Data Mining Ancient Script Image Data Using Convolutional Neural Networks

Shruti Daggumati
University of Nebraska-Lincoln
Lincoln, Nebraska, USA
sdagguma@cse.unl.edu

Peter Z. Revesz
University of Nebraska-Lincoln
Lincoln, Nebraska, USA
revesz@cse.unl.edu

ABSTRACT

The recent surge in ancient scripts has resulted in huge image libraries of ancient texts. Data mining of the collected images enables the study of the evolution of these ancient scripts. In particular, the origin of the Indus Valley script is highly debated. We use convolutional neural networks to test which Phoenician alphabet letters and Brahmi symbols are closest to the Indus Valley script symbols. Surprisingly, our analysis shows that overall the Phoenician alphabet is much closer than the Brahmi script to the Indus Valley script symbols.

CCS CONCEPTS

• **Computing methodologies** → *Machine learning*; *Machine learning approaches*;

KEYWORDS

Indus Valley Script, Brahmi Script, Phoenician Alphabet, Convolutional Neural Networks

ACM Reference format:
Shruti Daggumati and Peter Z. Revesz. 2018. Data Mining Ancient Script Image Data Using Convolutional Neural Networks. In *Proceedings of 22nd International Database Engineering & Applications Symposium, Villa San Giovanni, Italy, June 18–20, 2018 (IDEAS 2018),* 6 pages.
https://doi.org/10.1145/3216122.3216163

1 INTRODUCTION

From 3200 to 1300 BCE the Indus Valley Civilization thrived in northwestern South Asia, including areas of present-day India, Pakistan and Afghanistan [34]. The theories regarding the origin of the Indus Valley Civilization and its script ranges from antecedent indigenous roots to diffusionist explanations [34]. The Indus Valley script is an undeciphered script with over 400 different symbols and thousands of inscriptions, mostly on seals. Unfortunately, the inscriptions are too short for traditional decipherment techniques. Figure 2 shows the most frequent Indus Valley script symbols.

Decipherment efforts are usually aided by bilingual inscriptions, which have not been found yet. Another clue to a script could come from finding a similar but already known script with which the unknown symbols could be matched. That matching can give the unknown script symbols a tentative phonetic value. Usually, the already known scripts occur later in time than the unknown script. In this paper, we use neural networks and two known scripts to find tentative phonetic assignments to the Indus Valley script symbols. These two scripts are the following:

Phoenician alphabet: The first known script that we try is the Phoenician alphabet, which is frequently studied because it spread to a large part of Eurasia. The Phoenician alphabet is an abjad writing system, written from right to left, which consists of 22 letters representing consonants (see Figure 1) [6].

Brahmi syllabary: The second known script is the Brahmi script, which is the second oldest South Asian script. The origin of the Brahmi script is controversial. The Brahmi script is said to stem from the Phoenician alphabet [33]. The lack of intermediate archaeological artifacts between the end of the Indus Valley Civilization in 1300 BCE and the earliest Brahmi script in the late 4th to mid 3rd centuries BCE [32] makes the latter unlikely to be a descendant of the former. Unless all writing was done on perishable materials the intervening period.

The Brahmi script is a syllabary but for the same consonant C the Ca, Ce, Ci, Co, and Cu. forms are only minor variations of each other. In this paper, we focus on the syllabic symbols of Ca series, which are shown in Figure 3.

In this paper, we answer the following data mining questions:

(1) Which script, Phoenician or Brahmi, is more likely to be a descendant of the Indus Valley script?
(2) What tentative phonetic assignments can be given to the Indus Valley script symbols?

Section 2 describes the dataset of the ancient scripts and texts which we used as a data source in our research. Section 3 describes the neural networks that we used for the computerized comparison of the visual characteristics of pairs of symbols from two different scripts. Section 4 presents the experimental results. Section 5 discusses related work. Finally, Section 6 gives some conclusions and directions for further research.

2 THE DATASET USED

Image databases (see Chapter 8 in [23]) are ubiquitously used in numerous applications from recognizing text in images [10–12], affine invariant image retrieval [9], facial recognition systems and traffic monitoring [39]. Images in image databases can be represented by pixels, vectors, or constraints [14]. A recent image database called the Archaeoastronomy in Space and Time Database contains not only ancient Indus Valley script texts, but also Maya, Goldhüte, Proto-Byblos script, and the Minoan scripts [7]. For this paper, we use three different ancient scripts: Phoenician alphabet, Indus Valley script, and the Brahmi script. For the Phoenician alphabet, we

IDEAS 2018, June 18–20, 2018, Villa San Giovanni, Italy

ACM ISBN 978-1-4503-6527-7/18/06.
https://doi.org/10.1145/3216122.3216163

Figure 1: Phoenician alphabet letters

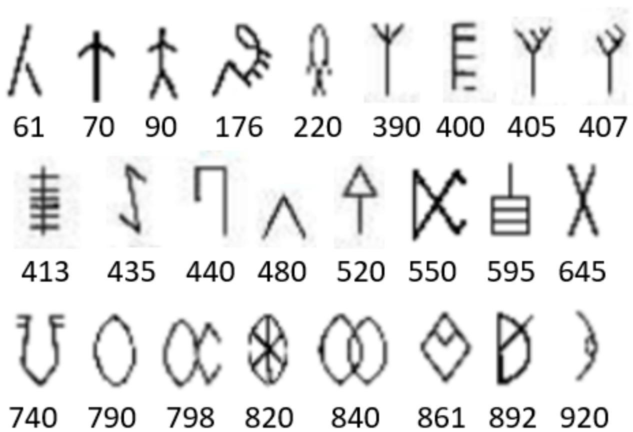


Figure 2: Twenty-Five of most frequent Indus Valley script symbols.

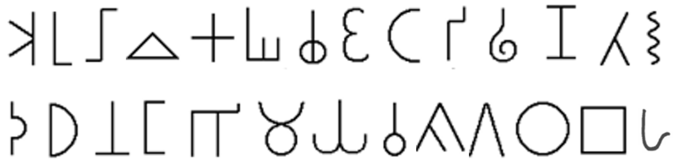


Figure 3: Brahmi script symbols standing for vowel symbols of the form CA, where C is a consonant and A is the vowel /a/.

use all 22 symbols (Figure 1), for the Indus Valley script we used 25 of the symbols (Figure 2), and for the Brahmi script we use 27 of the symbols (Figure 3). We use only the symbols with the highest frequencies because the Indus Valley script has over 400 symbols and symbols that occur only once or twice are likely to be insignificant [36].

The MNIST image database [16] contains 60,000 training images of handwritten black and white digits, where they use 10,000 images for validation images. In comparison to the MNIST database, our training and validation datasets contain far fewer images. We use transformations and distortions on the images to create a bigger dataset.

Each symbol has forty training images and eight validation images, that is a total of 48 images associated with it. Each image is 25x25 pixels, and it is black and white like the MNIST dataset [16]. In total, we have 3,552 images.

3 THE DESIGN OF THE SCRIPT IMAGE RECOGNITION NEURAL NETWORK

In deep learning, it is necessary to learn from features available in the data. In most neural networks learning is possible only if a large amount of data are available. However, convolutional neural networks (CNNs) also allow learning from small datasets. CNNs are multi-layer neural networks which assume that the data is of the form of an image, which allows the encoding of certain properties into the architecture, entailing a reduction in the number of parameters in the network [2].

Using ideas similar to Chollet [2], in order to use a smaller data set, we apply random transformations and normalization operations to the training image data set. The image is transformed in the following ways: rotation, translation, scaling, zooming, and flipping.

The created neural network uses Python and Tensorflow, and we also use the Keras wrapper. Tensorflow is an open source library for numerical computation, which uses data flow graphs. Keras is a wrapper that allows the use of a TensorFlow backend, providing modularity and Python-nativeness, allowing for out of the box implementations of common network structures.

The constructed neural networks have various levels of accuracy dependent on the script. The neural networks use three main layers: a two-dimensional convolutional layer, a two-dimensional pooling layer, and a dense layer.

The first convolutional layer applies 32, three by three filters, where it extracts three by three pixel subregions using the rectified linear unit (ReLU) activation function. The first pooling layer performs maximum pooling with a two by two filter and stride of 2 (which specifies that the pooled regions do not overlap). The second convolutional layer applies 64 of the three by three filters, with the ReLU activation function. The second pooling layer repeats the process of the first pooling layer. Lastly, the first dense layer has 1,024 neurons, with a dropout regularization rate of 0.4, which is the probability of any given element being dropped during training. The second dense layer has 22 (Phoenician), 25 (Indus), or 27 (Brahmi) neurons, one per target class.

3.1 Script Recognition

All twenty-two of the Phoenician symbols were used for the neural network. For the Brahmi script, we used only the more prominent vowels and the root consonants. For the Indus Valley Script, we used the most frequent symbols. According to Wells [35] each of the symbols has a root symbol and additional items are added.

Tables 1 and 2 show the different accuracy measures we have for the three datasets. Table 1 shows the accuracy of each neural network predicting validation data, which was not used in training. Table 2 shows the accuracy of the neural network is with a combination of training and validation data.

We see from Table 1 that as the number of epochs increased, the validation accuracy of the neural network recognizing the correct symbols also increased for all three scripts. When using training

Table 1: Validation Accuracy

	Number of Epochs			
	25	50	75	100
Phoenician	93.18	94.77	94.77	94.77
Brahmi	95.09	98.15	98.24	99.35
Indus	93.50	95.50	96.80	98.00

Table 2: Training and Validation Accuracy

	Number of Epochs			
	25	50	75	100
Phoenician	87.00	93.55	96.41	96.66
Brahmi	81.68	89.95	92.74	94.61
Indus	83.82	91.98	94.02	96.22

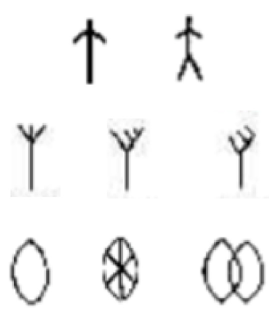


Figure 4: Indus Valley script symbols - each row contains symbols with similar shape/structure

and validation data as the epochs increased the accuracy also increased significantly from 87 percent to more than 96 percent for the Phoenician alphabet, 81 percent to almost 95 percent for the Brahmi script, and for the Indus Valley script it jumped from almost 83 percent to more than 96 percent. As the epochs increased, the validation accuracy started with a fairly high percent for all three datasets and increased marginally for the Phoenician alphabet dataset, increased slightly by four percent for the Brahmi script dataset and the Indus Valley Script dataset.

Overall, Table 2 has a lower accuracy for all cells in comparison to Table 1. If the validation accuracy is lower than the combined training and validation accuracy, then overfitting occurs. Overfitting is when the model does not generalize well and has become too accustomed to the training data, and when new data is presented it can react haphazardly.

3.2 Data Analysis

From Table 2, we see that the Phoenician neural network has the highest accuracy overall even at the lowest epoch. Referring back to the symbols in the language, we see that they are fairly distinct. This leads us to believe that the uniqueness of the Phoenician letters plays a key factor in the accuracy of our neural network.

Figure 5: Brahmi script symbols - each row contains symbols which have similar shape/structure

However, similar statements cannot be made about the Brahmi script. The accuracy of the Brahmi script was lower for the training and validation accuracy and the higher for the validation accuracy. The Brahmi script has numerous symbols which have similar shapes. Figure 5 shows that there is a similarity between either two or three symbols for the majority, and a similarity between five symbols for a single case (middle row of Figure 5).

For the Indus Valley script, there are more pairs or triples of symbols which are slightly more similar (see Figure 4). There are fewer symbols which are similar when compared to the Brahmi script, which explains why the Indus Valley script has a slightly higher accuracy than the Brahmi script has.

4 EXPERIMENT

Given that our neural networks perform well at recognizing and classifying each symbol set to its respective value, we sought to look at how the Phoenician and Brahmi datasets would be classified as input to the Indus Valley script neural network. These predictive measures were done using forty-eight sample images for each of the twenty-two Phoenician alphabet symbols and the twenty-seven Brahmi script symbols.

The Indus Valley script neural network was compiled a total of five times, due to neural networks running differently on the runs. From each of the compilations, we stored the weights according to the classification to determine how the Brahmi symbols and Phoenician alphabet would be classified according to the Indus Valley script. Among the five compilations, there are times where the symbol classified has an overall consensus regarding the mapped symbol. However, there are other times where there is an equal mapping to two symbols. In that case, we decide to use the symbol with a high strength match, where strength is the summed percentage of how similar the pair of symbols is. In the case of a few

Table 3: Average strength of Phoenecian alphabet and Brahmi script mappings to the Indus Valley script symbols, with duplicate mappings, and duplicate mappings removed.

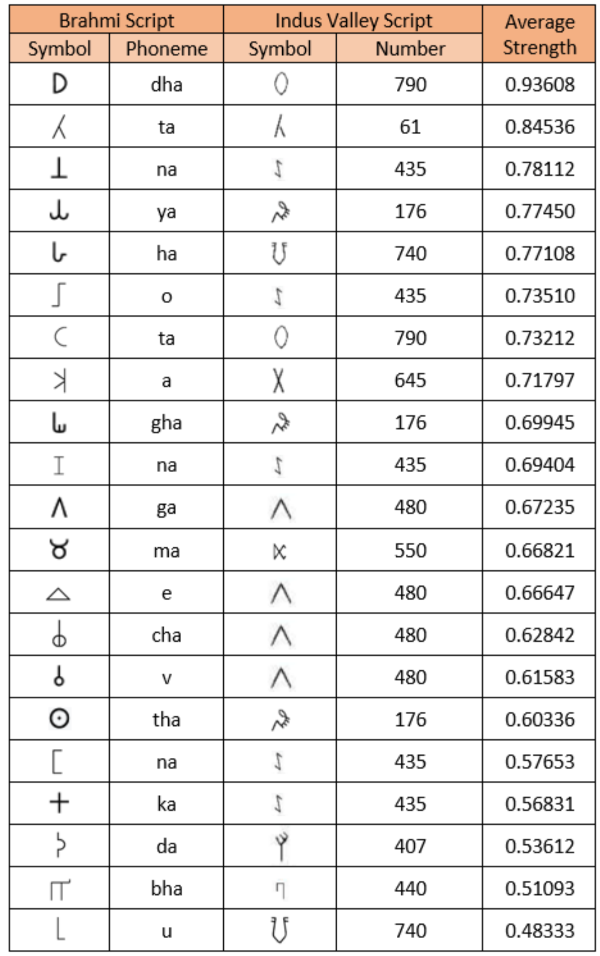
	Avg. Strength with Duplicates	Avg. Strength without Duplicates
Phoenician	0.6047	0.6546
Brahmi	0.6071	0.6490

symbols that have a random mapping for all five runs, we choose the mapping with the highest strength.

Figures 6 and 7 show the classification of each Brahmi and Phoenician symbol passed to the Indus Valley neural network. The tables are ordered from the strongest to the weakest match among the symbols. Regarding whether there might have been an interference among the symbols, as in the case of a double mapping from one Brahmi script symbol/Phoenician alphabet symbol to the Indus Valley script symbols, we used the following algorithm: based on the strength ordering if there is an Indus Valley script symbol which has already been mapped, then we choose to pick the next strongest symbol in relation to the Brahmi/Phoenician symbols. If this second symbol already exists among the chosen ones, then we decide to stick with the original strongest symbol. Regarding duplicates, we believe adding more Indus Valley script symbols could possibly reduce the number of multiple mappings and may give more insight regarding more similarities among the scripts.

Among the symbols chosen that match Brahmi, out of the twenty-seven Brahmi symbols we have numerous duplicates or even triplicates. For the Phoenician symbols, we have fourteen unique mappings and only eight duplicates. Moreover, the neural network choices usually yield visually similar pairs. Hence it is understandable why the neural network would classify a given Phoenician alphabet letter to the selected Indus Valley script symbol.

The average strength of Phoenician and Brahmi symbols passed to the Indus Valley script neural network (see Table 3) indicates that Phoenician alphabet has a stronger connection to the Indus Valley script than to the Brahmi script. Visually, we can see that the symbols which the neural network picked are more reasonable for the Phoenician versus the Indus Valley script symbol sets. The shape structure as to why a neural network may pick a selected symbol is quite clear for most symbols. Figure 8 shows how the Phoenician alphabet and Brahmi script symbols map to the Indus Valley script symbols. We see that there is more of a unique mapping for Phoenician than Brahmi. Table 3 shows the overall average strength for all the symbol mappings regardless of duplicates and when the duplicates are removed. We see that in this case Phoenician still has the best mapping. With duplicate mappings, we see that Brahmi has a slightly higher average strength. However, there are numerous mappings to symbol "480" in the Indus Valley script. Removing these and other duplicates shows that Phoenician has more unique mappings where 14 out of 22 symbols are uniquely mapped, whereas for Brahmi only 13 out of the 27 symbols were uniquely mapped.



Brahmi Script		Indus Valley Script		Average Strength
Symbol	Phoneme	Symbol	Number	
	dha		790	0.93608
	ta		61	0.84536
	na		435	0.78112
	ya		176	0.77450
	ha		740	0.77108
	o		435	0.73510
	ta		790	0.73212
	a		645	0.71797
	gha		176	0.69945
	na		435	0.69404
	ga		480	0.67235
	ma		550	0.66821
	e		480	0.66647
	cha		480	0.62842
	v		480	0.61583
	tha		176	0.60336
	na		435	0.57653
	ka		435	0.56831
	da		407	0.53612
	bha		440	0.51093
	u		740	0.48333

Figure 6: Brahmi script symbols passed into Indus Valley neural network

5 RELATED WORK

5.1 Decipherment Attempts

Various scholars have tried to decipher the Indus Valley script. However, it is difficult to decipher because the Indus Valley script is found only in short length inscriptions and has no bilingual text.

Sir Alexander Cunningham, one of the first people to encounter the Indus Valley script, assumed considered it to be an ancestor of Brahmi. That view is supported by many other scholars [20–22]. Scholars also generally suppose that the Indus Valley script expresses some Dravidian language close to Tamil [17–19, 35, 37, 38, 40]. However, some researchers state that the Indus Valley script should not be considered a language because it seems more similar to nonlinguistic signs such as those that symbolize family or clan names/symbols and religious figures/ concepts [5].

5.2 Computer-Aided Techniques

Scholars have used machine learning techniques to analyze images and read text found in them [10–12]. However, the use of neural

Phoenician Alphabet		Indus Valley Script		Average Strength
Symbol	Phoneme	Symbol	Number	
	w [w]		390	0.99102
	g [g]		70	0.90504
	ṭ [tˤ]		820	0.88454
	ʿ [ʕ]		790	0.80389
	l [l]		61	0.75773
	š [ʃ]		550	0.69893
	s [s]		413	0.67104
	r [r]		70	0.66967
	t [t]		645	0.66061
	q [q]		407	0.59651
	m [m]		550	0.57895
	ḥ [ħ]		595	0.55183
	k [k]		645	0.54641
	h [h]		90	0.48331
	p [p]		861	0.48017
	n [n]		550	0.58057
	d [d]		861	0.47028
	z [z]		440	0.45986
	ʾ [ʔ]		90	0.44916
	ṣ [sˤ]		440	0.39817
	y [j]		435	0.33351
	b [b]		820	0.33123

Figure 7: Phoenician alphabet symbols passed into Indus Valley neural network

networks to compare language families has yet to be explored. Entropic analysis on the Indus Valley script according to [20] states that the entropy of linguistic scripts closely matches existing linguistic systems. Markov chain models have been used to determine the chain of symbols which are normally found dependent among symbols [21]. Zide and Zvelebil [40] was one of the first attempts to decipher the Indus Valley script using a computer.

Revesz used different computer-aided techniques to analyze the evolution of the Cretan Script Family [26]. The evolution of the Cretan Script Family was analyzed using a feature-based similarity measure of script symbols and phylogenetic algorithms [24, 30] to derive a hypothetical script evolutionary tree, which showed a strong similarity between Cretan hieroglyph scripts and the Old Hungarian alphabet, which is called *Rovásírás* in the native language. The matching of Cretan hieroglyphs script symbols and the old Hungarian alphabet letters allowed to give new phonetic values for the former. Further, the new phonetic values aided the decipherment of Cretan hieroglyphs [25], including the Phaistos disk [27] and the Arkalochori Axe [29]. In addition, the Minoan Linear A Script was also matched with the Old Hungarian alphabet leading to a decipherment of Linear A too [28]. According to these decipherments, the Minoan language was closely related to the Hattic language in Anatolia (present-day Turkey) and to Hungarian. In fact, Revesz [28] proposed to group these languages together into the *West-Ugric* language family.

5.3 Database

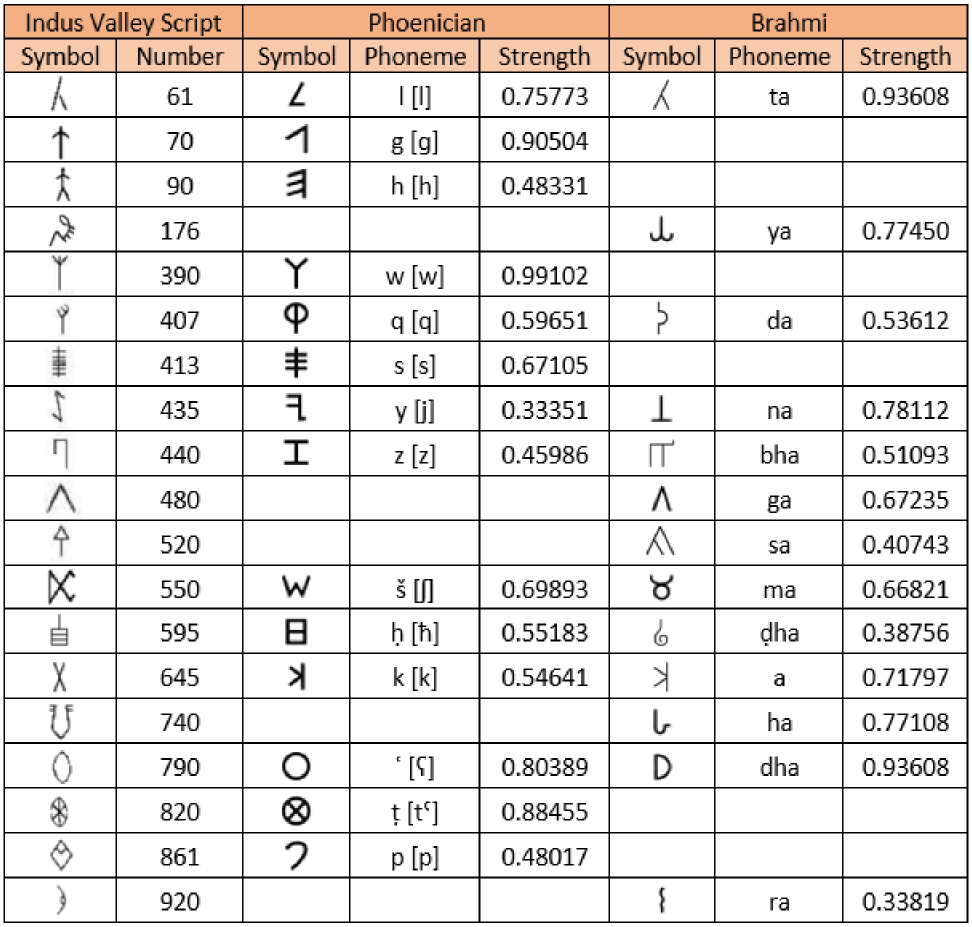
Bryan Wells and Adreas Fuls have created a database [36] with the signs identified by Bryan Wells in his Ph.D. dissertation [37]. The sign list according to Wells contains 695 distinct signs. The database shows statistics regarding each symbol's frequency and location. It also shows analysis on the distribution of the symbol with respect to others, i.e. precedes, follows, etc.

6 CONCLUSIONS AND FUTURE WORK

Our results show that the Indus Valley script is surprisingly closer to the Phoenician alphabet than to the Brahmi script. The computer science methods in this paper identified a strong connection between the Phoenician alphabet and the Indus Valley Script symbols and that can lead to some interesting scientific deductions. In particular, the neural networks-based matchings between the Phoenician alphabet letters and the Indus Valley symbols suggest some phonetic values for the Indus Valley symbols. In the future, such computer-generated phonetic suggestions may help linguists to decipher the Indus Valley script.

REFERENCES

[1] Ujjwal Bhattacharya and BB Chaudhuri. 2005. Databases for research on recognition of handwritten characters of Indian scripts. In *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*. IEEE, 789–793.
[2] Francois Chollet. 2016. Building powerful image classification models using very little data. https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html.
[3] H Devi. 2006. Thresholding: A Pixel-Level image processing methodology preprocessing technique for an OCR system for the Brahmi script. *Ancient Asia* 1 (2006).
[4] Walter Ashlin Fairservis. 1992. *The Harappan civilization and its writing: a model for the decipherment of the Indus script*. Brill.
[5] Steve Farmer, Richard Sproat, and Michael Witzel. 2016. The collapse of the Indus-script thesis: The myth of a literate Harappan civilization. *Electronic Journal of Vedic Studies* 11, 2 (2016), 19–57.
[6] Steven R Fischer. 2004. *History of writing*. Reaktion books.
[7] Andreas Fuls. 2018. Online Archaeoastronomy in Space and Time. http://www.user.tu-berlin.de/fuls/Homepage/.
[8] Swarajya Prakash Gupta and Kattalai Subrahmanyayyar Ramachandran. 1979. *The origin of Brahmi script*. Vol. 2. DK Publications.
[9] Sofie Haesevoets, Bart Kuijpers, and Peter Z Revesz. 2017. Affine-Invariant Triangulation of Spatio-Temporal Data with an Application to Image Retrieval. *ISPRS International Journal of Geo-Information* 6, 4 (2017), 100.
[10] Max Jaderberg, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. 2014. Synthetic data and artificial neural networks for natural scene text recognition. *arXiv preprint arXiv:1406.2227* (2014).
[11] Max Jaderberg, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. 2016. Reading text in the wild with convolutional neural networks. *International Journal of Computer Vision* 116, 1 (2016), 1–20.
[12] Keechul Jung, Kwang In Kim, and Anil K Jain. 2004. Text information extraction in images and video: a survey. *Pattern recognition* 37, 5 (2004), 977–997.
[13] Subhash C Kak. 1988. A frequency analysis of the Indus script. *Cryptologia* 12, 3 (1988), 129–143.
[14] Paris C Kanellakis, Gabriel M Kuper, and Peter Z Revesz. 1995. Constraint query languages. *J. Comput. System Sci.* 51, 1 (1995), 26–52.



Indus Valley Script		Phoenician			Brahmi		
Symbol	Number	Symbol	Phoneme	Strength	Symbol	Phoneme	Strength
	61		l [l]	0.75773		ta	0.93608
	70		g [g]	0.90504			
	90		h [h]	0.48331			
	176					ya	0.77450
	390		w [w]	0.99102			
	407		q [q]	0.59651		da	0.53612
	413		s [s]	0.67105			
	435		y [j]	0.33351		na	0.78112
	440		z [z]	0.45986		bha	0.51093
	480					ga	0.67235
	520					sa	0.40743
	550		š [ʃ]	0.69893		ma	0.66821
	595		ḥ [ħ]	0.55183		ḍha	0.38756
	645		k [k]	0.54641		a	0.71797
	740					ha	0.77108
	790		ʿ [ʕ]	0.80389		dha	0.93608
	820		ṭ [tˤ]	0.88455			
	861		p [p]	0.48017			
	920					ra	0.33819

Figure 8: Phoenician alphabet symbols and Brahmi script symbols passed into Indus Valley neural network

[15] David H Kelley and Bryan Wells. 1995. Recent Developments In Understanding The Indus Script. *The Quarterly Review of Archaeology* 16, 1 (1995), 15–23.

[16] Yann LeCun, Corinna Cortes, and Chris Burges. 1998. MNIST handwritten digit database. http://yann.lecun.com/exdb/mnist/.

[17] Asko Parpola. 1986. The Indus script: A challenging puzzle. *World Archaeology* 17, 3 (1986), 399–419.

[18] Asko Parpola. 2005. Study of the Indus script. In *Proceedings of the International Conference of Eastern Studies*, Vol. 50. 28–66.

[19] Asko Parpola. 2009. *Deciphering the Indus script.* Cambridge University Press.

[20] Rajesh PN Rao, Nisha Yadav, Mayank N Vahia, Hrishikesh Joglekar, R Adhikari, and Iravatham Mahadevan. 2009. Entropic evidence for linguistic structure in the Indus script. *Science* 324, 5931 (2009), 1165–1165.

[21] Rajesh PN Rao, Nisha Yadav, Mayank N Vahia, Hrishikesh Joglekar, R Adhikari, and Iravatham Mahadevan. 2009. A Markov model of the Indus script. *Proceedings of the National Academy of Sciences* 106, 33 (2009), 13685–13690.

[22] Shikaripur Ranganatha Rao. 1982. *The decipherment of the Indus script.* Asia Publishing House.

[23] Peter Z. Revesz. 2010. *Introduction to Databases: From Biological to Spatio-Temporal.* Springer, New York.

[24] Peter Z. Revesz. 2013. An algorithm for constructing hypothetical evolutionary trees using common mutations similarity matrices.

[25] Peter Z. Revesz. 2016. Bioinformatics evolutionary tree algorithms reveal the history of the Cretan Script Family. *Internation Journal of Applied Mathematics and Informatics* 10 (2016), 67–76.

[26] Peter Z. Revesz. 2016. A computer-aided translation of the Cretan Hieroglyph script. *International Journal of Signal Processing* 1 (2016), 127–133.

[27] Peter Z. Revesz. 2016. A computer-aided translation of the Phaistos Disk. *International Journal of Computers* 10 (2016), 94–100.

[28] Peter Z. Revesz. 2017. Establishing the West-Ugric language family with Minoan, Hattic and Hungarian by a decipherment of Linear A. *WSEAS Transactions on Information Science and Applications* 14 (2017), 306–335.

[29] Peter Z. Revesz. 2017. A Translation of the Arkalochori Axe and the Malia Altar Stone. *WSEAS Transactions on Information Science and Applications* 14, 1 (2017), 124–133.

[30] Peter Z. Revesz and Zhiqiang Li. 2016. Incremental phylogenetics by repeated insertions: An evolutionary tree algorithm. *International Journal of Biology and Biomedical Engineering* 10, 1 (2016), 148–158.

[31] Richard Salomon. 1996. *Deciphering the Indus Script.* JSTOR.

[32] Richard Salomon. 1998. *Indian Epigraphy: A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and the Other Indo-Aryan Languages.* Oxford University Press.

[33] Richard G Salomon. 1996. Brahmi and Kharoshthi. *The worldâĂŹs writing systems* (1996), 373–383.

[34] Upinder Singh. 2008. *A history of ancient and early medieval India: from the Stone Age to the 12th century.* Pearson Education India.

[35] Bryan Wells. 1998. *An introduction to Indus writing.* University of Calgary.

[36] Bryan Wells and Andreas Fuls. 2017. Online Indus Writing Database. http://caddy.igg.tu-berlin.de/indus/welcome.htm.

[37] Bryan K Wells. 2011. *Epigraphic approaches to Indus writing.* Oxbow Books.

[38] Bryan K Wells and Andreas Fuls. 2015. *The archaeology and epigraphy of Indus writing.* Archaeopress.

[39] Hang Yue, Laurence R Rilett, and Peter Z Revesz. 2016. Spatio-temporal traffic video data archiving and retrieval system. *GeoInformatica* 20, 1 (2016), 59–94.

[40] Arlene R Zide and Kamil V Zvelebil. 1976. *The Soviet decipherment of the Indus Valley script: Translation and critique.* Vol. 156. Walter de Gruyter.