Kufa Journal of Engineering
Vol. 15, No. 3, July 2024, P.P. 68-81
Article history: Received 6 January 2024, last revised 20 March 2024, accepted 9 April 2024

A DATASET FOR THE DIACRITICS IMAGES OF THE HOLY QURAN: TOWARDS A TACTILE-VISION SYSTEM FOR THE VISUALLY IMPAIRED

Mazin H. Aziz

PhD Lecturer, Computer Engineering Department, College of Engineering, University of Mosul, Mosul, Republic of Iraq. Email: mazin.haziz@uomosul.edu.iq

https://doi.org/10.30572/2018/KJE/150305

ABSTRACT

The only way for the Blind and visually impaired persons to read the holy Quran is by using a special paper edition of embossed Braille code. A new approach to reading the Holy Quran text for those persons was proposed by the author. One of the main demands for this approach is the classification of the dots and diacritics, which can be abbreviated as DaDs. This paper outlines the creation of a dataset of images for the DaDs of Al-Mushaf. A handheld scanner was developed for this purpose, and MATLAB programs were employed for DaDs segmentation. The final goal is the design of a tactile vision sensory substitution system based on an optical character recognition technique to help blind individuals read the Holy Quran in an alternative way to Braille codes. Approximately 1750 images were taken from two distinct Al-Mushaf versions with the proposed handheld scanner. Using the suggested techniques and algorithms, 6000 DaDs were retrieved from these images; however, only 4710 images of the DaDs, arranged in 22 classes, were selected after the repeated, incomplete, and non-DaDs were eliminated. The dataset was organized by DaD class, prepared to be used directly for machine learning purposes, and made available for public use upon request.

KEYWORDS

Vision Sensory Substitution (VSS), Optical Character Recognition (OCR), Image Dataset, Holy Quran, Dots and Diacritics.

1. INTRODUCTION

Tactile Visual Sensory Substitution (TVSS), dots and diacritics (DaDs) of Al-Mushaf, and the Braille translation of Al-Mushaf are very related to this work and are presented here briefly.

1.1. Tactile Visual Sensory Substitution (TVSS)

Visual sensory substitution by tactile stimulation was firstly proved by Paul Bach Rita at 1996 (Bach-Y-Rita et al., 1969), and it is based on the brain plasticity. TVSS consisted of an image sensor (camera), a processing unit, and a tactile actuators called the tactile display (Kaczmarek et al., 1991). The tactile actuators are arranged in a matrix form and attached to the skin of the blind or visually impaired. The actuators could be vibrators, electrical, thermal, or others that are used to stimulate the skin of the blind person. After sufficient training session, the brain of a blind individual will learn how to transform this tactile pattern into visual data that the brain can perceive, even though the stimulated pattern has very low resolution (Bach-y-Rita, Tyler and Kaczmarek, 2003). The tactile display could be attached to different parts of the blind person body like: back, front head, arm, or even the tongue (Bach-y-Rita and W. Kercel, 2003).

1.2. Dots and Diacritics (DaDs) of Al-Mushaf

Al-Mushaf writing is composed of two parts. The main part is the grapheme (the letters) which is the backbone of the initial handwritten version of AL-Mushaf. The subsidiary portion consists of the dots and the diacritics (DaDs) which are appended to the primary script afterword. The diacritics are small sized letters, marks, and strokes. The DaDs are positioned above, below, or in between the grapheme, which is located on the text's main line as seen in Fig. 1. The shape's similar letters are made less ambiguous by the addition of dots. Certain diacritical marks are

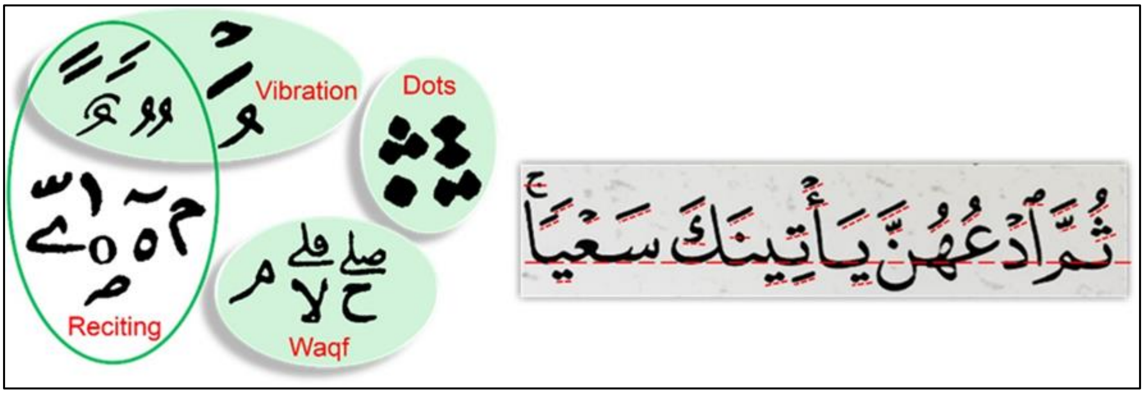


Fig. 1. An example showing the locations and orientations of DaDs relative to the base line in Al-Mushaf, indicated with red lines (right). The indications and purposes of (DaDs) groups (left).

employed as vibration indicators to assist the reader in pronouncing words correctly and preventing them from misinterpreting phrases or words. To highlight the Telawa norms for reciting, additional diacritical marks are employed. Other diacritical marks, known as Waqf signs, are employed to indicate when to begin, halt, or continue reading a phrase. Owing to calligraphy conventions, some diacritical marks are sketched parallel to the main line of the script while others have slanted angles. In addition, there is adornment surrounding the verse numbers (Ayat) and at the beginning of each chapter (Sura).

1.3. Braille Translation of Al-Mushaf

For many years, AL-Mushaf translation using Braille code was considered a priceless asset by many blind people throughout the Islamic world. The Braille translation of AL-Mushaf is very helpful, however, it just includes the word codes, including dots and basic vibration indicators. The information included in AL-Mushaf is not entirely provided in it. Raised dots arranged in distinct cells, with either six or eight dots per cell, make up Braille codes. (Rubini and Setyawan, 2020). Fig. 2 demonstrates an example of the Braille code for an Arabic sentence.

1.4. Research Objectives

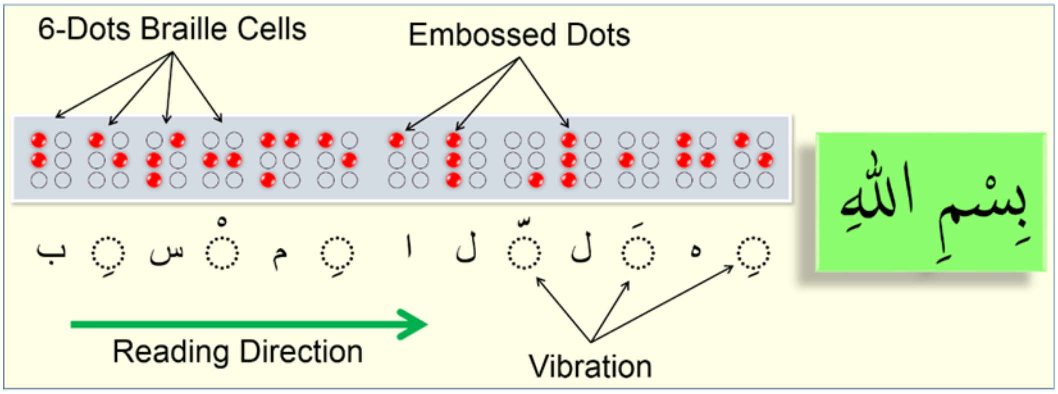


Fig. 2. Braille Translation of the Arabic Sentence (بسم الله).

This paper presents a step of a novel way of reading the Great Quran using the paper copy of AL-Mushaf that is available for public use, by the blind and visually impaired. This way is based on the Tactile Visual Sensory Substitution (TVSS) technique, and consists of handheld scanner that enables the blind person from scanning an ordinary paper edition of Al-Mushaf and sends the images to a processing unit that can prepare them to be displayed, in a simplified form, on a special tactile display unit, where the user can grope it to read, as shown in Fig. 3. One of the contents of the captured image Al-Mushaf that should be simplified is the DaDs, therefore, it is important to segment the DaDs, classify them and replace them with simplified synthetic DaDs. Hence DaDs classification which is based on machine learning techniques depends on the training phase using DaDs image data set, which is not available up to date for the Quranic script. The objective of this work is to implement a setup and procedure for the extraction of the DaDs from common paper editions of Al-Mushaf and the generation of a labelled dataset for the Quranic DaDs.

2. RELATED WORKS

The translation of Al-Mushaf into Braille has not been extensively studied. Few research studies have included some of the vibrating signals or reciting rules as an external Braille symbol

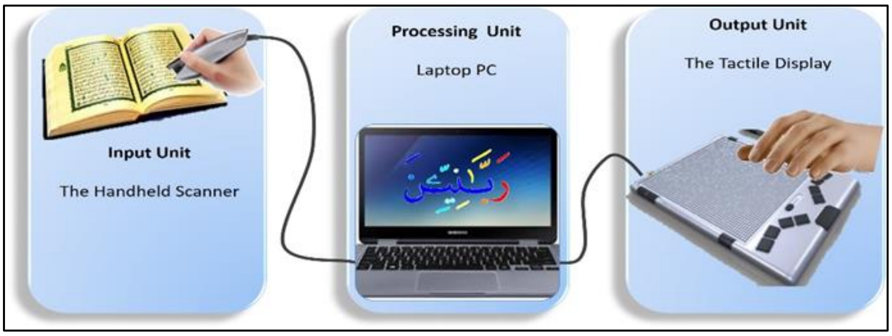


Fig. 3. The proposed tactile vision sensory substitution (TVSS) system for holy Quran reading.

(Abdallah Abualkishik and Khairuddin Omar, 2009) (Abualkishik and Omar, 2009)(Abualkishik and Omar, 2013), and occasionally a refreshable Braille display (Sulaiman, Rambli and Zuki, 2015) was employed. Although there are few studies on image processing for Al-Mushaf script, several (Kef, Chergui and Chikhi, 2016) (Eraqi and Abdelazeem, 2012)(Mousa, Sayed and Abdalla, 2013)(Eraqi and Azeem, 2011) (Fadel et al., 2019) (Mohammed, 2023) (Boubaker, Chaabouni, et al., 2014) deal with Arabic handwritten scripts. Deep learning was also used to extract Arabic text with Harakat from images(Maghraby and Samkari, 2023), and to segment words into graphemes(Elkhayati, Elkettani and Mourchid, 2022). Children's Arabic handwritten recognition was implemented using a Convolutional Neural Network (CNN)(Alheraki, Al-Matham and Al-Khalifa, 2023) and a Quranic OCR based on CNN is developed also(Mohd et al., 2021). The suggestion of using deep learning models for Arabic text diacritics recovering was presented in(Madhfar and Qamar, 2021). A dataset for Arabic handwritten letters by children was presented(Altwaijry and Al-Turaiki, 2021). A BERT model was used to restore diacritics synthetically(Nazih and Hifny, 2022). As far as I am aware, no previous research has linked the Holy Quran's image processing to the visual sensory replacement technique for the blind. Moreover, there was no dataset for the Quranic DaDs, which was the research gap that was addressed through this work.

3. THE PROPOSED METHODOLOGY

This section explains the proposed setup and the methods that were used for dataset generation.

3.1. SETUP

An off-the-shelf webcam was modified to work as a handheld image scanner and used to acquire DaDs images from ordinary paper editions of Al-Mushaf. This handheld scanner, shown in Fig. 4, is composed of the webcam's internal hardware, its optical parts and two lighting emitter diodes for illumination, housed in a plastic casing. To obtain the best quality focused image, the distance between Al-Mushaf papers and the camera lenses may be adjusted. The handheld scanner is connected to a laptop PC that was supplied with a program written in MATLAB to extract the DaDs images. The collected photographs have a resolution of 640x480 pixels, for a spatial image size on Al-Mushaf paper of 1.8 cm by 1.35 cm, as depicted in Fig.5. Each picture may include a word, a part of word (PAW) or more than a word with its accompanying DaDs. Images from two different editions of Al-Mushaf were obtained; they are both middle-sized, common paper versions, and based on Rewayat Hafs as seen in Fig. 6.

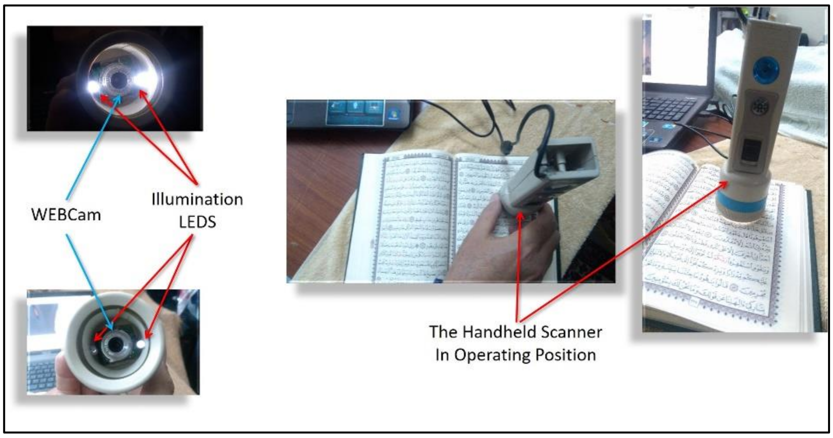


Fig. 4. The constructed handheld-scanner in operation.

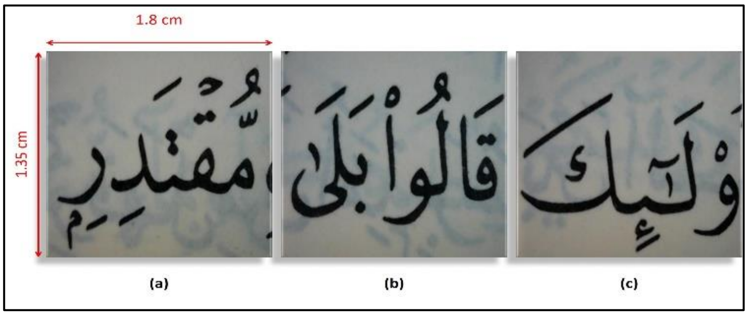


Fig.5. Three examples of the captured images, image resolution is (640×480 pixels) and image size is (1.8cm×1.35cm): (a) a complete word, (b) more than a word, and (c) part of a word (PAW).

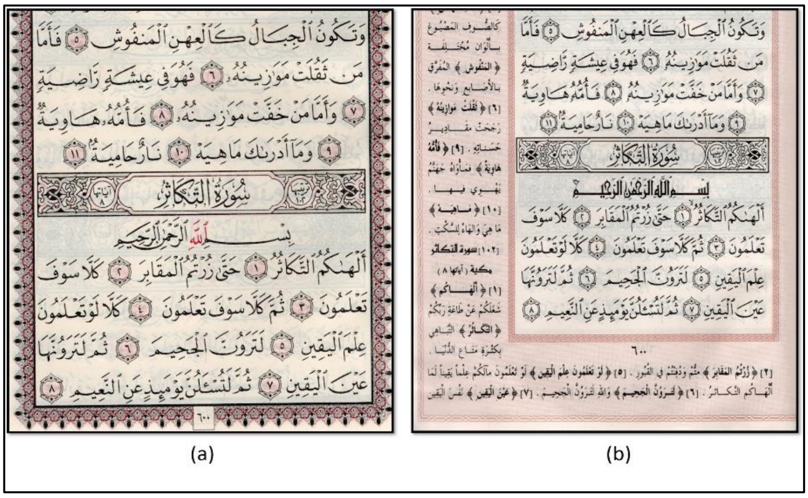


Fig. 6. Snapshots of the two versions of AL-Mushaf that were utilized to create our dataset. (a) Dar AL-Kalem AL-Taieb and (b) Dar AL-Fadjr.

3.2. METHODS

For clarity, Fig. 7 shows a colorful sample from the acquired images with the grapheme, dots, and diacritical marks displayed in red, green, and blue, respectively. The proposed DaDs dataset should not contain any part of the grapheme. Table 1 lists the 27 DaD kinds that are most commonly utilized in Al-Mushaf. For the dataset that needs to be developed, these DaD types represent potential classes. This table displays the calligraphic forms that Al-Mushaf utilized in his writing, in addition to the DaD's name in both Arabic and English. Our method for creating this dataset is as follows.

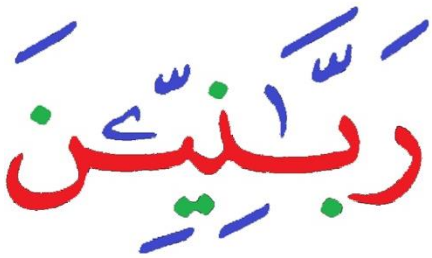


Fig. 7. The three main parts of Al-Mushaf text: The grapheme (red), the dots (green), the diacritics (blue).

3.2.1. IMAGE CAPTURING

More than 2000 photos from various locations throughout the two editions of Al-Mushaf were taken with the specially designed portable scanner and stored in a file. 1750 photos were chosen to be utilized in the next stages. The omitted photos were rejected due to human error, the possibility that they included non-script decorating elements, or the fact that they closely matched other images. The method of image capturing made sure that the collection of pictures followed the changes in the DaDs' objects' sizes, shapes, and tilts throughout Al-Mushaf's script.

Table 1. Classes of the DaDs that are used in Al-Mushaf (Rewayat Hafs especially).

Item No.	DaDs' name	DaDs' shape	DaDs' name in Arabic
1	A Dot		نقطة مفردة
2	Two Dots		نقطتان افقيتان
3	Vertical Two Dots		نقطتان عموديتان
4	Triple Dots		ثلاث نقاط
5	Fatha		فتحة
6	Damma		ضمة
7	Socon		سكون
8	Two Damma		ضمة مزدوجة
9	Tanween Damm		تنوين ضم
10	Two Fatha		فتحة مزدوجة
11	Tanween Fathh		تنوين فتح
12	Maddat		مدة
13	Shaddat		شدة
14	Alef Super Script		ألف خنجرية
15	Hamzat Wasl		همزة وصل
16	Hamzat		همزة
17	Hamzat Kaf		همزة كاف
18	Circle 1		دائرة ١
19	Circle 2		دائرة ٢
20	Yaa'		ياء
21	Meem Iqlab		ميم إقلاب
22	Meem Lazim		ميم وقف لازم
23	Jeem		جيم للجواز
24	La		لا للمنع
25	Sala		صلى
26	Qala		قلى
27	Tanween		تنوين

3.2.2. PREPROCESSING

Al-Mushaf is a printed version of a hand-written text using Arabic calligraphy art and printed on both sides of the paper. Sometimes the back-paper shadows may be visible from the front page and its effect should be removed. The RGB-saved images were read by the MATLAB program and were subjected to the preprocessing stage which is composed of three steps: image denoising using a median filter, RGB to grey converting, and image binarizing using a static thresholding technique. The median filter removes the pepper and salt noise, while a static threshold was applied to remove any remaining background noise and shadows from the back page text, resulting in a clear binary (black and white) image. Finally, the resulting images were saved as a subsidiary file to be used for further analysis. All the DaDs images were saved using the Tagged Image File Format (.tif) and named Diacr (n) where n is an indexing sequence number. Every image was standardized to 32 by 32 pixels, and the images for each class were saved in a separate subfolder.

3.2.3. LABELING AND SEGMENTATION

After preprocessing, the objects forming the image were labelled and separated based on their connectivity. Each group of connected pixels was assigned a unique index number, allowing the image to be converted into segments that could be easily identified by their index number. The labelling started in ascending order from the image's top-right corner to the bottom-left corner, and they could be colored for clarity. At this stage, the image contained two types of segmented objects: the grapheme and the DaDs, as illustrated in Fig. 8. Then, the grapheme should be identified and separated from the rest of the objects.

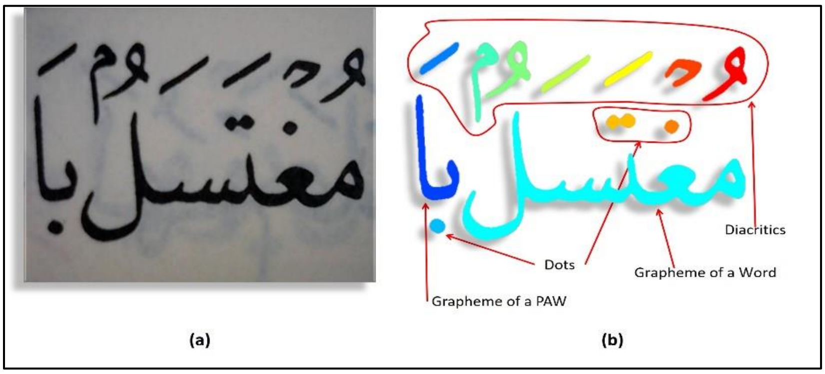


Fig. 8. (a) An input RGB-image. (b) The same image after preprocessing, labeling, segmentation, and coloring according to the labeling index.

3.2.4. GRAPHEME REMOVAL (DaDs EXTRACTION)

Grapheme separation was done based on the horizontal projection trajectory (HPT) technique (Boubaker, Tagougui, et al., 2014)(Al-Shatnawi and Omar, 2008). Equation (1) was used to determine HPT for the image, row by row. It is simple and direct but very effective; for every row in the image, the sum of the pixel values is computed. Since the DaDs are dispersed over multiple lines but the grapheme origin baseline in the Arabic text lies on the same horizontal line, the HPT will maximize around the grapheme origin line.

$$HPT(r) = \sum_{c=1}^{c=Lc} Pix(r,c) \quad (1)$$

where: Pix denotes the intensity of the pixel.

r for row, c for column of the pixel,

Lc denotes the last image's column.

The grapheme was defined as an item whose horizontal center line has the greatest HPT value and at least a portion of it lies in a certain window. After empirically selecting a window size of 20 rows, the grapheme is separated from the image, resulting in two distinct groups- DaDs-alone-group and grapheme-only-group. Fig. 9 displays the HPT curve as a front layer over a sample image that was generated from it. The grapheme baseline indicates where the HPT peak is situated. Finally, the grapheme is removed, resulting in an image of the separated DaDs as demonstrated in the same figure as an example.

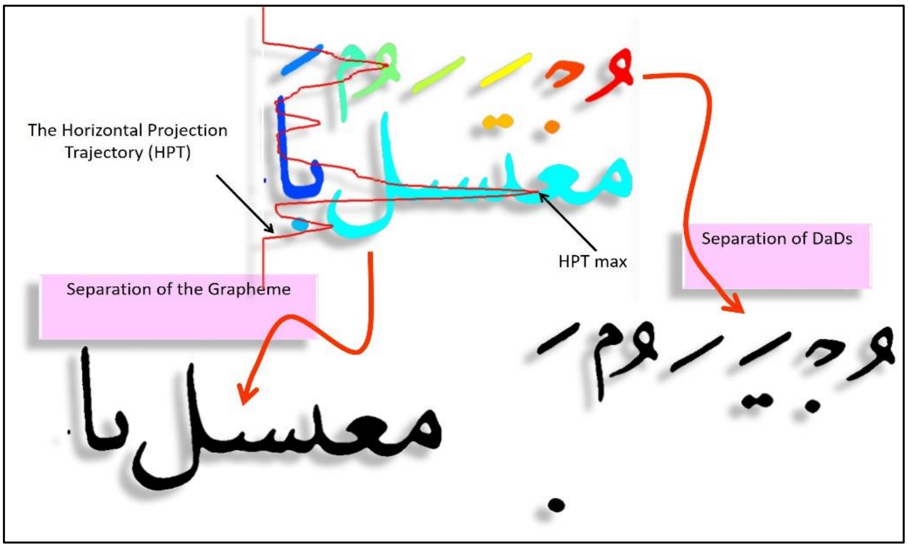


Fig. 9. The grapheme separation using the Horizontal Projection Trajectory (HPT) after image labeling and segmentation.

3.2.5. FALLS ALARM REJECTION and DaDs SEGMENTATION

During the extraction of DaDs from the image, it's possible that some incomplete DaDs or small noise objects were unable to be removed during the preprocessing stage. Any incomplete DaDs that intersected with the image's border were rejected, and small objects with an area below the minimum acceptable size for DaDs were eliminated. As a result, the image only contains the fully formed DaDs. To segment each DaDs-object into a separate image, the bounding box technique was utilized, and the resulting images were saved in a file.

3.2.6. DaDs RESIZING

Due to the distinct Al-Mushaf calligraphy, the segmented DaDs are extracted in different sizes. The last phase of the MATLAB program entails adjusting the DaDs' images to a uniform size of 32x32 pixels and saving them to a file. This resizing step is essential in expediting image dataset processing time, albeit resulting in minor shape distortion. Nevertheless, this distortion is negligible as the same procedure will be employed during the classification phase of the finalized system.

3.2.7. DaDs SORTING and SAVING

The downscaled DaDs images were saved in the order they were captured, resulting in over 6000 images. Only 4710 images were chosen for the final dataset after rejecting identical and non-DaDs images. Non-DaDs images were excluded manually, as they might be non-connected letters of Arabic text that look like some DaDs but are actually different. The last step involved sorting the DaDs images based on their classes, which were determined through ground truth labeling. They were then grouped and saved accordingly.

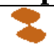
4. RESULTS AND DISCUSSIONS

This work has resulted in the generation of the first image dataset for the DaDs, used in Al-Mushaf text, extracted from two paper editions of Al-Mushaf utilizing a handheld scanner and methods tailored for this purpose. The dataset consists of a main folder and contains 22 subfolders, where each subfolder holds the images of one class of the DaDs. Each subfolder is named using the name of that class. The total number of DaD's images is 4710. Fig. 10 shows samples from two subfolders. Table 2 explains the distribution of the DaDs over different classes. The number of samples per class depends on how many times that class appears in Al-

Fig. 10. Samples from two folders of the generated dataset.

Mushaf. The number of classes has been reduced from 27, that were shown in Table 1 to only 22. The reduction is due to a proposed tactic that was benefited of assembling some DaDs from others; the three Dots can be assembled from the single Dot and the double Dots, the Dammatan from two Damma objects, Fathatan from two Fatha, Tanween Fathh from two Fatha, and Tanween Damm 1 from Damma and Tanween.

Table 2. The contents of the resultant dataset; 4710 images of 22 classes distribution.

Class No.	Class Shape	No. of Samples	Class No.	Class Shape	No. of Samples
1		500	**12**		100
2		500	**13**		100
3		250	**14**		100
4		400	**15**		100
5		400	**16**		150
6		150	**17**		100
7		150	**18**		40
8		500	**19**		100
9		350	**20**		110
10		150	**21**		110
11		200	**22**		150

5. CONCLUSION

Using two paper editions of Al-Mushaf, a DaDs image dataset was created. This paper presented the tools and the methodology for the generation of the dataset. The methodology is simple and mostly dependent on software, making it easy to use for enhancing the dataset. It is the first dataset for Al-Mushaf DaDs, and it is essential for any research on Al-Mushaf image processing and optical character recognition, especially when machine learning and artificial intelligence are needed. The images are organized in a way that allows for direct training and validation of a supervised ML classifier. In the future, the dataset can be expanded to include the DaDs of other Rewayat, which contain different types of diacritics than those used in Rewayat Hafs.

6. CONFLICT OF INTEREST

This work is part of the author's PhD research presented to the Electrical Engineering Department at the University of Mosul.

7. REFERENCES

Abdallah Abualkishik and Khairuddin Omar (2009) ‘Quranic Braille System’, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 2(10), pp. 3306–3312. Available at: http://www.waset.org/publications/7971.

Abualkishik, A. and Omar, K. (2013) 'Framework for translating the Holy Quran and its reciting rules to Braille code', International Conference on Research and Innovation in Information Systems, ICRIIS, 2013, pp. 380–385. doi: 10.1109/ICRIIS.2013.6716740.

Abualkishik, A. M. and Omar, K. (2009) 'Quran Vibrations in Braille Code Department of System Science and Management', in International Conference on Electrical Engineering and Informatics, pp. 12–17.

Alheraki, M., Al-Matham, R. and Al-Khalifa, H. (2023) 'Handwritten Arabic Character Recognition for Children Writing Using Convolutional Neural Network and Stroke Identification', Human-Centric Intelligent Systems. Springer Netherlands, 3(2), pp. 147–159. doi: 10.1007/s44230-023-00024-4.

Al-Shatnawi, A. and Omar, K. (2008) 'Methods of Arabic Language Baseline Detection – The State of Art', Journal of Computer Science, 8(10), pp. 137–143.

Altwaijry, N. and Al-Turaiki, I. (2021) 'Arabic handwriting recognition system using convolutional neural network', Neural Computing and Applications. Springer London, 33(7), pp. 2249–2261. doi: 10.1007/s00521-020-05070-8.

Bach-y-Rita, P. and W. Kercel, S. (2003) 'Sensory substitution and the human–machine interface', Trends in Cognitive Sciences, 7(12), pp. 541–546. doi: 10.1016/j.tics.2003.10.013.

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. Nature, 221(5184), 963-964.

Bach-y-Rita, P., Tyler, M. E. and Kaczmarek, K. A. (2003) 'Seeing with the brain', International Journal of Human-Computer Interaction, 15(2), pp. 285–295. doi: 10.1207/S15327590IJHC1502_6.

Boubaker, H., Chaabouni, A., Halima, M. B., El Baati, A., & El Abed, H. (2014, August). Arabic diacritics detection and fuzzy representation for segmented handwriting graphemes modeling. In 2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR) (pp. 71-76). IEEE.

Boubaker, H., Tagougui, N., El Abed, H., Kherallah, M., & Alimi, A. M. (2014). Graphemes segmentation for arabic online handwriting modeling. Journal of information processing systems, 10(4), 503-522.

Elkhayati, M., Elkettani, Y. and Mourchid, M. (2022) 'Segmentation of Handwritten Arabic Graphemes Using a Directed Convolutional Neural Network and Mathematical Morphology

Operations', Pattern Recognition. Elsevier Ltd, 122, p. 108288. doi: 10.1016/j.patcog.2021.108288.

Eraqi, H. M. and Abdelazeem, S. (2012) 'A new efficient graphemes segmentation technique for offline Arabic handwriting', in Proceedings - International Workshop on Frontiers in Handwriting Recognition, IWFHR, pp. 95–100. doi: 10.1109/ICFHR.2012.162.

Eraqi, H. M. and Azeem, S. A. (2011) 'An on-line Arabic handwriting recognition system: Based on a new on-line graphemes Segmentation Technique', in Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, pp. 409–413. doi: 10.1109/ICDAR.2011.90.

Fadel, A., Tuffaha, I., & Al-Ayyoub, M. (2019, May). Arabic text diacritization using deep neural networks. In 2019 2nd international conference on computer applications & information security (ICCAIS) (pp. 1-7). IEEE.

Kaczmarek, K. A., Webster, J. G., Bach-y-Rita, P., & Tompkins, W. J. (1991). Electrotactile and vibrotactile displays for sensory substitution systems. IEEE transactions on biomedical engineering, 38(1), 1-16.

Kef, M., Chergui, L. and Chikhi, S. (2016) 'A novel fuzzy approach for handwritten Arabic character recognition', Pattern Analysis and Applications, 19(4), pp. 1041–1056. doi: 10.1007/s10044-015-0500-4.

Madhfar, M. A. H. and Qamar, A. M. (2021) 'Effective Deep Learning Models for Automatic Diacritization of Arabic Text', IEEE Access, 9, pp. 273–288. doi: 10.1109/ACCESS.2020.3041676.

Maghraby, A. and Samkari, E. (2023) 'Arabic Text Recognition with Harakat Using Deep Learning', International Journal of Computer Science and Network Security (IJCSNS), 23(1), pp. 41–45.

Mohammed, A. (2023) 'Arabic / English Handwritten Digits Recognition using MLPs ', Al-Rafidain Engineering Journal (AREJ), 28(2), pp. 252–260.

Mohd, M., Qamar, F., Al-Sheikh, I., & Salah, R. (2021). Quranic optical text recognition using deep learning models. IEEE Access, 9, 38318-38330.

Mousa, M., Sayed, M. and Abdalla, I. (2013) 'Arabic Character Segmentation Using Projection-Based Approach with Profile' s Amplitude Filter', International Conference of Information and Communication Technology (ICoICT), pp. 122–126.

Nazih, W. and Hifny, Y. (2022) 'Arabic Syntactic Diacritics Restoration Using BERT Models', Computational Intelligence and Neuroscience, 2022(1), p. 8 pages. doi: 10.1155/2022/3214255.

Rubini, R. and Setyawan, C. E. (2020) 'Inclusion Education: Learning Reading Arabic Language And Alquran For Blind', Al-Bidayah: Jurnal Pendidikan Dasar Islam, 11(2), pp. 330–345. doi: 10.14421/al-bidayah.v11i2.348.

Sulaiman, S., Rambli, D. R. A. and Zuki, F. S. M. (2015) 'Putting the tactile feedback to Quranic verses and tajweed rules', ARPN Journal of Engineering and Applied Sciences, 10(23), pp. 17996–18003.