

Received 26 October 2022, accepted 10 November 2022, date of publication 17 November 2022,
date of current version 23 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3223099

APPLIED RESEARCH

Tactile Audio Responsive Intelligent System

MARALBEK ZEINULLIN¹ AND MARION HERSH¹

Department of Biomedical Engineering, University of Glasgow, G12 8QQ Glasgow, U.K.

Corresponding author: Maralbek Zeinullin (2420412z@student.gla.ac.uk)

This work was supported by the Center for International Programs “Bolashak.”

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of the College of Science and Engineering, University of Glasgow, under Application No. 300200167.

ABSTRACT For people with visual impairments, information encoded in a visual format creates certain barriers. To alleviate this, a large volume of research has been conducted in the field of assistive technology. In our work, we developed a special system that makes educational materials more accessible. The system consists of three components: the pre-labelled tactile graphics, an interactive labelling web tool and the phone application. Tactile graphics are used at schools for the blind and allow the students to understand non-textual information by touch. The digital version of the graphics first needs to be labelled by teachers using the developed web tool. Then, the phone app, which is based on the Android platform, will accompany those graphics with the audio descriptions. The fundamental purpose of the developed app is to allow the user to gain information without sighted assistance. We also conducted a study to evaluate the system. First, the structured interview was carried out to gather data about the participant's experience with the tactile graphics and mobile devices. Next, quantitative measurements were obtained through a series of experiments. Subsequently, a post-experimental session was carried out to record the participants' thoughts and opinions about the system. The results of the experiments demonstrated that the proposed mobile application allows the users to explore the graphics more efficiently.

INDEX TERMS Assistive technology, audio feedback, computer vision, fingertip detection, tactile graphics, visually impaired, deep learning, YOLOv3.

I. INTRODUCTION

Information exchange presents significant challenges to visually impaired people (VIP), particularly as most information is in a visual format and insufficient attention is given to accessibility. This issue is particularly important in school and post-school education, where information is often presented in the form of shapes, diagrams, maps, schemes and photos. The ability to understand the information presented in these two-dimensional formats is called graphical literacy. Several studies [1], [2], [3] have concluded that this enables individuals to communicate and obtain knowledge more effectively.

Graphics have three main advantages. They are concise, relatively easily memorable and can clearly represent relationships between data. A well drawn and labelled image

The associate editor coordinating the review of this manuscript and approving it for publication was Mehul S. Raval¹.

can present detailed information which does not require much more than a glance to understand. Several studies have demonstrated the ability to remember visuals [4], [5]. Charts and graphs can be used to represent the links between complicated data in an easy-to-understand format [6]. In summary, graphics can enable the assessment and understanding of a large amount of information relatively quickly and comprehensively. This, of course, holds for well-drawn graphics with appropriate text labels. Poorly drawn graphics like badly written text convey minimal information and may lead to confusion and misunderstandings.

Assistive technology systems have been defined as “equipment, devices and systems that can be used to overcome the social, infrastructure and other barriers experienced by disabled people that prevent their full and equal participation in all aspects of society” [7]. Tactile Graphics (TG) can be used as a form of assistive technology which enables visually impaired people to understand non-textual information by

touch. They consist of raised line versions of graphics which can be explored by touch. However, the original visual image needs to be adapted to produce a good TG, for instance by reducing the amount of information to avoid difficulties in reading and understanding the image. There are no strict rules for TG navigation, but a systematic approach is recommended [8] and similar techniques can be used to explore a TG to those used for reading Braille. The user can scan the tactile image with both hands or use one hand as an anchor and actively explore with the other. TG are particularly important in the fields of science, technology, engineering and mathematics (STEM) where data is frequently represented as diagrams, charts, figures and spatial maps.

As indicated above graphics require text labels. Adding Braille labels to TG would make the graphic very cluttered and difficult to read. The solution to this is the use of audio descriptions. The paper presents a phone app for doing this. It uses an advanced computer vision algorithm to track the user's fingers and lets them know in real-time what they are touching as they explore a TG. The paper also presents the methodology and results of experiments carried out to test the hypothesis that the proposed app allows VIP to develop better graphical literacy skills in an independent manner and makes spatial content more accessible.

II. RELATED WORK

A. THE BENEFITS OF TACTILE GRAPHICS

A number of studies have investigated users' experiences of TG. For instance, 52% of 59 visually impaired students were found to agree (or strongly agree) that they liked using TG and wanted more access to them [9]. 47% of them agreed that TG helped them to keep up with their sighted peers and to feel connected with the flow of class teaching. Authors of another study [11] asked users about their experiences with TG exploration and found that 74% of 76 participants expressed positive (medium to very high) attitudes. One of the limitations of this research is the lack of information about participants' backgrounds, including whether they attended a school for the blind or a mainstream school.

Participants in all these surveys commented on the importance of providing accompanying descriptions in Braille, audio or another accessible format. The interview results [12] support the idea that TG require additional textual information and are not sufficient on their own, especially in secondary schools.

Several studies have obtained information about the usefulness of TG from teachers and instructors. A recent study [34] of 241 teachers of visually impaired students in Canada and the USA found that almost all (98%) of them agreed (or strongly agreed) that exposure to TG at an early age was crucial for the student's further academic success. In another study [1] ($n = 306$), the authors found that visually impaired students who started using graphics (including tactile ones) in the early grades were more successful in the

later academic curriculum. One of the teachers in a recent study [2] reported that students with good graphical literacy skills were better at generalising and structuring information. The 24 teachers working with visually impaired students were interviewed in a longitudinal study [12] and all reported that there were situations where TG significantly contributed to effective learning. The main limitations of these studies are the lack of information about how long the teachers have been working with visually impaired people and how many visually impaired students they teach.

In summary, graphical literacy is a very important skill for visually impaired (and other) students which can facilitate obtaining information and learning during the education process. The results of the studies discussed above show that TG accompanied by accessible labels or descriptions are an effective learning material and that early exposure to them can contribute to academic success later on.

B. EDUCATIONAL SYSTEMS FOR VISUALLY IMPAIRED PEOPLE

The era of audio-tactile assistive technology started with Nomad [13]. The device was connected to a computer and a touch-sensitive surface was used to trigger corresponding pre-defined audio descriptions. The system was commercialised the following year, but was not very popular with users. This was largely due to technology limitations at the time rather than the device design [15]. In particular, screen resolution was poor and the available synthetic voices did not sound very natural. Several audio-tactile solutions were developed subsequently [10], [15], [16], [17], [18], [19], [20], [21], [22], [23] and are summarised in Table 1. The discussion in this section focuses on the approaches which used mobile devices and are therefore related to our system.

1) TPad

[25] is a mobile tablet-based (iPad Pro) educational application which allows users to explore tactile images. Users explore an embossed tactile image placed over a tablet screen with their hands. The system provides clarifying audio feedback when it detects the user touching the image. A 3D-printed plastic frame is used to hold the A4 paper in place and stop it moving. The TPad uses Scalable Vector Graphics (SVG) files with the information about the included images and objects to provide the audio descriptions. Different modes enable users to download pre-processed SVG files easily, for instance by scanning QR-codes located on the other side of the TG. A web interface has been developed to allow instructors to upload and organise graphics and send them to the server to which all TPods are connected. A comparative study showed that users obtained information about the TG faster and gave 70% more correct answers when using the Tpad compared to obtaining additional information from Braille text or a screen reader. The drawbacks of the system are potential difficulties with 3D printing the frame and embossing the tactile graphics since some educational organisations do not have 3D printers and Braille embossers.

TABLE 1. Educational systems for visually impaired.

Name	Hardware	Software	Tactile Output	Interaction Modes	Limitations
Touch screen based					
Nomad [13]	Electronic touch-pad connected to a computer	Nomad Kernel, CAD, Information Access System, Walkabout System	Swell paper	Direct touch	High cost, not portable
TTT (Talking Tactile Tablet) [15]	Electronic tablet connected to a computer	World Map, TTT Match Game, TTT Snakes & Ladders, TTT Tool	Swell paper	Direct touch	High cost, not portable
IVEO by View-Plus [16]	Electronic touch-pad	IVEO Player Pro	Braille paper	Direct touch	High cost
Refreshable Tactile Graphics [17]	STRESS Refreshable tactile display [24]	Xenomai, MaskGen, Photoshop	Refreshable tactile display	Direct touch	High cost
IMG (Interactive Multimodal Guide) [18]	Enclosure display, Arduino, proximity touch sensor	ZW3D 3D drawing software	3D printed model	Direct touch	High cost
Computer and web camera based					
The Tactile Graphics Helper (TGH) [10]	Laptop and webcam	Matlab	Swell paper	Finger pointing and voice commands	Not portable
Access Lens (AL) [19]	Laptop and webcam	OpenCV, Microsoft NET speech library, ABBY Fine Reader	Regular paper	Finger pointing, virtual edge menus and voice commands	Not portable
Tactile Graphics Reading Assistive Device [20]	Laptop and webcam	MediaPipe [30]	Swell paper	Finger pointing	Not portable
Depth camera based					
CamIO (Camera Input-Output) [21]	Microsoft Kinect Camera	Computer vision algorithms	3D-printed model	Direct touch	High cost
IAG (Interactive Audio Guide) [22]	Intel RealSense F200	Computer vision algorithms	3D relief surface	Finger pointing	Not portable
Markit and Talkit [23]	RGBD camera	OpenCV	3D-printed model	Direct touch	Not portable
Mobile phone and tablet camera based					
Tactile Graphics with a Voice [26]	iPod touch and iPhone	QR code reader	Braille paper	Finger pointing and voice guidance	Aiming the phone camera
TPad system [25]	Tablet	non-CV	Regular paper	Direct touch	High cost of 3D printing and Braille embosser
Touch and Hear Assistive Teaching System (THATS)	iPhone	Computer vision algorithm	Swell paper	Finger pointing	Aiming the phone camera
Molder [27]	IPad, 3D printer	OpenCV, SpeechRecognizer	3D printed object	Finger pointing and voice commands	Single finger exploration, High cost of 3D printing
TARS [28]	iPhone	MediaPipe [30], AVSpeechSynthesizer	Swell paper	Finger pointing	Slow fingertip detection time
TAURIS	Android Phone	Tiny YOLOv3, OpenCV	Swell paper	Finger pointing	Aiming the phone camera

2) TACTILE GRAPHICS WITH A VOICE

[26] is an application which runs on a smartphone and provides feedback by scanning QR codes placed on the TG. The main motivation for its development was to replace large Braille texts with more compact codes; thus, providing more information about TG in a smaller space. There are three different modes to help the user aim the phone camera: silent, audio instructions and finger pointing. Instructions navigate the user to a QR code by sensing the phone orientation, whereas the finger pointing mode helps an app to identify the correct QR code when multiple labels are visible. Overall, visually impaired users were satisfied with the system but found aiming a smartphone camera challenging.

3) THATS¹ (TOUCH AND HEAR ASSISTIVE TEACHING SYSTEM)

is a mobile app that allows the user to explore pre-defined tactile images by providing accompanying audio descriptions. As well as the app, the THATS team developed an online editor and linked it to a digital library. This enables instructors to both create tactile graphics and download ready-to-use ones. THATS implements widely used computer vision algorithms (background subtraction, image thresholding, etc) to detect the user's fingertip, in contrast to the deep learning approach used in our work. One of the drawbacks of this system is that it is only able to detect a single finger pointing action. However, studies show [32], [33] that it is easier to recognise tactile images using both than one hand. In summary, THATS is a very promising project which was developed to give visually impaired people free access to easy-to-use educational materials, but the app has not yet been launched and there are no publicly available experimental results of its use.

Researchers from Cornell University created a **Molder** [27] - an accessible tool for tactile maps design and exploration. This tool has four main components: a physical frame, a website for the model's creation, a mobile app for the model exploration and a server. The Molder was tested by the end users and results showed that participants with different vision abilities were able to create tactile models using this tool. The main disadvantage of the Molder is that it only supports a single-finger exploration. In addition, production of the 3D models is time-consuming and has a high cost.

4) TARS

[28] is another mobile application which provides audio descriptions whilst the user explores a tactile image. This app utilises Google's MediaPipe [30] hand-tracking system for fingertip detection. According to their experimental results, the app detects fingertips with 85.5% accuracy. However, it was not clearly stated whether this result is for all five fingertips or just the index one. In addition, a single-frame processing time was not specified. In the further work conducted by the same authors [29], it took two seconds to

process a single frame. Whereas real-time execution expects at least 15 fps (frames per second) [31]. Another limitation of the study is that the authors have not described how the tactile images and corresponding annotations are created.

In conclusion, technology development has significantly increased the computational power of edge devices and several research areas have shifted their focus from cumbersome systems that use relatively large PCs and cameras to small mobile devices with integrated computing and cameras. We have described five assistive technology solutions designed to make TG more accessible. TPad, Molder and the Tactile Graphics with a Voice devices have already been tested and shown promising results. However, there are still some drawbacks, i.e. the high price of the 3D printers or failure to read the QR code when it is obscured by the user's finger. There are indications that work on the THATS device may have been discontinued and TARS is not able to process in real-time on a mobile device yet. We have developed a new system which uses state-of-the-art technology to address the limitations of existing solutions.

III. TACTILE AUDIO RESPONSIVE INTELLIGENT SYSTEM (TAURIS)

A. MOBILE APPLICATION

The main aim of the application is to enable visually impaired users to explore TG without sighted assistance. The app tracks the positions of the users' fingers as they explore the image and provides information about what they are touching. The user holds the phone in one hand and uses the other to explore the tactile image. A phone holder could be used to enable the user to explore the TG with both hands.

The device uses swell paper with a tactile image on one side and a QR code on the other. The QR code has to be scanned to transfer the information about the TG to the device memory. After users receive a notification that the code has been successfully scanned, they can turn over the TG and start exploring it. First, the app will give brief information about the graphic and its main features. Then it will run a computer vision (CV) algorithm that detects fingertips and triggers predefined audio feedback. The algorithm used in this process was developed by the researchers. So-called Aruco markers are placed at the corners of the TG to enable the phone camera to capture the whole image. This image is divided into 2400 (60 × 40) cells and each cell has a pre-defined piece of information associated with it. Simultaneously, the CV model detects all visible fingers and outputs the position of the default finger, which has been initially set as the index finger, but could be changed using the app settings. Finally, the fingertip location is mapped onto the image cells and the corresponding data is read out to the user. It is important to note that it uses both square QR codes and square Aruco markers, but they have different functions. A QR code is used for downloading data from the server and a set of Aruco markers to enable the camera to determine the image location.

¹<https://thats.wiki.procvic.ro/en/>



FIGURE 1. Phone mounted on a phone holder.

The app is capable of processing ~ 15 frames per second but it will not trigger feedback initially. It will give audio output initially and then wait for the user to move their finger. If the location does not change significantly or remains within the single TG object it will remain silent. The developed CV model is based on Deep Neural Network (DNN) architecture. A major advantage of this approach is the high detection rate of each finger. Therefore, the user can explore an image with the whole hand and the algorithm will return the position of only the desired finger. In other studies [10], [22], [26], OrCam,² it is required to point with the index finger only, which, as it was mentioned before, is inconvenient.

As was mentioned above, Aruco markers [35] are an essential component of TAURIS. Four markers are located at each corner of the TG and function as reference points. At least three markers have to be visible for the system to operate properly. The app will notify the user if it is unable to detect at least three markers. The underlying algorithm, reference points and the DNN model are described in the next section. If the user prefers to explore the image with both hands [2], the phone can be mounted on a holder (Figure 1). The algorithm is able to distinguish the fingers of both right and left hands. For instance, when the user uses the index finger of the right hand, the system will trigger output for the objects touching that finger only. The app will save information about the current session if it is paused.

Users can request different types of descriptions depending on type of information they need. The available descriptions are: overview, basic, and detailed.

The **overview** is activated after the user scans the QR code and turns over the TG. If the user would like to hear this information again, they need to turn over the graphics in both directions again.

As soon as the user begins exploring the TG and their fingers are detected by the algorithm, the app provides a **basic** description. In this mode the app tells them which object their index finger is touching without providing any additional details.

If the user needs more thorough information, they can activate the **detailed** description mode. This can be achieved by holding the index finger still for three seconds above the target object.

To clarify the differences between what type of information is presented in each mode, the examples of descriptions for an Australian map are provided below.

Overview: This is the map of Australia. It consists of six states and two territories.

Basic: New South Wales

Detailed: Capital city is Sydney. The Australian Capital territory is located in this State as well.

The quality of the descriptions depends on the information provided. Descriptions should be concise and relevant. Long texts can be tiring and irrelevant information may confuse the user and also wastes their time. Also, it is essential that the most important information is provided. “How to Write Alt Text and Image Descriptions for the Visually Impaired”³ is a useful guide created by the Perkins School for the Blind.

B. COMPUTER VISION ALGORITHMS

1) TINY-YOLOv3

YOLOv3 is based on the YOLO [39] algorithm. YOLO's main advantage is the ability to process images in real-time. This is achieved by the unified architecture of its DNN. **Tiny-YOLOv3** is a compact and accelerated version of YOLOv3 which was designed for embedded and mobile systems. The smaller architecture size makes the Tiny-YOLOv3 very fast. YOLOv3 and its tiny version have 24 and 13 convolutional layers respectively. As expected, a higher detection speed comes at the price of lower accuracy. It has been reported [39] that the smaller version works 3.5 times as fast with just 10% accuracy trade off. After experimenting with different YOLO and TensorFlow models, only the tiny version of the YOLOv3 was capable of detecting objects in nearly real-time on a mobile device. In our project, the SCUT-Ego-Gesture Dataset was used for model development [40]. The dataset contains 59,111 labelled images of sixteen different hand gesture types.

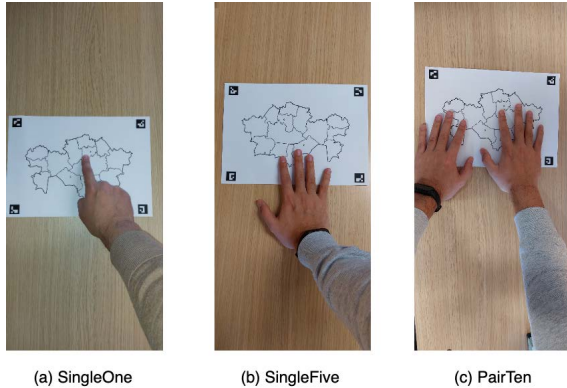
The model was trained using Google Colab free cloud service. First, we focused on the SingleOne gesture (Figure 2.a). Unfortunately, with this approach the algorithm often confused the index finger with other fingers due to their very similar appearances. This meant that the index finger could only be detected reliably when only one finger was visible in the frame. As there are few advantages in exploring a tactile image with one finger [41], another model was trained using the SingleFive set (Figure 2.b). The resulting model was able to accurately detect the fingertips when only one hand was in the field of view, but performed poorly when both hands were visible. So, it was decided to train another model using the PairTen set (Figure 2.c). Unlike the previous

²<http://www.orcam.com>

³<https://www.perkinslearning.org/technology/blog/how-write-alt-text-and-image-descriptions-visually-impaired>

TABLE 2. Fingertip detection accuracy under different lighting conditions.

Gamma value	Thumb (%)	Index (%)	Middle (%)	Ring (%)	Little (%)	Average (%)
Original (~600 lux)	84.1	92.4	79.4	99.1	85.1	88.0
80% (~500 lux)	84.6	89.7	74.9	98.9	87.0	87.0
60% (~380 lux)	78.2	78.3	68.4	87.3	81.9	78.8
40% (~220 lux)	61.2	55.5	64.6	64.9	83.2	65.9
20% (~40 lux)	0.1	0.1	3.5	1.4	4.5	1.9

**FIGURE 2.** Dataset images.

model, it detected fingertips correctly when both palms were visible and failed when only one was present.

We wanted the app to be flexible and convenient for users and give them the option to use one or both hands. Therefore, it was decided to merge the two image sets and train a third model. This model successfully detected fingertips in both scenarios (when either one or two hands were visible). Also, the model was able to differentiate between fingers on the left and right hands. So, if we program our app to detect the right index finger, it will detect the right-hand finger only. This model was used by the app during the experimental sessions. However, it should be noted that the algorithm requires at least one hand to be visible and does not allow exploration with a single finger. The effectiveness of the detection model can be improved by increasing both the number of dataset images and training epochs. Our model was trained on a dataset of 7500 images with 100 training epochs.

2) EVALUATION OF THE DETECTION MODEL

Evaluation is an essential part of the ML model development. This process is conducted using the test set. In order to obtain unbiased results, it is important that the images in the test set are different from those in the training one. To create a test set which is close to the real-life setting, the researcher recorded himself while exploring the TG both with one and two hands. In total, 200 images were collected and annotated. Results presented in Table 2 are obtained by running the model through this test dataset.

In the same table, you can find the results for the model performance under different lighting conditions. An ideal

lighting level requirement for schools in the UK is 500⁴ lux and higher, while 300 lux is considered to be an acceptable illuminance. Images for our test set were collected under 600 lux illumination. Afterwards, the gamma correction method was used to control the brightness and produce synthetic images (Figure 3).

$$O_{image} = 255 \times \frac{I_{image}}{255}^{\frac{1}{\gamma}}$$

where:

O_{image} = output pixel value [0, 255]

I_{image} = input pixel value [0, 255]

γ = gamma value

According to the results, we can conclude that the model operates adequately even under the illumination conditions which are below the norm (55.5% accurate for the index finger when the brightness is around 220 lux). While CV algorithms which are based on skin colour and coloured sticker detection do not demonstrate a robust performance. This is due to the fact that skin tone changes under different lighting conditions [42], [43]. We could not find any literature which analyses fingertip detection accuracies under different lighting conditions. Thus, we believe that the results of our work may help other researchers in the field.

3) ARUCO

Aruco is an open source library which is used to detect square markers (Figure 4). Those are the markers consisting of a wide black border which makes it easy to detect the marker in an image, and a binary matrix inside it which is used to determine its ID. This gives fast and robust marker detection. Experimental results show that the Aruco algorithm works faster than other marker detectors while maintaining accuracy [37]. If the phone camera is calibrated, the Aruco system can be used to estimate the device pose. After successful marker detection, an algorithm returns a marker ID number with the four corner coordinates. The library is based on OpenCV [38] and its code is written in C++. In the context of our system, we use markers with IDs: 0, 1, 2 and 3 which are placed in a clockwise order starting

⁴<https://www.lyco.co.uk/advice/lighting-for-schools-colleges-and-universities>

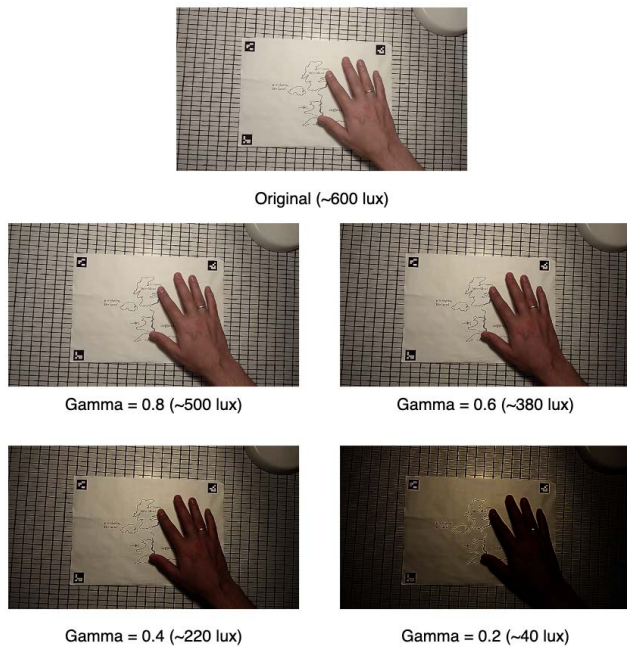


FIGURE 3. Image under different lighting conditions.

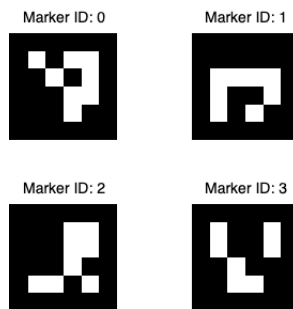


FIGURE 4. Aruco markers.

from the top left. Since we know which ID corresponds to each of the corners, the app can easily identify when one of the corners is not visible and then estimate its location by utilising the information about the remaining three points.

The phone notifies the user if during the exploration process more than two markers are not visible by making short vibrations. This helps the user to readjust the angle and position of the phone and make sure that the whole image is in the field of camera view. We also plan to add the feature which will assist the user to properly aim the camera. As it was mentioned above, the app knows the marker with which ID is missing. Thus, it can explicitly navigate the user and say that “top left” or “left side” (if two markers are not visible) is not in the field of view. We used markers with the same IDs for all the tactile images in our research.

Afterwards, these four points are used to build a top-down view of the image. The OpenCV library has a specific function that calculates a transform matrix and then constructs a warped (bird’s-eye view) image. The marker ordering needs to be consistent. If markers with different IDs are used or their order is changed the app will be unable to apply the transform

matrix algorithm. These steps allow the app to accurately detect the user’s fingertip position with respect to the tactile image even if the phone orientation is continuously changing. Figure 5 demonstrates the app working process.

C. TG ANNOTATION WEB TOOL

As well as the phone application, an online tool to add information about the tactile images was developed. The website interface is presented in Figure 6. The current version of this tool is designed for sighted users (sighted parents, teachers, instructors, etc.) and this is clearly a disadvantage which will need to be corrected in subsequent work. There are three stages. The user first uploads a tactile image in one of the two most common image formats (JPG or PNG). If they do not have their own images, they can obtain suitable images from an online library of tactile graphics. The list of online repositories can be found on the Harpo organisation website.⁵ The second stage is highlighting and annotating regions of interest (ROI). The user can deselect the highlighted region if it was selected by mistake. After marking all ROIs, the user presses the “finish” button to save all information to the server. Simultaneously, a PDF version of the tactile image with the frame of Aruco markers will be created. Also, a QR code will be automatically generated and placed on the second page of the PDF file. Lastly, the final document can be downloaded to local memory and printed (double-sided mode).

In conclusion, our approach differs from the existing solutions in several ways. First, by being capable of detecting all ten fingers separately, it enables a two-handed exploration for the users. Second, Aruco markers placed in the corners make the process of camera aiming more accessible. Third, a compact architecture of the neural network allows real-time execution even on a mid-range mobile device. Table 3 compares these features in the existing devices.

IV. EXPERIMENTAL STUDY

A. METHODOLOGY

The main aim of the experiments was to test the usability of the developed app with end-users. We also evaluated the effectiveness of the app in terms of accuracy and time compared to using Braille text and a screen reader. The research questions we investigated were as follows:

- 1) *Are there statistically significant differences in the **time spent** and **accuracy** achieved when using the TAURIS App compared to each of Braille text and a screen reader to provide additional information to support TG exploration?*
- 2) *Do the users **remember more information** after using the TAURIS compared to each of Braille text and a screen reader to provide additional information to support TG exploration?*
- 3) *Do the users **prefer the TAURIS App** to Braille text or a screen reader?*

⁵<https://piaf-tactile.com/tactile-graphic-resources/>

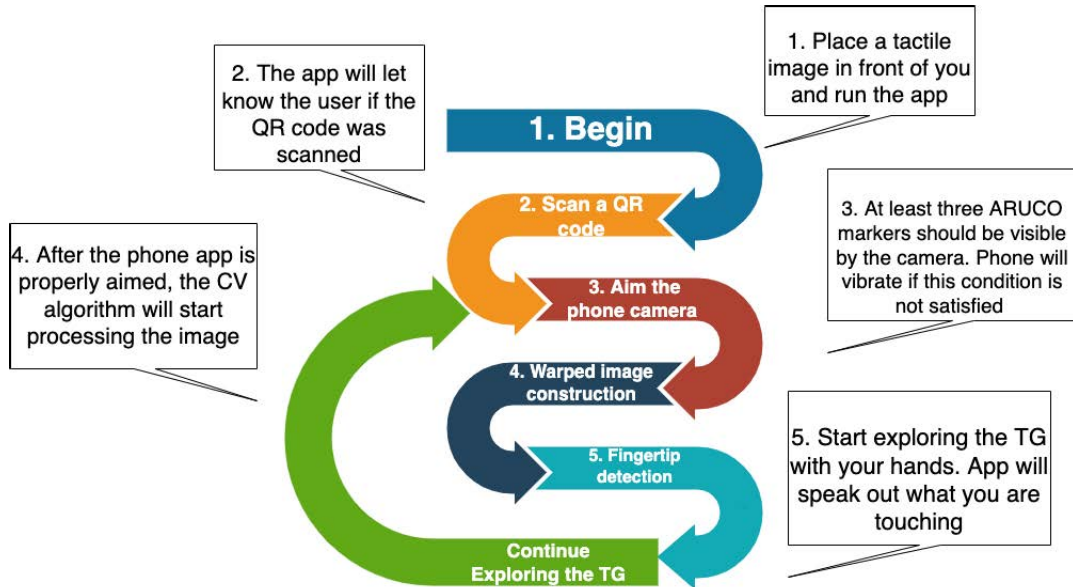


FIGURE 5. Flowchart of the app working process.

TABLE 3. Comparison of fingertip detection methods.

Device	Year	Fingertip detection method	Two-handed exploration	Mobile device	Real-time	Detection in low-light conditions
TAURIS	2022	Tiny-YOLOv3	✓	✓	✓	✓
THATS	2020	Not stated	x	✓	✓	Not stated
TARS	2020	MediaPipe	✓	✓	x	✓
Tactile Graphics with a Voice	2016	Colour-based skin detection	x	✓	✓	x
Tactile Graphics Helper	2015	Hand segmentation	x	x	✓	x
Access Lens	2013	Colour-based skin detection	x	x	✓	x
Tactile Graphics Reading Assistive Device	2022	MediaPipe	✓	x	✓	✓

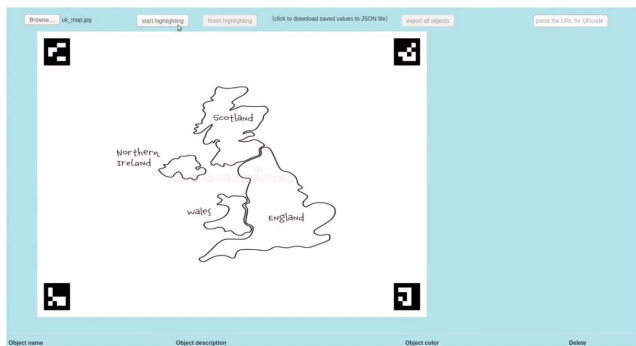


FIGURE 6. TAURIS web tool.

Answering these research questions required sets of both quantitative and qualitative data to be collected. Therefore, it was decided to utilise a mixed-method experimental

research design. This has the major advantage of enabling us to find out not only if the proposed system works better but also to understand how the participants feel about using it.

To achieve this our mixed-methods experimental study was divided into three phases. First, structured interviews were conducted to obtain data about participant experiences with tactile graphics and mobile devices. Then, a series of experiments was carried out to obtain quantitative measurements. One of the main challenges was to identify the right statistical method for the data analysis. After performing a series of Shapiro-Wilk tests [44], it was verified that most of the data were not normally distributed. Therefore, non-parametric statistics were used. Particularly, Wilcoxon signed-rank test was used for the within-subjects design (to test the significance of the interaction modes). Whereas, the Mann-Whitney U test was utilised for the between-subjects designs (to test the effect of age and vision loss). Finally, after

TABLE 4. Participants information.

ID	Gender	Age	Occupation	Vision	Braille literate	School attended
2	Female	42	Library worker	Blind	Yes	Mainstream school
3	Male	36	Library worker	Visually impaired	No	Mainstream school
4	Female	28	University student	Visually impaired	No	Both
5	Female	35	Library worker	Blind	Yes	Mainstream school
6	Female	45	Library worker	Visually impaired	Yes	School for the blind
7	Male	28	School worker	Blind	Yes	School for the blind
8	Male	35	Masseur	Blind	No	School for the blind
9	Male	18	School student	Visually impaired	Yes	School for the blind
10	Male	18	School student	Visually impaired	Yes	School for the blind
11	Male	37	Library worker	Visually impaired	Yes	Mainstream school
12	Female	65	Library worker	Blind	Yes	School for the blind

the experiments we recorded participants' impressions and comments as well as suggestions for improving the system.

B. PILOT STUDY

A pilot study was conducted with a single participant before carrying out the experiments to check whether all the materials were accessible and the experimental design feasible. As a result three changes were made to the experimental protocol and the investigator obtained a good idea of session duration which facilitated scheduling subsequent sessions.

The main change was giving the participants time to familiarise themselves with the descriptions first and only subsequently asking the question and starting to record the time for all three exploration modes. This was to avoid the response time for the first question being much longer than that for the remaining ones and the results not being normally distributed, as the pilot participant was found to require a significant amount of time to read the general TG descriptions in Braille.

The second change was replacing the device's default Samsung text-to-speech synthesizer (TTS) by the Google TTS, as the results showed that most of the participants found it more pleasant to listen to. The final change was replacing one of the Braille embossed texts, which was not fully readable.

C. PARTICIPANTS

Initially, it was planned to carry out the experiments in a Glasgow school. The first author contacted several schools, including a local school for children and young people with sensory impairments and complex learning needs, but COVID-19 restrictions at the time meant that the schools advised him to contact them again when lockdown measures had eased. Therefore, it was decided to gather data in the researcher's home country, Kazakhstan, where educational organisations were operating without major restrictions. Shymkent regional boarding school for visually impaired children agreed to take part in the study and permitted the

investigator to conduct the experiments if all health and safety measures were taken.

In total, there were 11 study participants plus the pilot study participant who was not included in the analysis. Due to the limited number of secondary school students available during the summer holidays, it was decided to involve the school staff and alumni as well. For that reason, the participants' age ranged from 18 to 65 years (average = 35.18 years, SD = 13.15).

Two participants currently attended the school, four had attended it in the past, four participants attended the mainstream school and one participant attended both schools. Information about their current occupations was collected as well. Two participants were secondary school students, and one was a university student. Six participants were employed at the school library, one was a teacher at the school for the blind and one was working as a masseur.

Five participants identified themselves as blind and six had low vision. Three participants had been visually impaired from birth, four lost their vision in childhood and the remaining four after the age of eleven. Eight participants were Braille literate and three said that they understand some Braille but are not fluent. Table 4 summarises the participant information.

D. PROCEDURE

1) PHASE 1 (INTERVIEW)

Informed consent was obtained before the start of the study. Digital versions of the information sheet and consent forms were sent to the participants in advance. We found it interesting that some participants used government issued rubber stamps with their signatures to sign the documents. The interviews were divided into two main parts. The first covered general demographic information and the second participants' experiences with the TG. The final question asked about their familiarity with assistive technology applications for smartphones. The questions used are listed in Appendix A.

2) PHASE 2 (DEVICE TESTING)

3) APPARATUS

The TAURIS App was installed and tested on a Samsung Galaxy A52 device running the Android 11 operating system. This mobile device was used in the experiments with all the participants. The device’s capable camera, decent chipset and most importantly a moderate price (under £300⁶) were the main criteria for the researchers. It was crucial to test the application on a mobile phone which is affordable to all potential users.

The TG were printed on A4 ZYTEX2 swell paper. Each TG had two versions: one with Braille labels and one without. The Braille labelled version was used by the participants who chose Braille text and the screen reader modes for their sessions. The version without labels was used for the TAURIS App experimental mode. The quality of the TG and the Braille texts were assessed by the library staff before production to ensure that the printed graphics were readable. All the experimental sessions were conducted under the same conditions. Participants were given up to ten minutes to familiarise themselves with the app before starting the actual testing session.

4) EXPERIMENTS

During the experimental session, each participant explored six different tactile images. Three tactile images were explored using the app and the remaining three with their choice of mode (Braille text or screen reader). Seven participants chose Braille and four a screen reader. The tactile graphics were divided into three categories: object, map and graph, with each category having a different type of associated graphics. Examples of the TG used in the experiments are illustrated in Figure 9. Brief descriptions of the modes and the object and graphic types are provided below.

a: EXPLORATION MODES

(1) *App mode.* The TAURIS app was used by the participants in this mode. Before each test, we asked participants whether they prefer to explore TG with one or both hands. All of them preferred to use both hands to explore the images. Thus, the phone was mounted on a holder for all sessions. The app used a phone camera to track the user’s fingertip locations in order to trigger the corresponding audio description. Figure 7 shows an example of TG used in the app exploration mode.

(2) *Braille mode.* In this mode embossed Braille text descriptions were provided with the TG. The user had to switch back and forth between the TG and the description sheet to read the description. For convenience, Braille legends were placed next to the objects illustrated on the TG (Figure 8).

(3) *Screen reader mode.* For this mode, text descriptions for the TG were printed on standard (rather than swell)

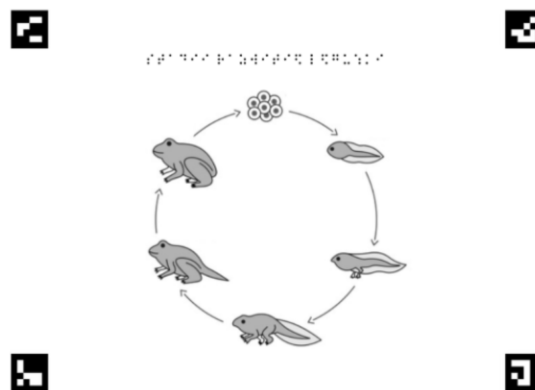


FIGURE 7. Example of TG used in the app exploration mode. No legends.

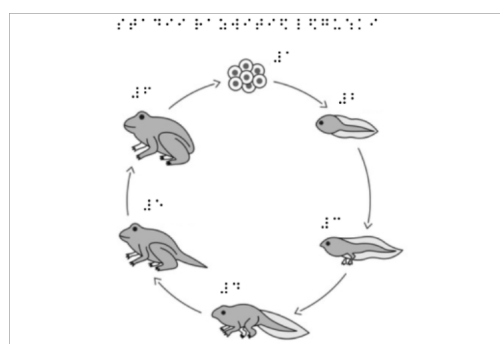


FIGURE 8. Example of TG used in the Braille and SR exploration modes. With Braille legends.

paper. Participants used their own mobile phones and their preferred apps to capture the document and convert printed text to speech. Two participants used a standalone device⁷ with optical character recognition (OCR) and audio output which converts printed text into speech since they did not have an appropriate document reader app installed. Response times were not affected by the choice of device, since all measurements were taken after the text was captured and spoken aloud once. Since the users who selected this mode were able to read Braille numbers (but were not fluent Braille readers), the TG with the Braille legends were used in this mode for their convenience (Figure 8).

b: GRAPHIC TYPES

Object: The images of the frog lifecycle and the space shuttle were used in this part. First, participants had to listen/read the general information about the object presented on the TG. They then explored the object thoroughly and answered questions about it. This task type was considered to be the most difficult one due to the amount of information provided.

Graph: Two different histograms were used. The associated tasks were the easiest, as most of the information could

⁶<https://www.pricerunner.com/Mobile-Phones/Samsung-Galaxy-A52-128GB-Compare-Prices>

⁷Standalone machine with a camera. Similar to this one <https://www.visionaid.co.uk/standalone-reading-machines/readeasy-evolve>



FIGURE 9. Examples of object, graph and map TG.



FIGURE 10. QR code printed on the other side of TG.

have been acquired by touch, e.g., the length and the location of the bar.

Map: The maps of Australia and Kyrgyzstan were used and participants used the same exploration algorithms as was described in the “Object” type.

A tactile image was placed in front of the participant at the start of each task. Then, a short summary of the image was presented in the appropriate exploration mode. In the app mode, the summary was provided in an audio format as soon as the QR code was scanned. A large QR code was placed on the other side of the paper at a point corresponding to the centre of the tactile graphic (Figure 10). The dimensions of the QR code in our study are much larger than in previous works. This makes it much easier for visually impaired users to scan the code because the risk of occlusions is reduced. Correspondingly, none of the participants experienced any difficulties in scanning the code during the experiments. In the screen reader mode, a text document with the summary was first scanned by the text reader app and then spoken out loud by the device. In the Braille mode, a sheet with the corresponding Braille text was used.

After the participants had familiarised themselves with the summary, the researcher read each question out loud and recorded the answers and response times. There were three general questions and one memory question for each object. Participants were free to continue exploring the graphic while answering the first three questions. The final memory question was used to assess how different exploration modes affect the participants’ ability to remember the information.

Since it was a test of memory, participants were not allowed to use the TG, Braille text information or the App while answering this question. Participants were not told whether their answers were correct, as this is what happens in, for instance, an examination situation.

As an illustration the questions for the TG with a map of Australia are presented below:

- 1) Which State or territory is surrounded by Western Australia, Queensland and South Australia (Answer: Northern Territory)
- 2) What is the capital of New South Wales (Answer: Sydney)
- 3) Which state is a separate island (Answer: Tasmania)
- 4) Memory question (Without TG). How many states and territories are there in Australia (Answer: 6 states and 2 territories)

The order of the questions in each task was the same for all participants. The order of the exploration modes always started with the app mode and was followed by the screen reader or Braille mode.

The following example shows the ordering of one session’s questions.

- 1) Object 1 (App)
- 2) Object 2 (SR/Braille)
- 3) Map 1 (App)
- 4) Map 2 (SR/Braille)
- 5) Graph 1 (App)
- 6) Graph 2 (SR/Braille)

This approach was used to ensure that participants all completed the tasks under identical conditions.

5) PHASE 3 (END-USER FEEDBACK)

The experiments were followed by a short feedback session where participants shared their initial thoughts and experiences of using the app. The participants were also asked for three things they liked and three they disliked about the app and to suggest additional features they would like to see in the app. Finally, participants were asked to answer six Likert scale questions. See Table 9.

Upon completion of the three sessions, participants were debriefed. First, the correct answers for the experimental session were revealed. Then, researchers presented a brief explanation of the hypothesis they tested. Finally, participants were asked whether they want to receive a report of the study outcomes.

V. RESULTS

This section has three subsections where we present the main findings of our study. They are ordered as follows: interview, the experimental tests on the device and end-user feedback about the device. Discussion of the experimental outcomes is divided into two parts, which respectively compare the app and Braille exploration modes and the app and screen reader exploration modes. The comparison included analysis of three main measures: time taken, accuracy (of responses to the first three questions) and memory accuracy (accuracy

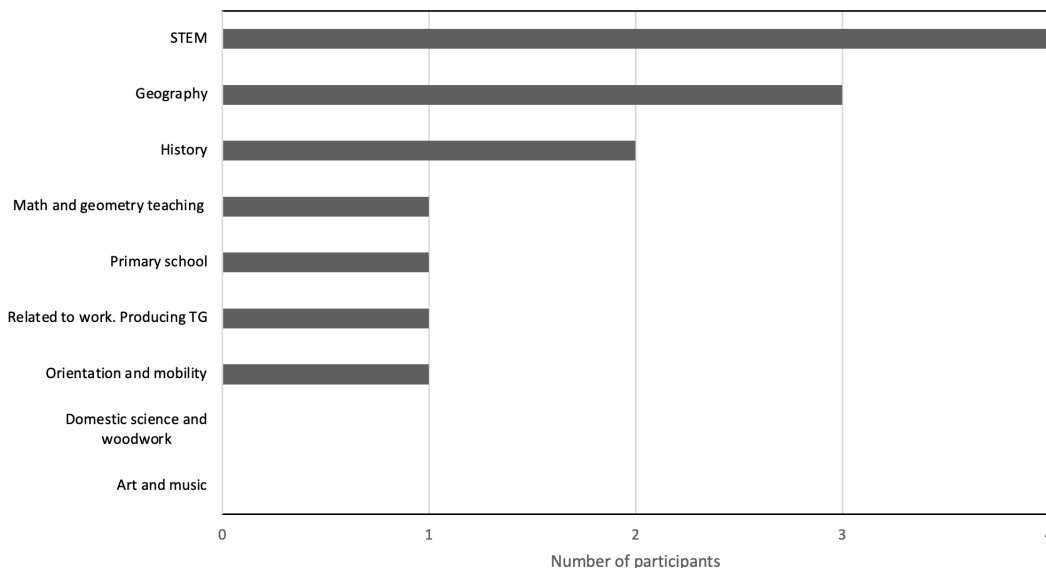


FIGURE 11. In which subjects TG were used.

of responses to the memory question). Comparisons of the data for different age and vision loss groups were also made.

A. INTERVIEW

1) EXPERIENCE OF USING TACTILE GRAPHICS

Six participants had used tactile graphics at school and the other five had not. TG had mainly been used in STEM subjects and were labelled with Braille text. More detailed information is presented in Figure 11. All the participants who had used TG reported that it was easier to understand the subject when TG were used in class. For example, TG helped them to learn the computer keyboard. In particular, they used the tactile version of the keyboard to learn where the different keys were located on the keyboard. In chemistry classes a tactile version of the periodic table was used to make it easier to understand the material. However, the TG used had the drawback of requiring a teacher to show each student individually how to navigate the TG before they could use it on their own.

2) EXPERIENCE WITH MOBILE DEVICES AND APPLICATIONS

During the study, participants were asked about their use of mobile devices. Seven participants had Android and four had iOS smartphones. Nine participants were actively using the phone camera and two participants said that they used it very rarely. Most of the participants (8/11) used a camera to take photos. Other apps which required a camera, and which were used by the participants were: currency and barcode readers, colour and photo identifiers, light detector, BeMyEyes, TapTapSee and Google Translate apps (Figure 12). The only difficulty in using these apps was aiming the camera. Eight participants (73%) indicated that they experienced difficulties

TABLE 5. App vs Braille text.

Mode	App	Braille
Average time (sec)	42.29	72.55
Average accuracy (%)	92.84	87.38
Memory accuracy (%)	85.67	78.09

with this, but did not comment on their difficulties in the open-ended interview questions.

B. DEVICE TESTING

1) COMPARISON OF APP AND BRAILLE TEXT

On average, it took less time for the participants to answer the questions and with greater accuracy while exploring the TG using the app compared to Braille (Table 5).

Results of the non-parametric Wilcoxon signed-rank test [45] showed that TG exploration mode had a statistically significant impact on the **time spent** to answer the questions. The p-value is 0.0005 ($p < 0.05$).

The same tests for the **average accuracy** and **memory accuracy** showed that the difference between the two groups was not significant ($p = 0.13$) and ($p = 0.67$) respectively ($p > 0.05$).

We also analysed the impact of participants' ages on their performance. Their average age was about 35 years. They were therefore divided into two groups: over 35 ($n = 5$, mean = 45) and under 35 ($n = 6$, mean = 27). From Table 6 it can be seen that the under 35 age group was answering faster but showed worse performance in remembering the information in both exploration modes. The percentage of correct answers was almost the same for both groups and a non-parametric Mann-Whitney U tests indicated that these small differences were not significant. The p-values of the impact of the age on

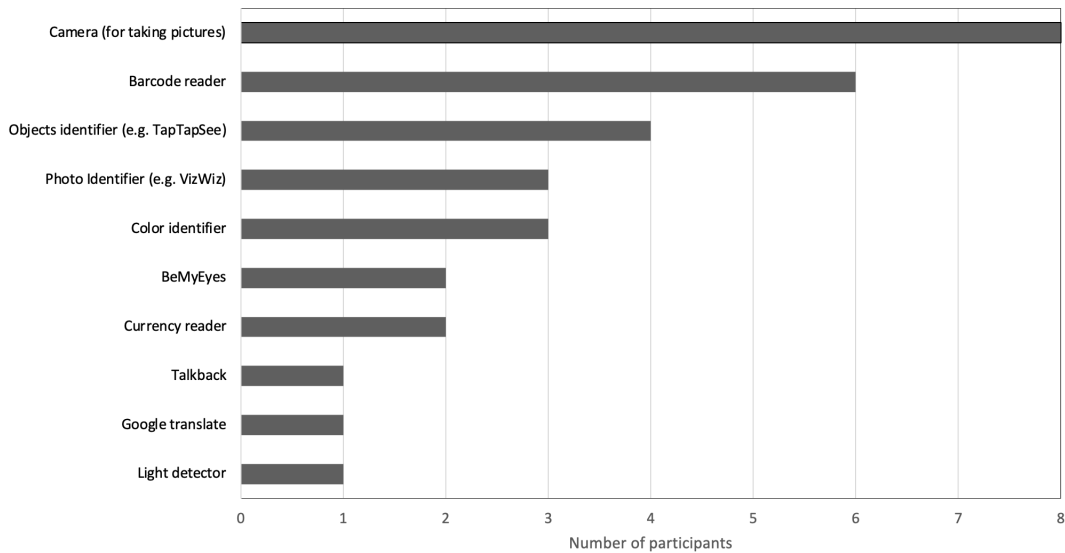


FIGURE 12. Apps used by participants.

time, accuracy and memory accuracy are **0.13, 0.85 and 0.28** respectively.

Five participants indicated that they were blind, either totally blind or blind and able to distinguish light and dark. Six participants stated that they were able to distinguish shapes and read very large print texts, classified as partially sighted. Comparison of the results of these two groups showed that, as might have been expected, participants who were able to see object shapes, spent less time exploring the TG and answering the questions (Table 6). However, blind participants memorised the information better in both modes. On the other hand, the Mann-Whitney U tests showed that these results are not statistically significant. The p-values of the impact of the type of visual impairment on time, accuracy and memory accuracy are **0.52, 0.99 and 0.57** respectively. The lack of significance may be due to the small number of participants and a larger sample size could support the hypotheses.

2) COMPARISON OF APP AND SCREEN READER

As can be seen from Table 7, participants had almost identical average response times in both modes, but there were noticeable differences in the accuracy of the answers, including the memory ones. The results of the Wilcoxon signed-rank tests showed that TG exploration mode was statistically significant for **average accuracy** ($p = 0.04$) but not for the **time spent** ($p = 0.97$) and the **memory accuracy** ($p = 0.07$).

As in the previous section, after running the Mann-Whitney U tests, it was found out that differences between age and vision loss groups were not statistically significant. The p-values of the impact of the age on time, accuracy and memory accuracy were **0.67, 0.45 and 0.99** respectively. The results of the same tests for the vision loss are **0.87, 0.97 and**

TABLE 6. Comparison of different age and vision impairment groups.

Mode	App		Braille	
	Under 35	Over 35	Under 35	Over 35
Age				
Time spent (sec)	31.33	50.52	68.89	75.29
Average accuracy (%)	94.42	91.67	86.56	88.00
Memory accuracy (%)	77.67	91.67	68.44	85.33
Vision	VI	Blind	VI	Blind
Time spent (sec)	37.89	45.98	57.78	83.63
Average accuracy (%)	95.36	90.96	85.39	88.87
Memory accuracy (%)	81.44	88.33	74.89	80.50

TABLE 7. App vs screen reader.

Mode	App	SR
Average time (sec)	40.46	40.83
Average accuracy (%)	95.65	71.69
Memory accuracy (%)	90.91	78.41

0.28. We believe that repeating the experiments with a larger sample may affect the significance of the results.

C. END-USER FEEDBACK

A post-experimental set of Likert-scale questions was carried out to investigate respondents' attitudes toward the app. According to the results presented in Table 9, a majority of the users were satisfied with the application. The overwhelming majority, 9 of 11 (82%) strongly agreed that they would be

TABLE 9. Results of the Likert scale question.

Statement	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
The app helped me to understand concepts better than Braille or text descriptions alone	6	2	1	2	0
I found it easy and intuitive to use the app	3	7	0	1	0
I found it easy to properly aim the camera	2	4	1	4	0
I am satisfied with the app response time	2	5	1	3	0
I liked the detailed description app feature	9	1	0	0	1
I would be interested in using this app to get better quality information during my classes on a daily basis	9	1	0	0	1

may lose focus while doing this. Thus, the results confirm that the TAURIS app is able to support visually impaired users to study effectively and independently.

Our study confirmed evidence from earlier research [9], [12], [34] that TGs on their own are not sufficient to aid visually impaired people in learning. The experiments further showed that the additional audio information provided by the TAURIS app improved the usefulness and information in the tactile images. This finding is consistent with those of [10] and [25]. Contrary to expectations, this study did not find statistically significant differences between different vision loss and age groups. However, one interesting finding was that blind participants remembered the information better, but partially sighted ones were faster. The lack of significance may be due to the small sample and it would be useful to repeat the research with a larger sample to investigate this.

Participant feedback will be used to improve the app in the future. The improvements will include adding a customisation option to increase the speed of speech output, change the synthetic voice and notifying users when light levels are too low and the algorithm is struggling to detect the objects. We may also consider creating a universal frame with built-in Aruco markers to increase the effective area of the printed tactile images.

There is one issue that is more difficult to resolve. Several participants said that they felt more comfortable reading and answering questions in the Kazakh language. Unfortunately, there is currently no speech synthesiser for Kazakh. This raises wider issues of access to speech synthesisers in a much wider range of languages and we encourage researchers and developers to consider this, with particular reference to the Kazakh language.

Together these results provide important insights into the effectiveness of each approach. Participants were not only able to respond more quickly when using the app mode, but they also understood and remembered the material better and these differences were statistically significant. The results of the open-ended interviews also provided important insights into the validity of the TAURIS application. The majority of participants agreed that the app approach was more effective for learning in terms of both the time required and the quality of the information obtained. In addition, more than half of

the participants agreed that it was relatively easy to point a camera with the help of Aruco markers.

VII. CONCLUSION AND FUTURE WORK

We have presented TAURIS, a new app that enables visually impaired people to access tactile graphics by providing audio descriptions in real time. It does not require specialised equipment or proprietary software; all the user requires is an Android smartphone with our application installed. The developed educational mobile application relies on a deep learning computer vision model to detect the user's finger. This gives robust and accurate detection even under poor lighting conditions. We also conducted a three-part study to evaluate the system. This consisted of a structured interview to obtain qualitative data about participants' experience with the TG and mobile devices, a series of experiments to obtain quantitative measurements of time and accuracy, and a post-experimental session to record participants' thoughts and opinions about the proposed system and suggestions for further developments. The results of the experiments demonstrated that the proposed mobile application supported users to explore TG more efficiently. These findings were borne out by user feedback obtained subsequently. To sum up, the current study has obtained a proof of concept for the TAURIS app, a comprehensive system that will make educational material accessible to visually impaired students. It should be noted that the app has been developed specifically for use in education, but it is not restricted to education and can be used more widely.

Proof of concept has been obtained, but, as indicated above, there is potential for improvement, particularly based on user feedback suggestions. This includes a universal frame with the Aruco markers placed at the corners to increase the TG effective area, user notification when light levels are too low so they can increase them, and an option for customising the speed of speech output. The most important limitation of our study is the small sample size. On the other hand, this is not uncommon in research which involves individuals with disabilities. Also, it was not possible to assess how one-handed and two-handed exploration affect performance. Therefore, there would be value in carrying out the same experiments with a larger number of participants in several

different schools for the blind and also different countries in order to obtain further useful insights, increase the power of the statistical analysis and enable further comparisons.

APPENDIX A INTERVIEW QUESTIONS

A. PERSONAL INFORMATION FOR STATISTICAL PURPOSES

1. Please select the option that best describes you
 - Secondary school student
 - University student
 - College student or a vocational trainee
 - A person who has left education in the last five years
 - Other (Please specify)
 2. What is your age?
 3. Which gender do you identify with?
 - Male
 - Female
 - Other
 4. What country do you live in?
 5. Which of the following best describes your vision?
 - Totally blind with no awareness of light
 - Blind and able to distinguish light and dark
 - Able to see shapes, but unable to distinguish detail or to read print
 - Able to read a large print text
 - Other (Please specify)
 6. Which of the following statements best describes you?
 - I have been blind or partially sighted since birth
 - I lost my sight between 0 and 3 years
 - I lost my sight between 3 and 11 years
 - I lost my sight between 11 and 35 years
 - I lost my sight after the age of 35
 7. Which of the following best describes the school you are attending now or attended most recently? (Check all that apply)
 - School for the blind
 - A mainstream school with resource room
 - The mainstream school (full-time classroom)
 - Other (Please specify)
 8. Can you read Braille?
 - Yes
 - No
 - I understand some Braille but I am not fluent
 9. Have you used tactile graphics?
 - Yes, in education
 - Yes, in other applications but not education
 - No
- Please, proceed to section B if your answer to the previous question was “YES”.**
- If your answer was “NO”, please answer the next two questions and you can skip section B.**
10. Would you like to use tactile graphics?
 - Yes
 - No (please specify the reason)
 - Unsure (please specify the reason)

11. For what reasons have you not used them?
 - Not provided by teachers
 - No training was provided on how to use them
 - Other (please specify)

B. YOUR EXPERIENCE WITH TACTILE GRAPHICS

12. What subjects have you used tactile graphics in? (Check all that apply)
 - STEM (Science, technology, engineering and mathematics)
 - Geography
 - History
 - Orientation and Mobility classes
 - Art and music
 - Domestic science and woodwork
 - Other (please specify)
13. How were these tactile graphics usually labelled?
 - Braille text
 - Large printed text
 - Audio descriptions
 - Not labelled
 - Other (please specify)
14. How often do you use tactile graphics in education?
 - Most of the time (3-4 times a week or more)
 - Sometimes (1-2 times a week)
 - Rarely (a few times a month or less)
 - Almost never
15. Please select the option that best applies to you
 - Tactile graphics are sufficient on their own without Braille and text descriptions and help me understand the concept better
 - Tactile graphics complement Braille and text descriptions and help me understand the concept better
 - Tactile graphics do not help me understand the concept better
16. Is it easier to keep up when tactile graphics are used in class?
 - Yes, they make a big difference
 - Yes, they help a bit
 - No, they make no difference
 - No, I can keep up without them
17. Indicate all of the following that would help you to use tactile graphics more effectively
 - Training in using tactile graphics
 - Being shown how to use/oriented to each tactile graphics
 - Audio descriptions provided with tactile graphics
 - Training in using tactile graphics with audio descriptions
 - Other (please specify)
18. Please give three examples of how you have used tactile graphics to support your learning with comments on how useful you found them
19. Please comment on any features of tactile graphics you have used that makes them easy or difficult to read and any strategies you have used to read new tactile graphics.
20. Do you have any suggestions on how tactile graphics could be improved?

C. YOUR EXPERIENCE WITH ASSISTIVE TECHNOLOGY PHONE APPLICATIONS

21. Which mobile phone do you use?
 Android smartphone
 IOS smartphone
 Another smartphone
 Phone without a camera
 I do not use a mobile phone
22. If you ever used an app which requires you to use a phone camera, which of the following apps have you used? (Check all that apply)
 Camera (for taking pictures)
 Currency reader
 Color identifier
 Photo Identifier (e.g. VizWiz)
 Light detector
 Barcode reader
 Object identifier (e.g. TapTapSee)
 Other (Please specify)
23. How often do you use an app which requires you to use a camera?
 Everyday
 Once or twice a week
 Once a month
 Almost never
24. Do you want applications to give you feedback to properly aim the camera?
 No, I can aim the camera without feedback
 Yes, I would use camera apps more frequently if there was feedback on how to aim the camera
 Yes, I would consider using camera apps if feedback to aim the camera was available
25. How the feedback should be conveyed to you while aiming the camera?
 Vibration
 Voice information
 Tones
 A combination of the above
 Other, please specify

REFERENCES

- [1] K. T. Zebehazy and A. P. Wilton, "Charting success: The experience of teachers of students with visual impairments in promoting student use of graphics," *J. Vis. Impairment Blindness*, vol. 108, no. 4, pp. 263–274, Jul. 2014.
- [2] L. P. Rosenblum, L. Cheng, and C. R. Beal, "Teachers of students with visual impairments share experiences and advice for supporting students in understanding graphics," *J. Vis. Impairment Blindness*, vol. 112, no. 5, pp. 475–487, Sep. 2018.
- [3] F. K. Aldrich and L. Sheppard, "Tactile graphics in school education: Perspectives from pupils," *Brit. J. Vis. Impairment*, vol. 19, no. 2, pp. 69–73, May 2001.
- [4] A. Paivio, *Imagery and Verbal Processes*. New York, NY, USA: Holt, Rinehart, Winston, 1971.
- [5] C. L. Grady, A. R. McIntosh, M. N. Rajah, and F. I. M. Craik, "Neural correlates of the episodic encoding of pictures and words," *Proc. Nat. Acad. Sci. USA*, vol. 95, no. 5, pp. 2703–2708, Mar. 1998.
- [6] J. Novak and A. Cañas, *The Theory Underlying Concept Maps and How to Construct and Use Them*. New Delhi, India: Institute for Human, 2008.
- [7] M. Hersh and M. Johnson, *Assistive Technology for Visually Impaired and Blind People*. Cham, Switzerland: Springer, 2008.
- [8] E. Berlá, "Behavioral strategies and problems in scanning and interpreting tactual displays," *New Outlook Blind.*, vol. 66, no. 8, p. 277, 1972.
- [9] K. T. Zebehazy and A. P. Wilton, "Straight from the source: Perceptions of students with visual impairments about graphic use," *J. Vis. Impairment Blindness*, vol. 108, no. 4, pp. 275–286, Jul. 2014.
- [10] G. Fusco and V. S. Morash, "The tactile graphics helper: Providing audio clarification for tactile graphics using machine vision," in *Proc. 17th Int. ACM SIGACCESS Conf. Comput. Accessibility (ASSETS)*, 2015, pp. 97–106.
- [11] D. Prescher, J. Bornschein, and G. Weber, "Production of accessible tactile graphics," in *Proc. Int. Conf. Comput. Handicapped Persons.*, Jul. 2014, pp. 26–33.
- [12] L. Sheppard and F. K. Aldrich, "Tactile graphics in school education: Perspectives from teachers," *Brit. J. Vis. Impairment*, vol. 19, no. 3, pp. 93–97, Sep. 2001.
- [13] D. Parks, "An audio-tactile tool for the acquisition, use and management of spatially distributed information by partially sighted and blind persons," in *Proc. 2nd Int. Symp. Maps Graph. Visually Handicapped People*, 1988, pp. 24–29.
- [14] R. Jacobson, "Navigating maps with little or no sight: An audio-tactile approach," in *Proc. Content Visualizat. Intermedia Represent. (CVIR)*, 1998, pp. 1–8.
- [15] S. Landau and K. Gorgey, "Development of a talking tactile tablet," *Inf. Technol. Disabilities*, vol. 7, no. 2, pp. 1–3, 2001.
- [16] J. Gardner and V. Bulatov, "Scientific diagrams made easy with IVEO™," in *Proc. Int. Conf. Comput. Handicapped Persons*, 2006, pp. 1243–1250.
- [17] G. Petit, A. Dufresne, V. Levesque, V. Hayward, and N. Trudeau, "Refreshable tactile graphics applied to schoolbook illustrations for students with visual impairment," in *Proc. 10th Int. ACM SIGACCESS Conf. Comput. Accessibility Assets*, 2008, pp. 89–96.
- [18] L. C. Quero, J. I. Bartolomé, and J. Cho, "Accessible visual artworks for blind and visually impaired people: Comparing a multimodal approach with tactile graphics," *Electronics*, vol. 10, no. 3, p. 297, Jan. 2021.
- [19] S. K. Kane, B. Frey, and J. O. Wobbrock, "Access lens: A gesture-based screen reader for real-world documents," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Apr. 2013, pp. 347–350.
- [20] M. I. H. S. Bahrin, H. M. Yusof, and S. N. Sidek, "Hands and fingers tracking for tactile graphics reading assistive device," in *Enabling Industry 4.0 Through Advances in Mechatronics*. Singapore: Springer, 2022, pp. 413–422.
- [21] H. Shen, O. Edwards, J. Miele, and J. M. Coughlan, "CamIO: A 3D computer vision system enabling audio/haptic interaction with physical objects by blind users," in *Proc. 15th Int. ACM SIGACCESS Conf. Comput. Accessibility*, Oct. 2013, pp. 1–2.
- [22] A. Reichinger, A. Fuhrmann, S. Maierhofer, and W. Purgathofer, "Gesture-based interactive audio guide on tactile reliefs," in *Proc. 18th Int. ACM SIGACCESS Conf. Comput. Accessibility*, Oct. 2016, pp. 91–100.
- [23] L. Shi, Y. Zhao, and S. Azenkot, "Markit and Talkit: A low-barrier toolkit to augment 3D printed models with audio annotations," in *Proc. 30th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2017, pp. 493–506.
- [24] J. Pasquero and V. Hayward, "STRess: A practical tactile display system with one millimeter spatial resolution and 700 Hz refresh rate," in *Proc. Eurohaptics*, 2003, pp. 94–110.
- [25] G. Melfi, K. Müller, T. Schwarz, G. Jaworek, and R. Stiefelhagen, "Understanding what you feel: A mobile audio-tactile system for graphics used at schools with students with visual impairment," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2020, pp. 1–12.
- [26] C. Baker, L. Milne, J. Scofield, C. Bennett, and R. Ladner, "Tactile graphics with a voice: Using QR codes to access text in tactile graphics," in *Proc. 16th Int. ACM SIGACCESS Conf. Comput. Accessibility*, 2014, pp. 75–82.
- [27] L. Shi, Y. Zhao, R. G. Penuela, E. Kupferstein, and S. Azenkot, "Molder: An accessible design tool for tactile maps," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2020, pp. 1–14.
- [28] Y. Hosokawa, Y. Miwa, and Y. Hashimoto, "Development of TARS mobile app with deep fingertip detector for the visually impaired," in *Proc. Int. Conf. Comput. Helping People With Special Needs*, 2020, pp. 435–445.
- [29] T. Miwa, Y. Hosokawa, Y. Hashimoto, and G. Lisi, "TARS mobile app with deep fingertip detector for the visually impaired," in *Proc. Int. Conf. Intell. Hum. Syst. Integr.*, 2020, pp. 301–306.

- [30] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C.-L. Chang, and M. Grundmann, "MediaPipe hands: On-device real-time hand tracking," 2020, *arXiv:2006.10214*.
- [31] A. Angelova, A. Krizhevsky, V. Vanhoucke, A. Ogale, and D. Ferguson, "Real-time pedestrian detection with deep network cascades," in *Proc. Brit. Mach. Vis. Conf.*, 2015, pp. 1–12.
- [32] M. W. A. Wijntjes, T. van Lienen, I. M. Verstijnen, and A. M. L. Kappers, "Look what I have felt: Unidentified haptic line drawings are identified after sketching," *Acta Psycholog.*, vol. 128, no. 2, pp. 255–263, Jun. 2008.
- [33] F. Bara, "Exploratory procedures employed by visually impaired children during joint book reading," *J. Develop. Phys. Disabilities*, vol. 26, no. 2, pp. 151–170, Apr. 2014.
- [34] K. T. Zebehazi and A. P. Wilton, "Quality, importance, and instruction: The perspectives of teachers of students with visual impairments on graphics use by students," *J. Vis. Impairment Blindness*, vol. 108, no. 1, pp. 5–16, Jan. 2014.
- [35] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marin-Jimenez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognit.*, vol. 47, pp. 2280–2292, Jun. 2014.
- [36] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.
- [37] F. J. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image Vis. Comput.*, vol. 76, pp. 38–47, Aug. 2018.
- [38] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision With the OpenCV Library*. Sebastopol, CA, USA: O'Reilly Media, 2008.
- [39] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.
- [40] W. Wu, C. Li, Z. Cheng, X. Zhang, and L. Jin, "YOLSE: Egocentric fingertip detection from single RGB images," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 623–630.
- [41] F. Leo, E. Cocchi, and L. Brayda, "The effect of programmable tactile displays on spatial learning skills in children and adolescents of different visual disability," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 7, pp. 861–872, Jul. 2017.
- [42] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 696–706, May 2002.
- [43] H.-C. Do, J.-Y. You, and S.-I. Chien, "Skin color detection through estimation and conversion of illuminant color under various illuminations," *IEEE Trans. Consum. Electron.*, vol. 53, no. 3, pp. 1103–1108, Aug. 2007.
- [44] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, nos. 3–4, p. 591, Dec. 1965.
- [45] F. Wilcoxon, "Individual comparisons by ranking methods," in *Breakthroughs in Statistics*. New York, NY, USA: Springer, 1992, pp. 196–202.



MARALBEK ZEINULLIN received the B.S. degree in biomedical engineering from Arizona State University, USA, in 2014. He is currently pursuing the Ph.D. degree with the University of Glasgow. His research interests include assistive technology and computer vision.



MARION HERSH is currently a Senior Lecturer (an Associate Professor) at the University of Glasgow, U.K. Her research interests include end-user driven and co-production approaches to working to the design of assistive and other technologies for disabled and older people, universal and sustainable design, and ethical engineering. Her projects include the spatial representations of visually impaired, communication technologies for visually impaired and hearing impaired, barriers experienced, and strategies used by autistic people and technology solutions to support independence, participation, and quality of life of older autistic people. She has over 200 publications, including books on assistive technology for blind and deaf people.

...