, S.; Wei, F.; Li, Z.; Zhou, M. DocBank: A Benchmark Dataset for Document Layout Analysis. *arXiv* **2020**, arXiv:2006.01038. [[CrossRef](#)]
10. Wang, J.; Krumdick, M.; Tong, B.; Halim, H.; Sokolov, M.; Barda, V.; Vendryes, D.; Tanner, C. A graphical approach to document layout analysis. In Proceedings of the International Conference on Document Analysis and Recognition, San José, CA, USA, 21–26 August 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 53–69.
11. Binmakhashen, G.M.; Mahmoud, S.A. Document layout analysis: A comprehensive survey. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–36. [[CrossRef](#)]
12. Gemelli, A.; Marinai, S.; Pisaneschi, L.; Santoni, F. Datasets and annotations for layout analysis of scientific articles. *Int. J. Doc. Anal. Recognit. (IJ DAR)* **2024**, *27*, 683–705. [[CrossRef](#)]
13. Pontelli, E.; Gillan, D.; Xiong, W.; Saad, E.; Gupta, G.; Karshmer, A.I. Navigation of HTML Tables, Frames, and XML Fragments. In Proceedings of the Fifth International ACM Conference on Assistive Technologies, Edinburgh, UK, 8–10 July 2002; Assets '02, pp. 25–32.
14. Leporini, B.; Buzzi, M. Visually-Impaired People Studying via eBook: Investigating Current Use and Potential for Improvement. In Proceedings of the 2022 6th International Conference on Education and E-Learning (ICEEL), Yamanashi, Japan, 21–23 November 2022; pp. 288–295.
15. Kim, H.; Smith-Jackson, T.; Kleiner, B. Accessible haptic user interface design approach for users with visual impairments. *Univers. Access Inf. Soc.* **2014**, *13*, 415–437. [[CrossRef](#)]
16. Wang, A.; Chen, H.; Liu, L.; Chen, K.; Lin, Z.; Han, J.; Ding, G. Yolov10: Real-time end-to-end object detection. *arXiv* **2024**, arXiv:2405.14458.
17. Li, J.; Xu, Y.; Lv, T.; Cui, L.; Zhang, C.; Wei, F. Dit: Self-supervised pre-training for document image transformer. In Proceedings of the 30th ACM International Conference on Multimedia, Lisboa, Portugal, 10–14 October 2022; pp. 3530–3539.
18. Chen, Y.; Zhang, J.; Peng, K.; Zheng, J.; Liu, R.; Torr, P.; Stiefelhagen, R. RoDLA: Benchmarking the Robustness of Document Layout Analysis Models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 15556–15566.
19. Huang, Y.; Lv, T.; Cui, L.; Lu, Y.; Wei, F. Layoutlmv3: Pre-training for document ai with unified text and image masking. In Proceedings of the 30th ACM International Conference on Multimedia, Lisboa, Portugal, 10–14 October 2022; pp. 4083–4091.
20. Zhong, X.; Tang, J.; Jimeno Yepes, A. PubLayNet: Largest Dataset Ever for Document Layout Analysis. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 20–25 September 2019; IEEE: Piscataway, NJ, USA, 2019. [[CrossRef](#)]
21. Pfitzmann, B.; Auer, C.; Dolfi, M.; Nassar, A.S.; Staar, P. Doclaynet: A large human-annotated dataset for document-layout segmentation. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 14–18 August 2022; pp. 3743–3751.
22. Jaume, G.; Ekenel, H.K.; Thiran, J.P. Funsd: A dataset for form understanding in noisy scanned documents. In Proceedings of the 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), Sydney, NSW, Australia, 20–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; Volume 2, pp. 1–6.
23. Wang, Z.; Xu, Y.; Cui, L.; Shang, J.; Wei, F. Layoutreader: Pre-training of text and layout for reading order detection. *arXiv* **2021**, arXiv:2108.11591.
24. Zhang, L.; Hu, A.; Xu, H.; Yan, M.; Xu, Y.; Jin, Q.; Zhang, J.; Huang, F. Tinychart: Efficient chart understanding with visual token merging and program-of-thoughts learning. *arXiv* **2024**, arXiv:2404.16635.
25. Moured, O.; Baumgarten-Egemole, M.; Müller, K.; Roitberg, A.; Schwarz, T.; Stiefelhagen, R. Chart4blind: An intelligent interface for chart accessibility conversion. In Proceedings of the 29th International Conference on Intelligent User Interfaces, Greenville, SC, USA, 18–21 March 2024; pp. 504–514.
26. Sechayk, Y.; Shamir, A.; Igarashi, T. SmartLearn: Visual-Temporal Accessibility for Slide-based e-learning Videos. In Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 11–16 May 2024; pp. 1–11.

27. Wang, L.L.; Cachola, I.; Bragg, J.; Cheng, E.Y.Y.; Haupt, C.; Latzke, M.; Kuehl, B.; van Zuylen, M.N.; Wagner, L.; Weld, D. Scia11y: Converting scientific papers to accessible html. In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility, Virtual Event, 18–22 October 2021; pp. 1–4.
28. Stockman, T.; Metatla, O. The influence of screen-readers on web cognition. In Proceedings of the Accessible Design in the Digital World Conference (ADDW 2008), York, UK, January 2008.
29. Pathirana, P.; Silva, A.; Lawrence, T.; Weerasinghe, T.; Abeyweera, R. A Comparative Evaluation of PDF-to-HTML Conversion Tools. In Proceedings of the 2023 International Research Conference on Smart Computing and Systems Engineering (SCSE), Kelaniya, Sri Lanka, 29 June 2023; IEEE: Piscataway, NJ, USA, 2023; Volume 6, pp. 1–7.
30. Morris, M.R.; Johnson, J.; Bennett, C.L.; Cutrell, E. Rich representations of visual content for screen reader users. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–11.
31. Wang, L.L.; Bragg, J.; Weld, D.S. Paper to HTML: A Publicly Available Web Tool for Converting Scientific Pdfs into Accessible HTML. *ACM SIGACCESS Access. Comput.* **2023**, 1–11. [[CrossRef](#)]
32. Peng, Y.H.; Chi, P.; Kannan, A.; Morris, M.R.; Essa, I. Slide Gestalt: Automatic Structure Extraction in Slide Decks for Non-Visual Access. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; pp. 1–14.
33. Khurana, R.; McIsaac, D.; Lockerman, E.; Mankoff, J. Nonvisual interaction techniques at the keyboard surface. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–12.
34. Gadde, P.; Bolchini, D. From screen reading to aural glancing: Towards instant access to key page sections. In Proceedings of the 16th international ACM SIGACCESS Conference on Computers & Accessibility, Rochester, NY, USA, 20–22 October 2014; pp. 67–74.
35. Vtyurina, A.; Fournery, A.; Morris, M.R.; Findlater, L.; White, R.W. Verse: Bridging screen readers and voice assistants for enhanced eyes-free web search. In Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility, Pittsburgh, PA, USA, 28–30 October 2019; pp. 414–426.
36. Ahmed, F.; Borodin, Y.; Soviak, A.; Islam, M.; Ramakrishnan, I.; Hedgpeth, T. Accessible skimming: Faster screen reading of web pages. In Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, Cambridge, MA, USA, 7–10 October 2012 pp. 367–378.
37. Safi, W.; Maurel, F.; Routoure, J.M.; Beust, P.; Dias, G. Blind browsing on hand-held devices: Touching the web... to understand it better. In Proceedings of the Data Visualization Workshop (DataWiz 2014) associated to 25th ACM Conference on Hypertext and Social Media (HYPERTEXT 2014), Poznan, Poland, 10–13 September 2014.
38. Maurel, F.; Dias, G.; Routoure, J.M.; Vautier, M.; Beust, P.; Molina, M.; Sann, C. Haptic Perception of Document Structure for Visually Impaired People on Handled Devices. *Procedia Comput. Sci.* **2012**, *14*, 319–329. [[CrossRef](#)]
39. Chase, E.D.; Siu, A.F.; Boadi-Agyemang, A.; Kim, G.S.; Gonzalez, E.J.; Follmer, S. PantoGuide: A Haptic and Audio Guidance System To Support Tactile Graphics Exploration. In Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility, Virtual Event, 26–28 October 2020; pp. 1–4.
40. Maćkowski, M.; Brzoza, P. Accessible tutoring platform using audio-tactile graphics adapted for visually impaired people. *Sensors* **2022**, *22*, 8753. [[CrossRef](#)] [[PubMed](#)]
41. Chang, R.C.; Yong, S.; Liao, F.Y.; Tsao, C.A.; Chen, B.Y. Understanding (Non-) Visual Needs for the Design of Laser-Cut Models. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; pp. 1–20.
42. Prescher, D.; Weber, G.; Spindler, M. A Tactile Windowing System for Blind Users. In Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility, New York, NY, USA, 25–27 October 2010; ASSETS '10, pp. 91–98.
43. metec, A. The “Laptop for the Blind”. Available online: <https://metec-ag.de/> (accessed on 1 October 2024).
44. Artistic Document Layout Dataset (ArtDocLay). Available online: <https://github.com/moured/layout-for-all> (accessed on 1 October 2024).
45. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. In Proceedings of the 1996 IEEE Symposium on Visual Languages, Boulder, CO, USA, 3–6 September 1996; pp. 336–343.
46. Zhao, H.; Plaisant, C.; Shneiderman, B.; Duraiswami, R. Sonification of Geo-Referenced Data for Auditory Information Seeking: Design Principle and Pilot Study. In Proceedings of the International Conference on Auditory Display, Sydney, NSW, Australia, 6–9 July 2004.
47. Russell, D.M. What Are You Reading? *Interactions* **2023**, *30*, 10–11. [[CrossRef](#)]
48. Shein, E. Shining a Light on the Dark Web. *Commun. ACM* **2023**, *66*, 13–14. [[CrossRef](#)]
49. NV Access Limited. NVDA Screen Reader 2024. Available online: <https://www.nvaccess.org/download/> (accessed on 1 October 2024).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.