Multimedia Tools and Applications
https://doi.org/10.1007/s11042-020-09825-2

Extending Gardiner's code for Hieroglyphic recognition and English mapping

Reham Elnabawy[1] • Rimon Elias[1] • Mohammed A.-M. Salem[1,2] • Slim Abdennadher[1]

Received: 20 May 2020 / Revised: 6 August 2020 / Accepted: 2 September 2020
Published online: 21 September 2020

Abstract
"Knowledge is power"… Writing is the main way to preserve the humanity knowledge across the ages. Therefore, an automatic and accurate mapping from ancient scripts to modern live language is a must to have. Such a system will support knowledge transfer, tourism and education. This article presents a new algorithm to segment and recognize the ancient Egyptian Hieroglyphs from images and produce the corresponding English meaning. The algorithm used image processing along with Optical Character Recognition (OCR). Then, the meaning behind the image containing Hieroglyphs is interpreted based on the context of the mapped English sentence. Gardiner's sign list is a standard list used to classify the Hieroglyphics symbols such that similar shapes are grouped in the same category. Hence, Hieroglyphics script is mapped to the English language. Hieroglyphics could be either read from left to right or from right to left based on the face orientation of the Hieroglyphic symbol. However, Gardiner's code does not help to automate the reading direction of the Hieroglyphics script. In this work, an extension of the list is proposed to resolve the ambiguity in the reading direction. The results of the segmentation and recognition of Hieroglyphics symbols are demonstrated and compared to similar work for Chinese character recognition. Moreover, the results obtained from the state-of-the-art used in Hieroglyphic character recognition compared to the results obtained from the proposed algorithm on Hieroglyphic character are addressed. The mapped English sentence is then compared to some defined patterns. When a match is found, the input gets structured and reformatted accordingly. Tests on the defined patterns were conducted, and the results were successful, however, some additional results are generated that show equivalent synonyms to the input English sentence.

Keywords Optical character recognition (OCR) · Hieroglyphics · Gardiner's code · Recognition accuracy · English mapping

✉ Reham Elnabawy
rehamhossam312@gmail.com

Extended author information available on the last page of the article

1 Introduction

The inspiration driving this article was originated from the way that numerous individuals are keen on learning Hieroglyphic language to know how the old pharaohs developed their civilization and to know the historical background of Egypt. The issue emerges from the way that numerous voyagers while visiting the monumental sights are enthusiastic about understanding the pictographs composed on the walls freely by different methods as opposed to asking a tour guide. So it was thought to design a new algorithm to enable any language of symbolic representation such as Chinese language or Hieroglyphics to be easily recognized to a well-known language such as English.

The problem addressed in this article is: *"How to maintain the reading order of Hieroglyphics script and produce an accurate English meaning based on image processing techniques?"* The main objective behind this article is to recognize Hieroglyphs and map them to English so that the meaning behind a certain Hieroglyphic image could be accurately obtained. Hieroglyph is taken from the Greek hiero *'heavenly'* and glypho *'composing'*. Hieroglyph is a character that is generally utilized in pictorial composition. There are two kinds of Hieroglyphs: Ideograms and Phonograms. The Ideograms are pictures that speak to what the item communicates. On contrary, Phonograms are pictures that portray the old Egyptian language sound [1]. The principal fruitful symbolic representations decipherment was performed by Champollion, a French researcher, who deciphered the Rosetta stone, aka Rashid stone. This uncovered the historical background of the old Egyptians and their imagination in living with no facilities however, they had the capability to develop an exceptional progress which is loaded with riddles that are scanned for until today [25].

In old time, the Greeks and the Romans were astounded by the Hieroglyphic language. Researchers like Pythagoras were propelled by the inventiveness and advancement the antiquated Egyptians utilized [10]. Going before 3000 BC, Egyptian language seemed first recorded as a hard copy and continued being overwhelming in Egypt until the eleventh century AD where the Coptic showed up after the Hieroglyphics appearance [27]. A while later, the Arabic language progressively spread and turned into the transcendent language in Egypt. Right now, the antiquated Egyptian language is never again utilized and it is viewed as a dead language [2]. Some Hieroglyphic images have different implications.

The ideal significance or sound is resolved dependent on the situation of the Hieroglyphic image in the succession that will be read [28]. The normal reading criteria of the Hieroglyphic letters is critical and it ought to be read from upper till lower Hieroglyph. On the off chance that there is more than one pictograph in the same row, at that point, a unique criteria is pursued which is reading depending on the face direction of the Hieroglyphic image [16]. The nineteenth century American writing was influenced by the decipherment of the Egyptian Hieroglyphics [29]. Hieroglyphic language points to ideas and thoughts that give proximate view to the concealed significance behind the pictorial portrayals [4]. Amid Egypt intrusion in 1798, Napoleon's researchers were bewildered and enchanted to find antiquated sanctuaries and tombs [3]. Further-more, in [5], the old Egyptians messages in the pyramid were gathered and appeared. Moreover, chronicled sentence structure of the old Egyptians is acquainted to accomplish the interests of the language specialists and Egyptologists [11]. Another principle tended to theme is image segmentation. In [23], segmentation depends on isolating the frontal area from the foundation. Channels can be utilized to evacuate any boisterous parts in the picture with the goal that the division will be much simpler [14]. Image segmentation is viewed as a very critical stage to get exact outcomes.

There are various strategies that can be utilized to section a picture, for example, Active Contour Segmentation [21]. Chinese character acknowledgment is viewed as a standout amongst the most testing frameworks for character acknowledgment. The procedure pipeline for Chinese character acknowledgment is design portrayal, character order, learning/adjustment, and logical handling.

On contrary, when looking at the proposed pipeline in this article appeared in Fig. 1 to the state-of-the-art of the Chinese character acknowledgment, the *"post preparing"* stage in the proposed pipeline is excluded where the evacuation of clamor is an unquestionable requirement to get exact outcomes [7]. A correlation between the algorithms in the proposed algorithm alongside the state-of-the-art of the Chinese character acknowledgment connected on a Hieroglyphic picture in terms of division precision is appeared in the Section named *"Experimental Results"*.

This article is mainly focusing on Hieroglyphic recognition from the image processing perspective and does not take into consideration the deep learning perspective. The bottleneck of deep learning is the dataset, however, in Hieroglyphic character recognition, there is no available dataset for training the deep learning models. Thus, image segmentation techniques are followed to recognize the Hieroglyphic language.

Papers [6, 13, 15, 17–19, 22, 24, 31, 33] show some detection and recognition techniques used in image processing and image enhancement in general but not in Hieroglyphic character recognition. Further illustration is shown below.

Leng et al.'s [19] main aim was an automatic feces detection and trait recognition system based on a visual sensor that could greatly alleviate the burden on medical inspectors and overcome many sanitation problems, such as infections. Leng et al. [19] proposed light-weight practical framework that contains three stages: illumination normalization, feces detection, and trait recognition. The segmentation scheme used was free from training and labeling. The feces object is accurately detected with a well-designed threshold-based segmentation scheme on the selected color component to reduce the background disturbance.

Yang et al. [31] used advanced digital image processing technologies and deep learning methods which are employed for the automatic color classification of stool images.

Leng et al. [17] used three processes to enhance discriminant ability in DCT domain, and the relationship between them are summarized and discussed systematically. Leng et al. [17]

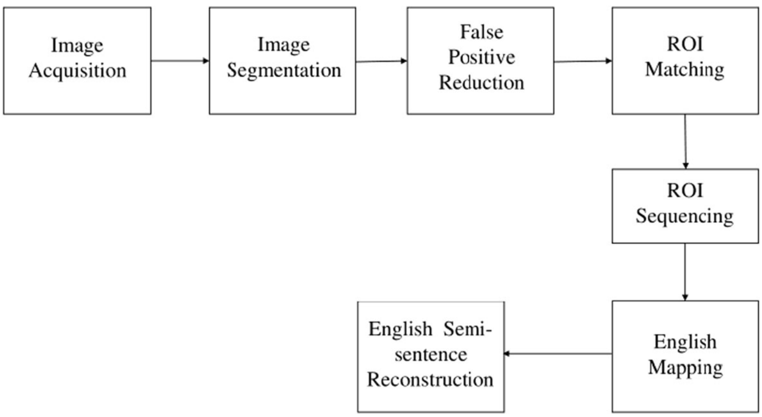


Fig. 1 Abstract block diagram for the proposed algorithm

proposed dynamic weighted discrimination power analysis (DWDPA) to enhance the discrimination power (DP) of the selected DCT coefficients.

Leng et al. [18] extracted the features of left and right palmprints with two-dimensional discrete cosine transform (2DDCT) to constitute a dual-source space. Normalization was utilized in dual-source space to avoid the disturbance caused by the coefficients with large absolute values. Thus, complicated pre-masking was needless and arbitrary removing of discriminative coefficients was avoided.

Zhang et al. [33] proposed Mask-Refined R-CNN (MR R-CNN), in which the stride of ROIAlign (region of interest align) was adjusted. Furthermore, the original fully convolutional layer is replaced with a new semantic segmentation layer that realizes feature fusion by constructing a feature pyramid network and summing the forward and backward transmissions of feature maps of the same resolution. The segmentation accuracy is substantially improved by combining the feature layers that focus on the global and detailed information.

Chu et al. [6] proposed an effective object detection algorithm for the small and occluded objects, which was based on multi-layer convolution feature fusion (MCFF) and online hard example mining (OHEM).

Jindal et al. [13] proposed an algorithmic approach that was used to zoom a given image in wavelet domain and to get a sharper image using various interpolation techniques.

Mander and Jindal [22] focused on a compression technique known as Block Truncation Coding (BTC) as it helped in reducing the size of the image so that it took less space in memory and easy to transmit. Thus, BTC was used to compress grayscale images.

Kaur and Jindal [15] focused on watermarking techniques. The techniques of watermarking had their advantages and outperformed better when combined with wavelets transformations (DWT) followed by interpolations.

Mittal and Jindal [24] presented a proposed algorithm that was adopted to enhance a given original input image in the domain of wavelets and the results have been proved with the help of PSNR values. The proposed algorithm was used further for contrast equalized images providing improvement in PSNR values and enhancement in images.

The state-of-the-art of Hieroglyphic character recognition was done as an automatic Egyptian Hieroglyph recognition by retrieving images as texts. It was taken from the computer vision and image processing perspectives however, with different approaches and techniques manipulated compared to the techniques used in the proposed algorithm in this article. Also, the state-of-the-art algorithm did not solve the ambiguity in the reading direction in the Gardiner's sign list which was, on contrary, addressed in this article's proposed algorithm [9]. Last but not least, the image matching techniques used in the state-of-the-art were RANSAC, Bag Of Words (BOW) and Single Descriptor, as mentioned in [9], which consumed a lot of time to find the best match compared to the HOG used on top of Euclidean distance in the proposed algorithm in this article. Further details about the state-of-the-art algorithm for Hieroglyphic character recognition is illustrated in section 3.

2 Proposed algorithm

Figure 1 contains the abstract block diagram for the proposed algorithm that shows the new algorithm proposed in this article. Since the way that the most habitually utilized configuration of the Hieroglyphics pictures found in momentous sights contains vertical lines as a division between each two Hieroglyphic sections, this article managed this configuration in the majority of its information pictures.

On the other hand, there are different arrangements of the Hieroglyphics pictures that could contain symbolic representations in isolated lines rather than sections. Additionally, there could be pictures that appeared as a square without the presence of any vertical or flat line. As a general rule, not every one of the pictures are arranged vertically upwards as there could be a few pictures that are put at a point making the caught picture be slanted. So in like manner, to mull over this tendency, a change is done toward the start of the calculation to pivot the picture with the goal that it shows up vertically upwards.

2.1 Image acquisition

A picture that contains pictographs is entered as an input as appeared in Fig. 2. At that point, Canny edge detection is performed on this picture to decide the edges in the picture as appeared in Fig. 2. From that point onward, Hough transform is connected to decide the absolute number of vertical lines in the picture so from Fig. 2, the entire number of vertical lines is 3. From this, the complete number of sections, which is C, is gotten by increasing the quantity of vertical lines in the picture, signified as V, as appeared in "$C=V+1$".

A line is signified as a vertical line in the event that it is arranged at 90 ° with regard to the width of the picture. The quantity of vertical lines is resolved as pursues: The height of the whole picture is determined and if the vertical line separated is inside a specific range (i.e., from a specific number till the whole picture stature), at that point, it will be considered a vertical line for the whole picture. At that point, by getting the width of one section in the picture through isolating the entire width of the whole picture by the quantity of vertical lines accessible, every segment will be cut independently and spared as a different picture to be prepared for the segmentation stage. In any case, in the event that the info picture is isolated into rows containing pictographs, at that point, this structure is adjusted by supplanting the vertical lines by the horizontal lines in the former advances where the edge of a horizontal line will be at 90 ° with regard to the whole picture height. In addition, the stature of a row is gotten

Fig. 2 Input image containing Hieroglyphs [8] and Canny edge detection applied respectively

by separating the entire height of the whole picture by the complete number of horizontal lines. This is considered the automatic separation of the column images in the entire image. A sample of the saved separated columns is shown in Fig. 3 where the leftmost column will be read first (i.e., the rightmost column in Fig. 2 since the face orientation of the Hieroglyphs is looking to their left side, then, the reading order will be from right to left) and then, the following columns will be read from the left to the right direction. By this, the columns reading order is determined.

2.2 Image segmentation

This is the most imperative stage where ordering is enforced to each Hieroglyphic character by taking a record dependent on the appearance sequence. This is gotten by playing out the accompanying advances. The initial step done is Canny edge detection which is acquired to get the edges, at that point, connected components labeling using 8-connectivity is utilized to give each region (i.e., extracted object) an interesting list that will be utilized for the segmentation intent. The segmentation is done by means of bounding box where the minimum and maximum boundaries of the width and the height of the Hieroglyph are detected to decide the limit that the object is occupying. At that point, a structure will be developed that has the base line picture pixel, most extreme line picture pixel, least segment picture pixel and greatest segment picture pixel as attributes. The bounding box of each separated district is spared in a list. Thus, the regions are arranged ascendingly dependent on the top edge of the bounding boxes (i.e., the base column pixel trait). Consequently, the objects in a single section are listed from top to bottom. Nonetheless, on the off chance that there is more than one Hieroglyph in the same row, at that point, in light of the face side of the pictograph, the reading sequence is resolved. The methodology which is utilized to get the face heading is acquired dependent on coordinating the picture that contains the symbolic representations with the informational index pictures that contain pictographs of various face directions (i.e., left and right looking faces) using HOG to get the best match. The dataset images are named by Gardiner's codes

Fig. 3 Input image divided into separated columns

[20]. Notwithstanding, so as to demonstrate the face diverse introductions for the symbolic representations that could be either looking towards right side or left side, for example, a fledgling like pictograph, a letter *"L"* is added to the Gardiner's code to demonstrate that the symbolic representation is looking towards its left side and a letter *"R"* is added to the Gardiner's code to demonstrate that the symbolic representation is looking towards its right side. For further illustration, if $x \in$ Hieroglyphic characters that have two orientations, then, let x_L represents a Hieroglyphic character with a face looking towards the left direction and x_R represents a Hieroglyphic character looking at the right direction. An example of the indices given to the symbolic representations is shown in Table 1.

Table 1 demonstrates the base and greatest traits for every one of the pictographs in Fig. 3 (i.e., the furthest left segment picture). As indicated in Table 1, the Hieroglyphic pictures are arranged ascendingly dependent on the top y-coordinate characteristic and every pictograph is given an exclusive index used for the reading objective. On the other hand, the fourth and fifth pictures are given a similar index which is 4. This indicates that the two pictures (i.e., symbolic representations) are located in the same row so consequently, they are both given a similar index. By this, every symbolic representation is given an exclusive index that will be utilized for the reading reason. The sequence of appearance is resolved where the pictographs of the same row will be coordinated with their reciprocals in the data set so that on the off chance that they are left looking, at that point, the furthest right symbolic representation will be read first and the succeeding symbolic representations will be read from right to left as appeared in Fig. 2. On contrary, in the event that they are right looking, at that point, the furthest left symbolic representation will be read first and the succeeding pictographs will be read from left to right.

Table 1 Index given to the extracted Hieroglyphs in Fig. 3, column 1

Image	Top y-coordinate	Bottom y-coordinate	Top x-coordinate	Bottom x-coordinate	Index
	2	36	15	25	1
	36	69	4	32	2
	72	85	2	36	3
	89	123	2	11	4
	92	121	19	35	4
	126	139	2	35	5
	130	142	2	24	6
	142	151	9	26	7
	152	184	3	31	8

Table 2 shows the Gardiner's code before adding the extension and after adding the extension where the extension denotes that all of the Hieroglyphic characters in Table 2 are looking to the left direction.

2.3 Post processing

After obtaining the Hieroglyphs in independent pictures, the pictures enter a post-processing stage to get just the region of interest (ROI). The purpose of ROI is to extract the region that totally occupies the Hieroglyphic character isolated from any other objects that are subset figures of the same Hieroglyphic character so that the following stages are done on the correct Hieroglyph. In the segmentation stage, in some cases there could be a picture that is a subset of another picture and it will be managed ordinarily as though it is a different picture where it will be mapped to its relating English language simply equivalent to different pictures. A subset picture could be effectively decided in the mapping stage where any two continuous rehashed symbolic representations having a similar Gardiner's code, implies that there are two adaptations of a similar picture (i.e., complete and subset pictures).

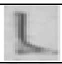
2.3.1 Segmentation enhancement

The primary stage is disregarding the little pictures (i.e., the pictures that are of moderately little size), treat them as clamor (i.e., denoising) and filter the picture to dispose of any clamor.

2.3.2 False positive reduction

The second stage is picture resizing resulting in having all the segmented Hieroglyphic pictures of a similar size to give increasingly exact outcomes when utilized in the matching stage. This is obtained by fixing a size for every one of the pictures which was 30 by 30. This size was likewise given to the pictures in the data set to guarantee the most ideal exactness when heading off to the matching stage. Any segmented picture that has a size not as much as that of a typical picture is denoted as a clamor and evacuated. A normal image is an image that

Table 2 Hieroglyphic characters' Gardiner's code

Hieroglyphic symbol	Gardiner's code before extension	Gardiner's code after extension
	F12	F12L
	G13	G13L
	F16	F16L
	H06	H06L
	F10	F10L
	F01	F01L

belongs to an approximation of a set of images. From here, came the importance of using False Positive Reduction technique and box plot. The outliers are any value greater than the third quartile or less than first quartile [30]. In the algorithm performed in this article, the data in the box were the area of the segmented images calculated (i.e., size of the segmented image) to be able to determine the outliers that refer to a small-sized segmented image [12]. An outlier is determined if an image (i.e., Hieroglyph) is of an area that is less than the first quartile or greater than the third quartile [26].

2.4 Image matching

The obtained pictograph will be taken independently and contrasted with the data set pictures utilizing HOG. HOG is a shape-extraction system that is utilized by ascertaining the Euclidean separation between the element vectors acquired from both the query and the data set pictures. Equation (1) demonstrates the Euclidean separation where $\overrightarrow{u}$ and $\overrightarrow{v}$ are the query and the data set images element vectors respectively. The HOG element vector incorporates the calculation of the quantity of pixels that add to a specific edge at a specific inclination. The best match is resolved in terms of the data set picture that is obtained from the minimum distance. At that point, after computing the Euclidean separation, standardize the outcome to show signs of improved results. Figure 4 shows a detailed flow chart for the image matching phase that includes the feature extraction technique used which is HOG along with the distance function used (i.e., Euclidean distance).

$$d\left(\overrightarrow{u},\overrightarrow{v}\right)=\left\|\overrightarrow{u}-\overrightarrow{v}\right\|=\sqrt{(u_1-v_1)^2+(u_2-v_2)^2+\ldots+(u_n-v_n)^2} \tag{1}$$

Concerning the way of knowing whether a Hieroglyph is looking towards the right side or the left side, the Hieroglyph, that should be mapped, is examined using HOG matching to get its equivalent matched image in the data set. Then, the data set image name (i.e., Gardiner's code) is saved for the usage in the English mapping phase. All the Gardiner's codes of the data set images consist of 3 characters preceding the *"L"* or *"R"* such as *M17L* and *Q03* where in case of having a one-digit number in the Gardiner's code, a *"0"* is placed preceding that digit just to generalize the case that *"L"* and *"R"* are always placed in the fourth position. In reality, the initial Gardiner's code for *Q03* is known as *Q3*.

2.5 English mapping

Subsequently, after retrieving the best match, the pictograph is converted into English with reference to a specific table that incorporates the Gardiner's codes [20] and their relating English interpretations. The mapped English sentence of Fig. 2 is *"The properties that God gave are to be useful like a beautiful soul and never destroy a desert or a palm, be a warrior and fight time to leave good reputation and wish your beloved ones in your prayers some good wishes."* Fig. 5 shows the Hieroglyphic word types that were used to grammatically format the English sentence correctly. In order to reformat and restructure the mapped English sentence to be mapped correctly in terms of grammar, five phases were passed by to get the desired output as shown in Fig. 6. The resulted English sentence obtained is a better format and structure in terms of grammar compared to the English sentence obtained from the last phase as the main aim in this article is to get a corresponding meaning of the Hieroglyphic sentence in English that better illustrates the context of the sentence that enables the reader to understand

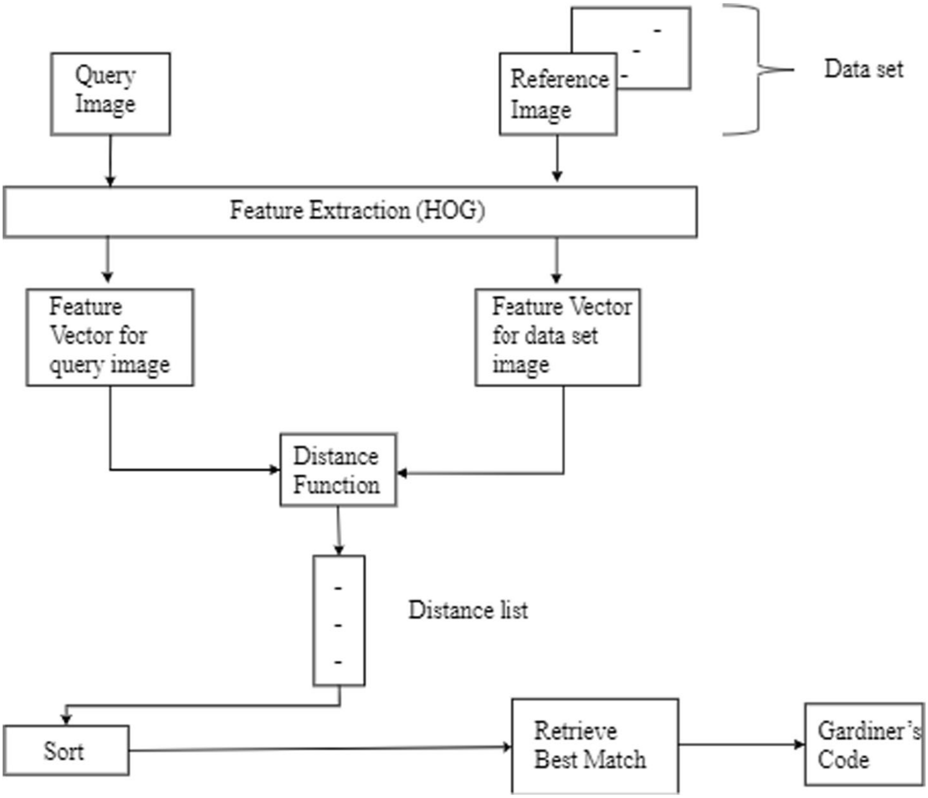


Fig. 4 Block diagram for image matching used in the algorithm

Hieroglyphs independently. As in English, and probably in all languages, a word does not necessarily have one meaning. An example is the word "*date*" in English, which could be used to describe the fruit or the time.

Furthermore, the type of a word is not necessarily the same, for instance the word *"leaves"* could be used as a plural noun or as a verb. Similarly, Hieroglyphic has words that witness the same behavior. Multiple mappings are found in the dataset, and are separated by a *'/'*. This means that for example, the word *"important"* has the following mappings: *"essential/great"*. The same goes for multiple types. For each single mapping, there should exist a corresponding type, so that the number of mappings of a word is equal to the number of types of the same word. This is to ensure handling each mapping independently in order to process a correct English sentence.

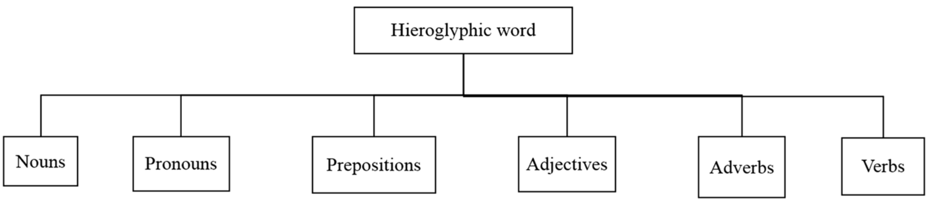


Fig. 5 Hieroglyphic word types

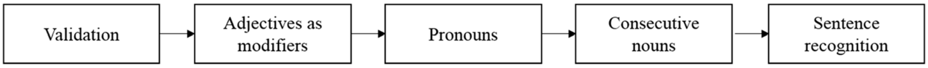


Fig. 6 Semi-sentence structuring and reformatting

In Fig. 6, the first phase is *"Validation"* phase. In this phase, since each word may have multiple mappings associated with multiple types, many incorrect sentences may be produced since some types should not fall after each other. For example, if a sentence consists of three words and each word has three mappings, then, the number of possible sentences is 27. However, some of these sentences could be grammatically incorrect and should be excluded. This is achieved in this phase as it checks whether it is correct for a certain word to fall after some other word or not. If it detects that some words cannot come together, then, the sentence is rejected. Concerning the data set of the words, it has around 12,000 words. Many of its words either did not have types or the types were not detailed enough. Therefore, these detailed types needed to be added manually to the data set. Most of the words are either verbs or nouns. So all other types, such as adverbs, were added manually, in addition to some of the nouns.

The rest of the data set remain without a type. The idea behind this phase came from the following: Imagine English had no spaces. For example, the sentence *"Thegladiatorisstrong"*. Trying to take *"t"* or *"th"* would fail, but taking *"the"* would succeed. Also taking *"the g"* would fail. Moving to the next word, *"glad"* and *"gladiator"* would succeed. But if *"glad"* was picked, it would ruin the rest of the sentence, so *"gladiator"* is picked then *"is"* and *"strong"* are picked. Hieroglyphics can be looked at in that same manner. Pretending *"iator"* is a word, if a scribe wanted to say *"glad"*, he/she would never write *"iator"* right after *"glad"* in order not to confuse readers. This is the main reason Egyptians thought a space is not necessary. What happens is that the output word is compared with the words in the data set that includes all the words classified into different types such as verb, noun, pronoun, adjectives and adverbs. Only those valid sentences can make it to the next phase. All valid sentences are ready to be structured. One at a time, valid sentences go through four phases representing different grammar rules, in order. In phase 2, which is *"Adjectives as modifiers"*, a valid sentence is taken without any previous alterations. It looks for adjectives modifying a noun. In English, when adjectives come as modifiers to nouns, they precede the noun, for example *"the excellent plan"*.

On the other hand, in Hieroglyphic, the adjective follows the noun, literally *"the plan excellent"*. There could be also more than one adjective following each other to describe a noun. The aim of this phase is to ignore all adjectives that act as modifiers. The noun is the more important word as it indicates what is being talked about. Imagine the noun was removed, *"the excellent"* would have no meaning, as the obvious question would be *"the excellent"* what? On the other hand, if the adjective was removed, *"the plan"* would be understandable even if it lost a property that the adjective provides. Therefore, what this phase tries to achieve is to combine a noun with its modifier adjective(s) and treat them as one word, which is a noun. The mappings are concatenated and their order is switched so that the noun follows the adjective(s). As for the type, there is no need to state the adjective now, hence the type of the combined word is just a noun. If no modifier adjectives were found, the sentence will proceed to the next phase without any changes.

On contrary, if modifier adjectives were present, some words will be combined as explained, and the sentence will proceed to the next phase with less number of words. In the next phase, *"Pronouns"* phase, a very common use of pronouns is with nouns. In English, when used with nouns, pronouns always precede nouns, for example, *"my book"*, *"this man"* and *"their house"*. In Hieroglyphic, there is no single rule for all pronouns. Some pronouns precede nouns, others follow nouns. In a similar manner as the previous phase, this phase tries to combine a noun with its corresponding pronoun, if any. Furthermore in case of a pronoun that follows a noun, the order is switched before the two words get combined. The result type of this combination is a noun. The effect of the previous phase can be seen when a combined word of a noun and its adjective(s) is encountered. For example, after combining *"plan"* and *"excellent"* in the previous phase and passing *"excellent plan"* to this phase, the next word could be *"this"*, which when found after a noun is combined with it after switching it with the noun resulting in *"this excellent plan"*. In the following phase, *"Consecutive nouns"* phase, the main aim is to look for any consecutive nouns and try to guess what kind of relationship exists between the nouns. If more than two nouns followed each other, then the program assumes this relationship as connection, since lists were common in Hieroglyphics. When exactly two nouns come together, it is hard to tell their relationship so to solve this problem, *"and/or/of"* is added between the nouns. Accordingly, when the user reads the output, they can understand which word they should pick from the context. In the final phase, *"Sentence recognition"* phase, the mapped and formatted sentence from the previous phases is obtained. There could be multiple mappings for a single sentence since every word could have multiple synonyms.

3 Experimental results

Table 3 includes the total number of hieroglyphic images (i.e., characters) that were segmented and processed from Fig. 2. In addition, it specifies the resolution of all the images which is unified through all segmented images as 30 × 30. Also, it indicates the total number of different classes that refer to the Gardiner's category given to the Hieroglyphic images.

Considering the variability and the availability of training/testing data sets, the robustness of unsupervised (i.e., rule-based techniques) are much higher than supervised techniques since the achieved accuracy does not rely on training data set. In general, accuracy of deep learning is significantly high and promising. However, success of this technique is relatively relying on the existence of appropriate dataset where there are no available datasets for training deep learning models for Hieroglyphic recognition. In Hieroglyphic character recognition, this challenge is not yet addressed and there is no adequate dataset for training. Therefore, the usage of conventional image segmentation and processing techniques are the most appropriate techniques for the Hieroglyphic character segmentation.

The mapped English sentence is reconstructed again at the end as discussed in section 2.5 to enable better format of a grammatically semi-correct sentence compared to the mapped English sentence directly obtained from the Hieroglyphic recognition. To approve the calculation executed in [32] versus the calculation actualized in the proposed Hieroglyph, the outcomes are presented in Table 4. The Chinese language is viewed as a comparative language

Table 3 Detailed overview for number of Hieroglyphic images used

Number of images	41 hieroglyphic images
Resolution	30 × 30
Number of classes	7 classes

to the Hieroglyphics because of the way that its characters are pictorial (i.e., symbolic) and they are having uncommon sounds when articulated. The state-of-the-art of Chinese character acknowledgment was obtained by methods for Neural Networks. The issue articulation was *"How to recognize a given information character to be spoken in English?"* By showing Neural Networks to change loads, the connection among sources of information and yields could be resolved effectively. Therefore, boisterous characters which are never observed are connected to Neural Networks for characterization. To take care of this order issue, preparing, refreshing loads by utilizing the contrasts between genuine yield and wanted yield and the rule were executed to stop Neural Networks preparing. Table 4 demonstrates the correlation between the calculation done in the proposed algorithm and the state-of-the-art of the Chinese character acknowledgment both on pictographs to make sense of the distinction between the two methodologies in terms of segmentation exactness (i.e., cutting the symbolic representations) and the accuracy of the relating English mapping. As appeared in Table 4, an example of segmented symbolic representations was tried to gauge the precision level of the segmentation which influences the accuracy of the English mapping.

The three segments of Table 4 are as per the following: Hieroglyphs Gardiner's codes, proposed algorithm actualized in this article and the work executed in the state of the art of Chinese character acknowledgment. In the principal passage, the proposed algorithm segmentation precision is 80% nonetheless, the state-of-the-art of segmentation exactness is 65% which is viewed as a low precision rate in respect to the proposed work. The mean of the precision rates of the proposed algorithm is 66.64% and that of the state-of-the-art is 55.27%. So this implies that the framework actualized in the proposed work gave more exactness and better outcomes contrasted with the state-of-the-art of Chinese character acknowledgment. This could be also proved from Fig. 7 where the left image is the original Hieroglyphic image that has been processed through the state-of-the-art algorithm used in Chinese character

Table 4 Statistics on different approaches for Hieroglyphs character recognition in terms of recognition accuracy

Image	Hieroglyph Gardiner's code	Proposed work	Zhang, X. Y. et. al [33]
	F12	80%	65%
	G13	73%	59%
	F16	33%	10%
	H06	60%	67%
	F10	65%	49%
	F01	52%	63%
	M03	76%	69%
	G37	87%	61%
	N41	54%	59%

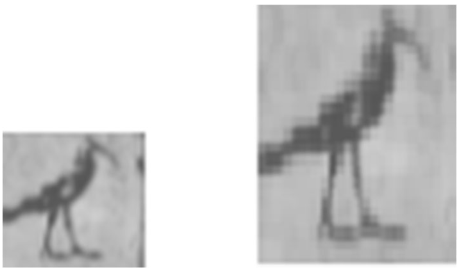
recognition and the right image is its corresponding output image where the output image is of low resolution and blurry as a result of the magnification performed by the algorithm which eliminates the recognition accuracy.

Other matching techniques were used before HOG. However, HOG gave the best results in terms of accuracy. Template matching technique was used where the data set image should be exactly the same as the query image which will not always be the case as there might be images of different orientations and different scale. That is why the accuracy percentage was extremely low when matching the data set images with a certain query image. Furthermore, the SURF matching technique was used which gave more precise and accurate results. In SURF matching, the more the data set size, the lower the accuracy of matching obtained. This is because of the fact that the data set image could be in a different orientation (e.g., rotation) rather than the query image which makes accuracy and precision in this aspect relatively low. The last matching technique tested was HOG which gave the best accuracy in terms of matching precision.

Concerning the error analysis behind the fact that there are some Hieroglyphs not 100% recognized correctly, the reason behind this is that there might be still some remaining noise or on coincidence, the Hieroglyph might not be fully occupying the bounding box where this case is rarely happening since that false positive reduction discards any falsely recognized Hieroglyph. A visual view of the error analysis in terms of recognition accuracy is shown in Fig. 8.

Concerning the state-of-the-art algorithm of the Hieroglyphic recognition, the following differences were manipulated compared to the proposed algorithm in this article. Instead of False Positive Reduction technique used in the proposed algorithm, texture synthesis and approximation were used to remove and discard any falsely detected shape as a Hieroglyph. The Texture Synthesis works by filling the gaps layer after layer from the outside in. Whereas, the approximation is sufficient in masking the undesired parts in the image such that the classification algorithm is able to train on them. This stage when tested compared to the proposed algorithm on the same set of Hieroglyphic character gave inaccurate results as, in case of having a Hieroglyph that its shape includes separated lines, it will be treated as if it is more than one Hieroglyph resulting in inaccurate recognition. Moreover, concerning the reading order maintenance, the state-of-the-art algorithm used the following criteria. The reading order is determined based on the Hieroglyph's position. This distinguishes horizontally and vertically aligned Hieroglyphs. For horizontally aligned Hieroglyphs, the x values are used to sort them, while for vertically aligned Hieroglyphs, the y values are used as shown and illustrated in Fig. 9. The approach shown in Fig. 9 does not include the usage of image matching to determine the reading order unlike what was proposed in this article that uses HOG on top of Euclidean distance to know the reading direction and to get the corresponding

Fig. 7 Sample of a Hieroglyph tested by the state-of-the-art algorithm used in Chinese character recognition



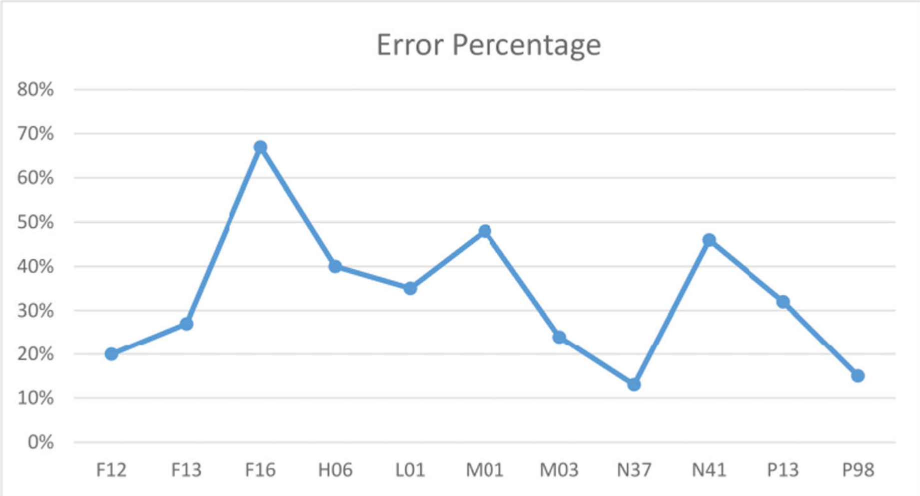
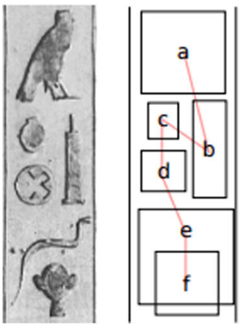


Fig. 8 Visual representation of error analysis

Gardiner's code at the same time which improves time efficiency compared to that in Fig. 9 that still did not get the Gardiner's code at that phase.

3.1 Image recognition evaluation

Concerning the measurement of the recognition accuracy of the manual segmentation versus the automatic segmentation, an approach called intersection over union was performed. Figure 10 shows an example of the areas of manual segmentation versus automatic segmentation respectively where there is an overlap between them as shown in the rightmost figure in Fig. 10. As shown in Fig. 10, the areas A_1 and A_2 are calculated for the manually segmented image and the automatically segmented image respectively. The intersection area over the union area percentage is calculated as shown in Eq. (2). As shown in Fig. 11, the percentage of the intersection area (i.e., the overlapping area) over the union area is calculated. The average percentage is 67.37%. Concerning the mapped English sentence, it is reformatted and structured correctly based on Fig. 6. Since that ancient Egyptians did not have as many words as people have nowadays in different languages, some words in Hieroglyphic, especially nouns, tend to have multiple interpretations. This fact leads to the generation of multiple sentences in



$a \to b$	No parts of a and b are overlapping, therefore they are ranked according their y-values (top to bottom).
$b \to c$	The center of c lies within the y-boundaries of b, therefore they are horizontally aligned. This means that b and c are ranked according their x-values (right to left).
$c \to d$	Same as for $a \to b$
$d \to e$	Same as for $a \to b$
$e \to f$	e is completely overshadowing f in the x-dimension, therefore they are not horizontally aligned and will also use the y-values to sort them.

Fig. 9 Reading order determination

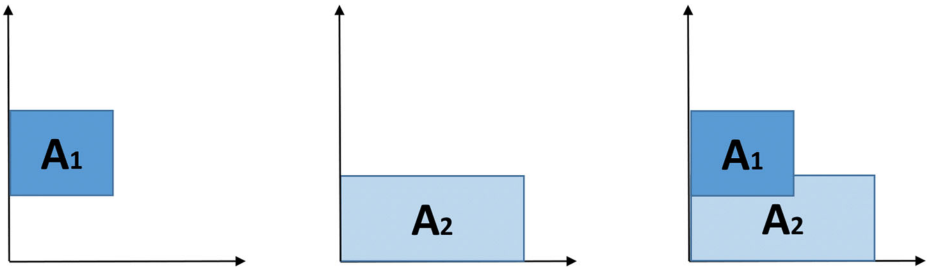


Fig. 10 Sample of Hieroglyph's area in its manual form and automatic form respectively along with the overlapping area

the result when translating an input that has a noun. So to avoid freezing the software, the results are limited.

$$\frac{A1 \cap A2}{(A1 \cup A2) - (A1 \cap A2)} \times 100 \tag{2}$$

4 Conclusion

This article has discussed the Hieroglyphic character recognition not only since it is needed for the knowledge transfer, but also, to facilitate the automatic applications used for tourism, history, education, etc. Recent related work on image segmentation and character recognition has been reviewed and discussed.

The main contribution is to facilitate the process of reading the Hieroglyphs correctly. Since that the Gardiner's code denotes only the symbolic shape of Hieroglyphs, then, a contribution in the Gardiner's code was done to not only determine the category of the Hieroglyph but also,

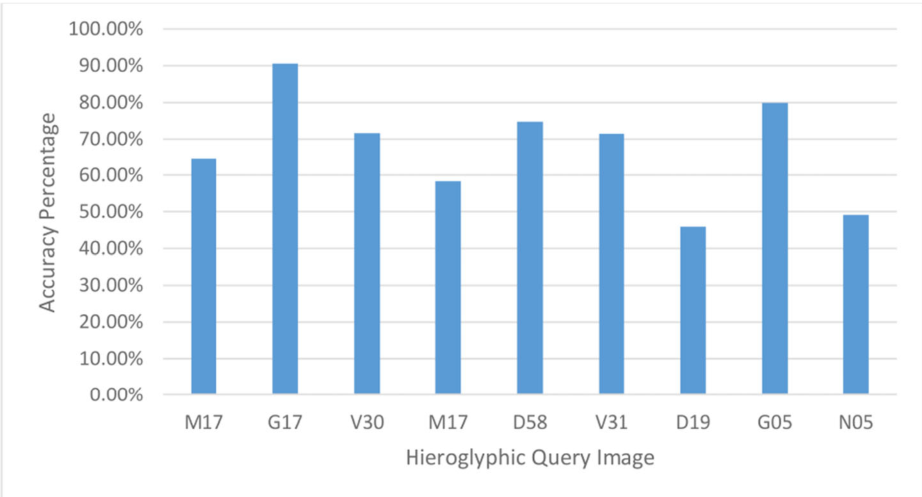


Fig. 11 Segmentation accuracy based on intersection-over-union

determine the face orientation of the Hieroglyphic symbol to denote the order of reading. By utilizing HOG, the query picture (i.e., the picture containing the symbolic representation) is contrasted with the data set pictures to locate the best match. This reference picture that signifies the best match, its Gardiner's code is known and a substring on the Gardiner's code is done on the fourth character (e.g., *"M17L.png"*) to determine whether the pictograph is right or left looking.

The resulted English mapping was processed in multiple stages to be structured and formatted correctly in terms of grammar. Each identified English word has a number of mappings and corresponding word types, which would generate many different combinations when mapped to English, some of which may be grammatically wrong. Hence, the input should be validated. This can be done by checking the types of the words. The algorithm checks if a certain type can follow some other type. If not, then all combinations that have these two consecutive types are rejected. The remaining combinations that make it through the validation phase, are ready to get mapped. Again by checking the types of the words, the algorithm looks for certain patterns that exist between those words. When a pattern is detected, the sentence gets mapped and structured accordingly, so that the output is a correct English phrase.

Data availability Data is available upon request.

Compliance with ethical standards

Conflict of interest Not available.

Code availability No software application, however, there is a custom code available upon request.

References

1. Allen JP (2000) Middle Egyptian: an introduction to the language and culture of hieroglyphs. Cambridge University Press, Cambridge
2. Allen JP (2013) The ancient Egyptian language: an historical study. Cambridge University Press, Cambridge
3. Bell DA (2018) Napoleon: a very short introduction. Oxford University Press, Oxford
4. Bowen T, Evans MM (2018) What does knowledge look like? Interpreting diagrams as contemporary hieroglyphics. Vis Commun
5. Chrysikopoulos V (2018) Defining Hidden in Ancient Egypt. In Proceedings of the 4th Biennial of Architectural and Urban Restoration. Host of the Itinerant Congress Hidden Cultural Heritage: Under Water, Under Ground and Within Buildings (p. 103). CICOP Italia
6. Chu J, Guo Z, Leng L (2018) Object detection based on multi-layer convolution feature fusion and online hard example mining. IEEE Access 6:19959–19967
7. Chung HK, Leung JC, Wong VM, Hsiao JH (2018) When is the right hemisphere holistic and when is it not? The case of Chinese character recognition. Cognition 178:50–56
8. Drucker J, McGann J (2000) Images as the text: pictographs and pictographic rhetoric. Info Design J 10(2): 95–106
9. Franken M, van Gemert JC (2013) Automatic Egyptian hieroglyph recognition by retrieving images as texts. In Proceedings of the 21st ACM international conference on Multimedia, pp. 765–768
10. Guthrie KS, Fideler DR (1987) The Pythagorean sourcebook and library: an anthology of ancient writings which relate to Pythagoras and Pythagorean philosophy. Red Wheel/Weiser, Newburyport
11. Hansen EJ (2018) Poe, Egypt, and "Egyptomania". In: Poe, Egypt, and "Egyptomania", In Poe and place (pp. 265–288). Palgrave Macmillan, Cham
12. Hintze JL, Nelson RD (1998) Violin plots: a box plot-density trace synergism. Am Stat 52(2):181–184

13. Jindal H, Kasana SS, Saxena S (2016) A novel image zooming technique using wavelet coefficients. In proceedings of the international conference on recent cognizance in Wireless Communication & Image Processing (pp. 1–7). Springer, New Delhi
14. Kashyap R, Tiwari V (2018) Active contours using global models for medical image segmentation. Int J Comput Syst Eng 4(2–3):195–201
15. Kaur S, Jindal H (2017) Enhanced image watermarking technique using wavelets and interpolation. Int J Image Graph Signal Process 11(7):23
16. Leach N (2005) The hieroglyphics of space: Reading and experiencing the modern metropolis. Routledge, Abingdon
17. Leng L, Zhang J, Xu J, Khan MK, Alghathbar K (2010) Dynamic weighted discrimination power analysis in DCT domain for face and palmprint recognition. In 2010 international conference on information and communication technology convergence (ICTC) (pp. 467-471). IEEE
18. Leng L, Li M, Kim C, Bi X (2017) Dual-source discrimination power analysis for multi-instance contactless palmprint recognition. Multimed Tools Appl 76(1):333–354
19. Leng L, Yang Z, Kim C, Zhang Y (2020) A light-weight practical framework for feces detection and trait recognition. Sensors 20(9):2644
20. Loprieno A (Ed.) (1996) Ancient Egyptian literature: history and forms (Vol. 10). Brill
21. Ma C, Luo G, Wang K (2018) Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of MR images. IEEE Trans Med Imaging 37(8):1943–1954
22. Mander K, Jindal H (2017) An improved image compression-decompression technique using block truncation and wavelets. Int J Image Graph Signal Process 9(8):17
23. Marmanis D, Schindler K, Wegner JD, Galliani S, Datcu M, Stilla U (2018) Classification with an edge: improving semantic image segmentation with boundary detection. ISPRS J Photogramm Remote Sens 135: 158–172
24. Mittal A, Jindal H (2017) Novelty in image reconstruction using DWT and CLAHE. Int J Image Graph Signal Process 9(5):28
25. Parkinson RB, Diffie W, Fischer M, Simpson RS (1999) Cracking codes: the Rosetta stone and decipherment. Univ of California Press, Berkeley
26. Pfannkuch M (2006) Comparing box plot distributions: a teacher's reasoning. Stat Educ Res J 5(2):27–45
27. Richter TS (2009) Greek, Coptic and the 'language of the Hijra': the rise and decline of the Coptic language in late antique and medieval Egypt. From Hellenism to Islam: Cultural and Linguistic Change in the Roman Near East, 401–46
28. Roth AM (2000) How to read Egyptian hieroglyphs: a step-by-step guide to teach yourself, by mark collier and bill Manley, illustrated by Richard Parkinson. 179 pages, 200 b/w illustrations. Berkeley: University of California Press, 1998. $18.95 (paper on board) ISBN 0-520-21597-4. Rev Middle East Stud 34(1):89–90
29. Schotter J (2018) Hieroglyphic modernisms: writing and new Media in the Twentieth Century. Edinburgh University Press, Edinburgh
30. Williamson DF, Parker RA, Kendrick JS (1989) The box plot: a simple visual method to interpret data. Ann Intern Med 110(11):916–921
31. Yang Z, Leng L, Kim BG (2019) StoolNet for color classification of stool medical images. Electronics 8(12):1464
32. Zhang XY, Yin F, Zhang YM, Liu CL, Bengio Y (2017) Drawing and recognizing Chinese characters with recurrent neural network. IEEE Trans Pattern Anal Mach Intell 40(4):849–862
33. Zhang Y, Chu J, Leng L, Miao J (2020) Mask-refined R-CNN: a network for refining object details in instance segmentation. Sensors 20(4):1010

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Reham Elnabawy[1] · Rimon Elias[1] · Mohammed A.-M. Salem[1,2] · Slim Abdennadher[1]

[1] Faculty of Media Engineering and Technology, German University in Cairo, Cairo, Egypt

[2] Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt