The Classification of Ancient Sumerian Characters using Convolutional Neural Network

Ahmed H. Y. Al-Noori[1][a], Moahaimen Talib[2][b] and Jamila Harbi S.[2][c]
[1]*Computer Engineering Department, Al-Nahrain University, Bagdad, Iraq*
[2]*Computer Science Department, Al-Mustansiriyah University, Baghdad, Iraq*

Keywords: Sumerian Characters, Convolutional Neural Networks, Deep Learning, Features Extraction.

Abstract: Recently, many sophisticated techniques have been used to classify ancient characters just like Phoenician, Egyptian hieroglyphs, and Maya glyphs. This paper proposed a new technique based on Convolutional Neural Network to classify (CNN) characters from Sumerian texts. The work was motivated by the challenges that faced with the status of Sumerian tablets which some had been broken and distorted by erosion factors. This technique includes taking the dataset of the Sumerian character's features and apply to these characters. Then, after initializing the weights for the output neurons, the layers of the CNN are prepared. Finally, the performance of this network is evaluated. This technique shows significant results and time-consuming.

1 INTRODUCTION

The Sumerians ancient, one of the greatest ancient of Mesopotamia, had commonly known for inventing the first civilizations in the world, building the first cities, invented the wheel, and use time units. Furthermore, they represent the first people that use writing and inventing the writing system. This system was used to organize their daily life to write a contract, buying, selling, agricultures, and enacting laws. This writing is known as **Cuneiform writing**.

Cuneiform writing was written on clay tablets and seals by using a wood instrument. The Sumerian written almost all their history and legends on these clays. For instant, the legends of their gods and the way the gods create a human being, the history of their heroes, and the achievements of their kings(Kramer, 1963). However, according to the Archaeologists, the ages of these tablets are return to 30 centuries ago (4500 BC)(Kassian, 2014). Throughout that history, these tablets are suffered from too many damages and erosion factors. That made most of these tablets are distorted and damage. For decades, many techniques have been used to extract and classified the characters and symbols in these tablets. The extraction of these characters is very important to recover the contents of these tablets since they are vulnerable to destruction and stolen.

Deep learning including **Convolutional Neural Networks** (CNN) represents one of the efficient techniques used in pattern recognition fields to recognize the data regularities and patterns. In the previous work (Talib & Harbi, 2017) Sumerian characters are extracted from these tablets and put their texture features into a dataset by using Discrete Wavelet Transform and Split Region Methods. In this work, the extracted dataset of the characters features to recognize these characters by applying CNN. This process is done by setting and initializing the parameters of the input characters then preparing layers of the CNN (input, convolution, subsampling layers) after initialize the weights for the output neurons. The demand to extract and understand the ancient texts (including cuneiform texts) gave attention to lots of researchers. Edan (2013) design an algorithm for recognizing cuneiform symbols. This algorithm is based mainly on K-mean to cluster the symbols. Then, multi-layer neural networks are applied to classify the symbol within the same cluster. Majeed, Beiji, Hiyam, and Jumana (2015) proposed a method based on the wavelet algorithm to obtain the text from the Sumerian clay tablet. Yang, Jin, Xie,

[a] https://orcid.org/0000-0002-0540-6849
[b] https://orcid.org/0000-0002-5106-849X
[c] https://orcid.org/0000-0003-1872-160X

and Feng (2015) proposed an approach to enhance Deep Convolutional Neural Network to recognize handwritten Chinese characters. This enhancement including deformation, non-linear normalization, imaginary strokes, path signature, and 8-directional features. SASAKI, HORIUCHI, and KATO (2015) improve the system for recognizing ancient Japanese characters in order to read ancient documents. They CNN to extract the features and used a support vector machine (SVM) to classify these features. On the other hand,(Tsai, 2016) used a deep convolutional neural network to classified the three different types of scripts of handwritten Japanese. This work focuses on the classification of the type of script, character recognition within each type of script, and character recognition across all three types of scripts.

2 WHAT IS DEEP LEARNING?

Deep Learning is a field of machine learning that is used to learn computers to do what can humans able to do in real life. Exactly just like, when a person understand and learn from his/her expertise. Conventional machine learning has algorithms that use the computations to learn from data or information directly without depending on predefined equations to be used as a model. Machine learning algorithms are widely used in pattern recognition fields. For instant, face recognition, recognize texts from audio or videos (speech recognition), and recognize car numbers from its plates (O'Shea & Nash, 2015). Furthermore, they are used in smart technology such as auto-driving of the cars which used to detecting lights and people crossing the street. Deep Learning (DL) represents the most efficient technique since it is providing better performance than other machine learning algorithms based on the experimental results. The main behind this is that DL mimics brain functions. Furthermore, its methods include multi-layer processing, which can give better time consuming and high accuracy performance. Sub-sampling layers give better results, by using CNN and auto-encoders when their number increased then better timing and clarity for the images are obtained (Hijazi, Kumar, & Rowen, 2015). In general, there are three important types of neural networks that form the basis for most pre-trained models in deep learning: Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), and Convolution Neural Networks (CNN). In the next section, CNN will be discussed in detail, which is mainly used in this work.

3 CONVOLUTION NEURAL NETWORKS (CNN)

A CNN is known as a special type of neural network and a main branch of the DL. It represents the best choice for pattern recognition and specifically in the image processing area. . A CNN is comprised of one or more convolutional layers, also sometimes it contains a subsampling layer and after the subsampling layer, there are one or more fully connected layers just like the conventional neural network (Patterson & Gibson, 2017). The Design of CNN is based on the mechanism of the visual of the human, i.e. the cortex of the visually in the brain of a human. Many cells exist in the cortex of the visually, these cells have the job of detecting light in sub-regions that are small or overlapped in the visual area for the human eye. These areas are known as receptive areas and the cells on its work as local filters for input space. Moreover, if the cells have higher complexity then they will have larger receptive areas. CNN’s convolution layer represented the function that is implemented by the visual cortex’s cells (Patterson & Gibson, 2017).

3.1 Advantage of CNN

Recently, CNN has taken the attention of many researchers, especially in image processing fields because of the advantages include within it. These advantages can be summarized as follows(Hijazi et al., 2015):

1. **Ruggedness to Shifts and Distortion in the Image:** the detection with using CNN is rugged to distortions for example change in shape due to camera lens, different lighting conditions, different poses, presence of partial occlusions, horizontal and vertical shifts, etc.
2. **Fewer Memory Requirements:** Theoretically, the fully connected layers can be used to have all the features to be extracted, for example, If an image of size 32×32 with a hidden that has 1000 features, then 106 coefficients order is needed, with a very large memory needed. However, these coefficients will be used in several locations across space in the convolutional layer, this will lead to reducing the memory usage drastically.
3. **Easier and Better Training:** In CNN, the number of parameters is reduced drastically. Which make CNN better time consuming when compared with the traditional neural network. In

addition, if a neural network is built and try to make it equivalent to CNN, then the standard neural network may have more noise while training due it has parameters more than the CNN, and its performance is less than the CNN.

3.2 CNN Architecture

In general, CNN consist of three different layer types. These layers are convolution, pooling, and fully connected layers. If the mentioned layers are stacked together then the CNN architecture has been formed. Figure 1 illustrates the main layers of simple CNN.

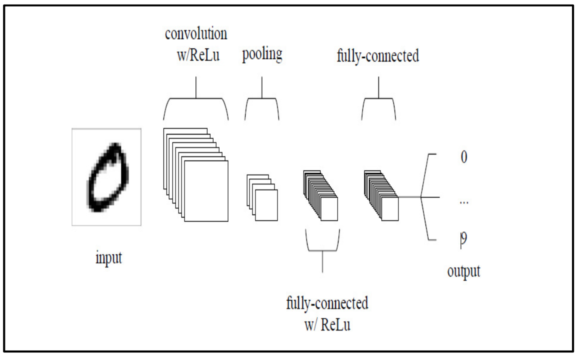

Figure 1: Layers of simple CNN architecture.

The basic functionality of this CNN can be classified into the following areas:

1. As found in other forms of Neural Network, the input layer will take the pixel values of the image.
2. The convolutional layer will determine the output of neurons of which are connected to local regions of the input through the calculation of the scalar product between their weights and the region connected to the input volume. The rectified linear unit (ReLu) aims to apply an ’elementwise’ activation function such as sigmoid to the output of the activation produced by the previous layer.
3. The pooling layer will then simply perform down sampling along the spatial dimensionality of the given input. Furthermore, the number of parameters within that activation is reduced.
4. The fully-connected layers will then perform the same duties found in standard Neural Networks and attempt to produce class scores from the activations, in order to be used for classification. Also, it is suggested that ReLu can be used between these layers to improve performance.

4 SUMERIAN CHARACTERS CLASSIFICATION USING CNN

As mention before, the CNN technique is very efficient in pattern recognition fields in general and especially what is concerned with image processing. It can perform effective work for classifying the extracted characters according to their features. Furthermore, contrary to other machine learning, CNN has the ability to extract the features from a character without the requirement to feature extraction pre-process. For that reason, and based on their efficiency in classification, CNN has been chosen in work.

In order to help with classifying the extracted Sumerian characters, the following algorithm is presented:

1. **Creates and Initializes All of the Parameters for a CNN:** The layers of the CNN should be initialized and prepared. To do that, the structure array containing three layers (The Input layer, convolution layer, and the subsampling layer0 must be prepared for this purpose.
2. **Perform an Evaluation of the Current Network on the Training Batch:** After finishing the setting of the parameters, the size of each character must be stored in a matrix. Then, create a structure array that includes variables that need to be used by the layers of CNN.
3. **Calculate Gradients using Back-propagation:** Creating some fields in the array of the CNN. These fields will be used in calculating the gradients. This is required to define the outputs for each field and consider that only the convolution layer will have a sigma function value. Then, taking the size of one field value of the structure of the CNN, and multiplying it with later field value. Finally rotating the output value to be used in the next step.
4. **Update the Parameters by Applying the Gradients:** Checking each convolutional layer and update each field in it. Then, the feed-forward field and backward fields are updated.

Figure 2 illustrates the main steps of the proposed algorithm:

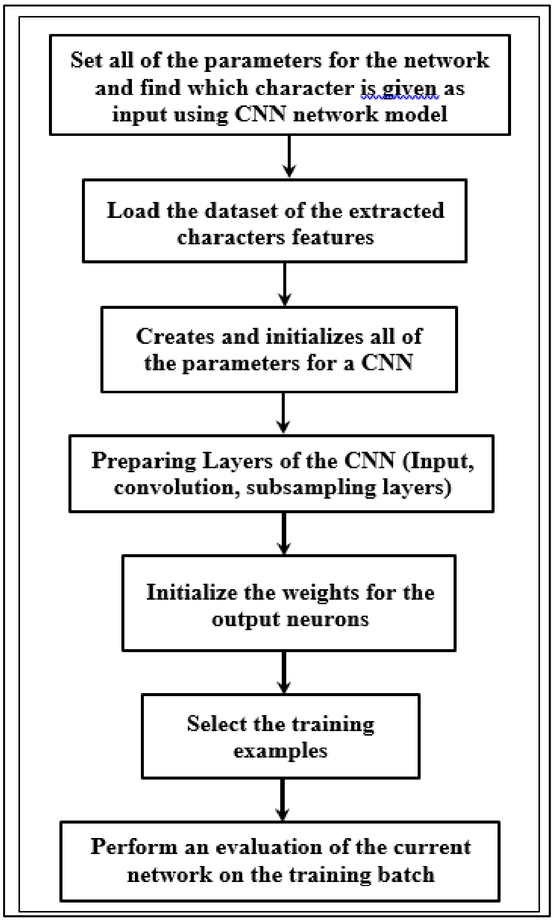

Figure 2: The proposed algorithm for Sumerian characters extractions.

5 EXPERIMENTAL RESULTS

5.1 Setup and Evaluation

The proposed algorithm was tested by using Matlab 2015b. The Sumerian characters are extracted from images of 20 tablets. These images are collected from the Cuneiform Digital Library Initiative (CDLI) at Cornell University(cdli, 2017) as demonstrated in Figure 3. According to this figure, the first image represents the input image of the tablet. The second and third represent the wavelet and inverse wavelet images respectively. The fifth image represents the output of the proposed algorithm using CNN techniques. the sixth image represents the mean square error(MSE) for the output.

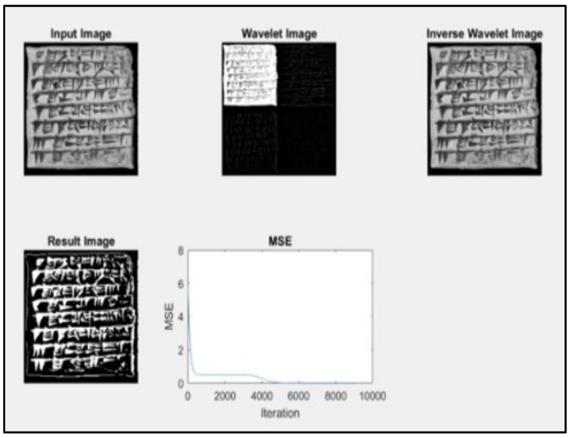

Figure 3: The Input /Output of the proposed algorithm.

In order to evaluate the performance of the classifier, the confusion matrix has been used. The Confusion matrix is a two by two table that contains four outcomes (true positive TP, true negative TN, false-positive FP, and false-negative FN) produced by a classifier (Table 1). These outcomes represent essential performance measures, which are accuracy, specificity, and sensitivity. These performance measures can be derived directly from the confusion matrix. Figure (4) demonstrates a plot of the confusion matrix, which has rows and columns and diagonal cells for a confusion matrix. The rows stand for the predicted class (the Output Class). While the columns stand for the true class (The Target Class).

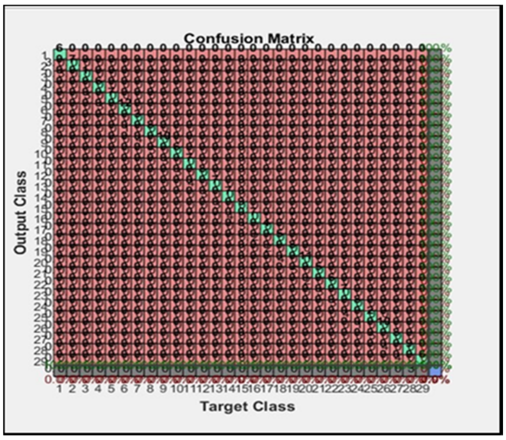

Figure 4: The plot of the confusion matrix.

In order to know how many or what is the percentage that the trained network examples had correctly estimated their classes by observing, the diagonal cells should be observed. In this figure, it shows performance with 100%. Table 1 shows the results of the confusion matrix where:

- TN is the number of correct predictions that an instance is negative,
- FN is the number of incorrect predictions that an instance is positive,
- TP is the number of incorrect of predictions that an instance negative,
- FP is the number of correct predictions that an instance is positive.

Table 1: Confusion Matrix.

Confusion matrix		Predicted	
		Negative	Negative
Actual	Negative	TN=0	FP=0
	Positive	FN=0	TP=100%

Finally, the measurement of each sensitivity (the probability that has a condition of the identification's test that the correct-characters are correct-characters), specificity (the probability that has a condition of the identification's test for the not correct-characters witch are not correct-characters), and accuracy are 100% for each. That means only the correct-characters are classified and resulted from the system, While e, not correct-characters are didn't resulted from the system. These measurements are checked by visually comparing with the correct-charters dataset in Figure 5 and the resulted correct-characters in Figure 6.

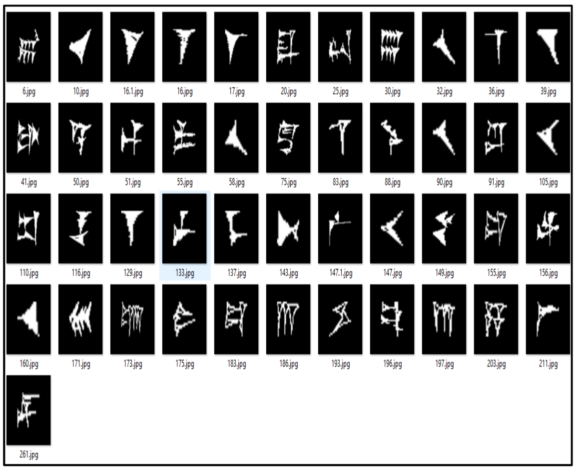

Figure 5: The dataset of the images of the correct-characters.

Figure 6: The correct resulted characters from the tablet image.

5.2 Result and Discussion

In this section, a comparison has been made between the proposed system(CNN based system) and the Ancient Cuneiform Text Extraction based on Automatic Wavelet Selection (**ACTEBWS)** (Majeed et al., 2015). This comparison can be summarized in the following point:

1. The proposed system takes 20 CDLI (cdli, 2017) images as input, while the ACTEBWS only one image as an input has been taken.
2. In both systems, the Wavelet transform has been applied.
3. The proposed system applies Region Splitting to extract characters from each tablet images, while the ACTEBWS extracts the whole text from the tablet without the ability to split each character separately.
4. The proposed system consume less time in handling one image and extracts characters from the tablet image. On the other hand, ACTEBWS consumes more time for the same image without the ability to extracting characters from each image.
5. The proposed system builds its dataset based on the correct Sumerian characters and the incorrect characters at the same time.

6. The usage of deep learning-based CNN in the proposed system gives the high performance of classification. Furthermore, CNN has the ability to extract features directly from the image without the required feature extraction pre-process.
7. The proposed system has achieved a 100% recognition rate when compare with ACTEBWS, which is not achieved this rate.

Figure 7 illustrates the difference between the extraction processes of two systems, where the images in the second column represent the input images while images of the third column represent the output of both systems.

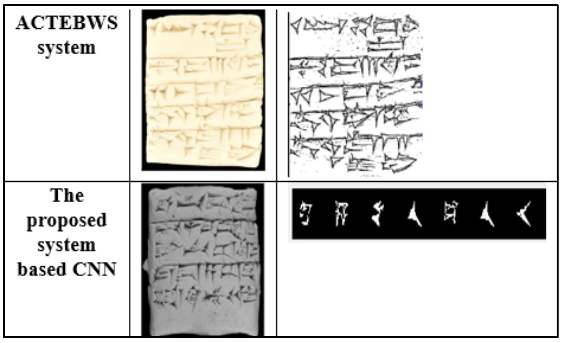

Figure 7: The comparison between ACTEBWS and the proposed system.

6 CONCLUSIONS AND FUTURE WORKS

In this paper, a proposed system including a deep learning-based convolutional neural network (CNN) has been used to extract and classified cuneiform characters from the Sumerian tablets. The proposed system achieves high classification performance with high accuracy recognition for each extracted character, especially when compared with other systems. Furthermore, the system shows high accuracy for the extracted character when it is matching with the reference of cuneiform corrected characters. For future work, we suggest applying the proposed system to other ancient and complex characters just like Assyrian cuneiform or hieroglyphics charters.

REFERENCES

cdli. (2017). Cuneiform Digital Library Initiative.

Edan, N. M. (2013). Cuneiform symbols recognition based on k-means and neural network. *AL-Rafidain Journal of Computer Sciences and Mathematics, 10*(1), 195-202.

Hijazi, S., Kumar, R., & Rowen, C. (2015). Using convolutional neural networks for image recognition. *Cadence Design Systems Inc.: San Jose, CA, USA*, 1-12.

Kassian, A. (2014). Lexical matches between Sumerian and Hurro-Urartian: possible historical scenarios. *Cuneiform Digital Library Journal, 4*, 1-23.

Kramer, S. N. (1963). Cuneiform Studies and the History of Literature: The Sumerian Sacred Marriage Texts. *Proceedings of the American Philosophical Society, 107*(6), 485-527.

Majeed, R., Beiji, B. Z., Hiyam, H., & Jumana, W. (2015). Ancient Cuneiform Text Extraction Based on Automatic Wavelet Selection. *International Journal of Multimedia and Ubiquitous Engineering, 10*(6), 253-264.

O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

Patterson, J., & Gibson, A. (2017). *Deep learning: A practitioner's approach*: " O'Reilly Media, Inc.".

Sasaki, H., Horiuchi, T., & Kato, S. (2015). Japanese Historical Character Recognition using Convolutional Neural Networks. *ICIC express letters. Part B, Applications: an international journal of research and surveys, 6*(12), 3159-3164.

Talib, M., & Harbi, J. (2017). Sumerian Character Extraction by Using Discrete Wavelet Transform and Split Region Methods. *Kurdistan Journal of Applied Research, 2*(3), 62-65.

Tsai, C. (2016). Recognizing handwritten Japanese characters using deep convolutional neural networks. *University of Stanford in Stanford, California*, 405-410.

Yang, W., Jin, L., Xie, Z., & Feng, Z. (2015). *Improved deep convolutional neural network for online handwritten Chinese character recognition using domain-specific knowledge.* Paper presented at the 2015 13th international conference on document analysis and recognition (ICDAR).