SVM CLASSIFIER FOR THE PREDICTION OF ERA OF AN EPIGRAPHICAL SCRIPT

Soumya A and G Hemantha Kumar

Department of Computer Science & Engineering, R V College of Engineering, Karnataka, India
soumyaa@rvce.edu.in
Department of Studies in Computer Science, University of Mysore, India
ghk2007@yahoo.com

ABSTRACT

Inscriptions are the main source for reconstructing the history and culture of ancient civilizations. The scripts of modern Indian languages have evolved over centuries and finally transformed to the present form. Modern readers find difficulty in interpreting a script of olden days. The characters have changed over time. Hence for reading ancient scripts the period has to be determined, so as to have knowledge of which character set of ancient days is to be employed for automatic reading. Prediction of the era of a given ancient script is a follow-on member of an ancient script recognition system and can be used as a component of the OCR system for ancient scripts. This knowledge can be used by archaeologists and historians for further explorations. In this paper we demonstrate period identification of various ancient Kannada scripts using SVM classifier. A system is proposed for prediction of the era and it is being done by examining a few characters in Kannada script of various periods referred to as test characters. These test characters are sampled from the script automatically and matched with the characters available for different periods using machine intelligence.. This classifier is tested on quite number samples of Kannada epigraphical document images belonging to nine different periods.

KEYWORDS

Inscription, Epigraphy, Paleography, Support Vector Machine (SVM), Optical Character Recognition (OCR)

1. INTRODUCTION

Inscriptions have remarkable importance to mankind. The information culled out from the inscriptions provides us the knowledge of history, culture, astronomy, medicine, management, political, religious, social, economic, tax, administrative, educational conditions and also languages and scripts that prevailed during yester years[1][2][3][4] . **Epigraphy** is the science of identifying the inscriptions on rocks, pillars, temple walls, copper plates and other writing materials. It is a primary tool of archaeology when dealing with literate cultures. Many inscriptions do not contain enough historical details to fix their authorship conclusively. For example, from the inscriptions with the name Rajaraja, it is not always clear whether Rajaraja the First or the Second or the Third is intended. To assign dates to such inscriptions and to identify the kings, palaeography is the main tool. **Paleography** is the study of ancient handwriting and the practice of deciphering and reading historical manuscripts. The paleographer must have the knowledge of: first, the language of the text and second, the historical usages of various styles of handwriting, common writing customs, and scribal/notarial abbreviations.

DOI : 10.5121/ijp2p.2011.2202

Scripts denote the writing systems employed by the languages. Any language can be written in any script. Having or not having 'own script' is neither a status nor any impediment for language. Three important varieties of scripts that were prevalent in ancient India were: Indus valley script, Brahmi Script and Kharosti script. The scripts of modern Indian languages have evolved from one of these scripts over the centuries. The evolution of the script is dependent on many factors: the writing material, (Stone, Copper, Palm leaf, Paper etc), writing tools, modes of writing and the background of the scribes. In India, prior to invention of writing / printing papers, Palmyra leaves and birch leaves were used for writing purposes. As they could not be long lasting, engraving on rocks, pillars and plates made of copper/ gold / silver came into practice. In India currently there are 13 scripts and 23 official languages for communicating at state level. Apart from these, there are many languages & dialects used by number of people.

The Kannada script has been used to write in Kannada language. Kannada which belongs to highly acclaimed Dravidian language family is the official language of the State of Karnataka and is one of the most enriched languages in India with is long historical heritage. Its earliest written records dates back to about the third century B.C. and the modern script has evolved gradually from the ancient script known as Brahmi. Kannada script has been evolved in its present form, undergoing several twins and turns. In fact, the Indian linguists have demarcated the whole of this evolutionary process in to four broad phases.

- **Poorvada Halegannada or Pre-ancient Kannada -** The first written record in the Kannada language is found in Emperor Ashoka`s Brahmagiri decree, which dated back to as early as 230 BC.
- **Halegannada or Ancient Kannada-** marks the second phase, covering a time period of the 9th to 14th centuries CE.
- **Nadugannada or Middle Kannada-** third phase started from the 14th century and continued till 18th century CE.
- **Hosagannada or Modern Kannada –** the Kannada works that were produced at the end of the 19th century and also much later are categorized as Hosagannada or Modern Kannada.

Kannada script has traveled a long way from the earlier Brahmi model as indicated in Figure 1. The fonts have evolved over the centuries and it has undergone many changes till now. It has undergone a number of changes during the regimes of Shatavahana, Kadamba, Ganga, Rashtrakuta, Chalukya, Hoysala,Vijayanagara and Wodeyar dynasties, and transformed finally into a shape which we are using today [2]. This poses a challenge in understanding these ancient scripts. The stages of evolution, though generally understood by epigraphists, are not known to the educated layman of that province. There was also a certain amount of regional variation. Thus even for the experts, it is difficult to assign dates to many inscriptions, whether complete or fragmentary. The expert epigraphists, who decipher these scripts and translate them into the regional languages, are expected to become extinct in near future and also the importance of inscriptions to mankind is remarkable. Hence there is a dire need for automation of classification and recognition of ancient epigraphical scripts. This work helps us in integrating the new technological concepts with underlying cultural features of the scripts, thereby minimizing the divide between “Man and the History”. The age of the document reflects to which era the script belonged to and the particular dynasty in which it existed.

In automation of deciphering epigraphical scripts, the characters present in the script are to be identified and recognized. The characters have evolved over years and transformed to the shape we are using today. Every era has its own character set, which is to be determined before recognition. Hence there is a need for dating of inscriptions to a particular period, which is the initial step towards decipherment.

The ancient scripts are engraved on rocks, copper, silver or gold plates. The images of these inscriptions are captured and later subjected to digitization process. The script of document images that has been digitized is to identified and later recognized. Hence the present problem addresses off line character recognition. Handwriting identification and recognition are of great practical interest in the extraction of discriminating and invariant information from a handwritten specimen. One of the major difficulties in off-line word recognition originates from the great variations observed in different samples of scripts from the same writer over time or from different scriptors. There is no perfect mathematical model that can describe such extreme variations and hence it is very difficult to find the characteristic features that are invariant with different writing styles.

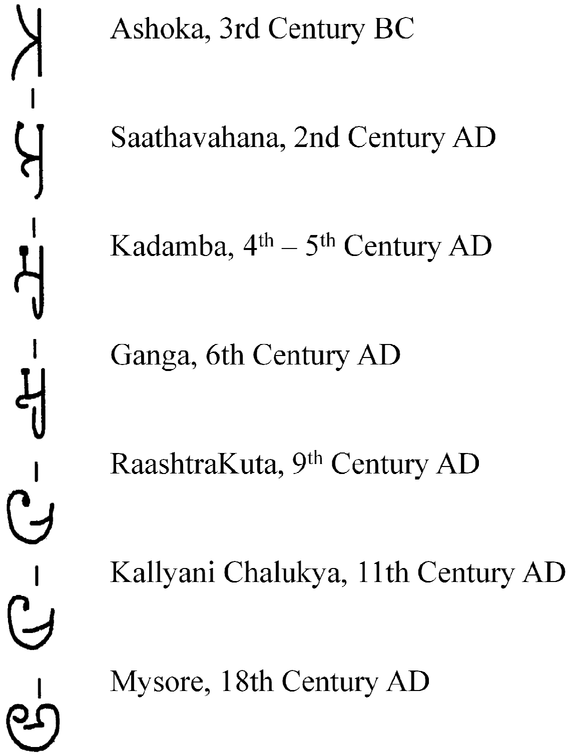

Ashoka, 3rd Century BC
Saathavahana, 2nd Century AD
Kadamba, 4th – 5th Century AD
Ganga, 6th Century AD
RaashtraKuta, 9th Century AD
Kallyani Chalukya, 11th Century AD
Mysore, 18th Century AD

Figure 1: Changes in character shape since 3rd century B.C to 18th century A.D

2. LITERATURE REVIEW

High accuracy OCR systems are reported for English with excellent performance in presence of printing variations and document degradation. Recognizing English characters is much simpler as there are only 26 letters and each letter is quite distinct from others compared to recognition of Indian language characters. For Indian and many other oriental languages- OCR systems are not yet able to successfully recognize printed / handwritten document images of varying scripts, quality, size, style and font. Compared to European languages, Indian languages pose many additional challenges.Many researchers have been working on script recognition for more than three decades but there are very few tools to identify these scripts. Indian languages are characterized with the properties:

(i) Large number of vowels, consonants, and conjuncts.

(ii) Have a base character along with vowels attached, forming single character called the compound character.

(iii) Most scripts spread over several zones.

(iv) Lack of standard test databases of the Indian languages.

There are many conventional OCR systems available for present day Kannada Script, but very few work on ancient Kannada epigraphical scripts are reported. Ancient Script recognition poses additional challenges.Some of researchers are making special efforts to make this knowledge more easily available.Nevertheless, it remains to be one of the most challenging problems in pattern recognition.

Anasuyadevi [2000] has worked on recognition of ancient Indian scripts [5]. She has proposed [2003] a fuzzy neural network approach for the recognition of Brahmi characters.

A hybrid neural network architecture for age identification of ancient Kannada Scripts was proposed by K Harish Kashyap, Bansilal, P Arun Koushik [2003]. After pre-processing the characters, the work is implemented in two phases. The first phase which identifies the base character incorporates an Artificial Neural Network (ANN). ANN is trained by Back propagation algorithm to identify the present day base character corresponding to input character. In the second phase, for identification of age pertaining to the base character, a Probabilistic Neural Network (PNN) - a Bayesian classifier is used taking the advantage that no training is involved prior to classification [6].

The research work by Srikanta Murthy K [2005] for transforming epigraphical objects into machine recognizable form, involves the preprocessing techniques - for removal of noises, segmentation of lines and characters, thinning and classification of the epigraphical document belonging to different period [16]. Two preprocessing techniques for removal of noise have been proposed – the first algorithm is based on a rectangle fitting wherein the height of the character to be retained is assumed to be greater than the noisy pixel. The second algorithm employs a template to obtain the minimum majority of white pixels. Segmentation of lines and characters are carried out using- a Partial Eight Direction Based Line Segmentation (PEBLS) algorithm wherein horizontal Projection profile is applied to identify the base and supplementary reference lines. The second approach is based on Nearest Neighbor Clustering (NNC), could be used even when the document is skewed. Three thinning algorithms - two-step algorithm, fully parallel thinning algorithm and rotation invariant 4- step algorithm have been designed. Classification of the epigraphical document belonging to different period is carried out using - a method based on texture features, a method based on invariant moments which are invariant to rotation, translation and scaling, and for accurate estimation of the period, a neural network based approach is adopted [17][18][19].

3. PROPOSED METHODOLOGY

Age identification of ancient scripts is important, which enables to know which character set is to be employed for automatic reading of age old scripts. It is a problem characterized under the genre of pattern recognition and image analysis. Figure 2 illustrates the steps towards classification of epigraphical document images to their respective era.

The Age prediction classifier involves the following steps:

I. ***Image Acquisition***: The present work proposes to take the objects of the inscriptions as the inputs. These objects may be camera grabbed or scanned if the inscriptions are engraved on a plate. The photographs of inscriptions on rock or pillars may not result in good objects. Hence estempages are produced [Dani A.H, 1986; Sircar D.C, 1965] and later scanned to get the object image of the inscription.

II. The input image of the inscription may be degraded due to the presence of the broken characters, erased characters, touching characters, distortion due to fossils settled, irrelevant symbols engraved by the scribes and so on. The non uniform spacing between the lines and characters of epigraphical document and the skew could complicate the process of deciphering the script. Hence ***the input document image is preprocessed for removal of noise and skew correction, followed by segmentation of characters.***

III. Appropriate ***Feature extraction method is devised*** for measuring the relevant shape information contained in a pattern so that the task of classifying the pattern is made easy by a formal procedure.
The features extracted have the following properties:
• Easy to extract, which reduces the complexity of the program.
• Distinct, which eases the classification process.
• Independent of font type and size, which is a big advantage.

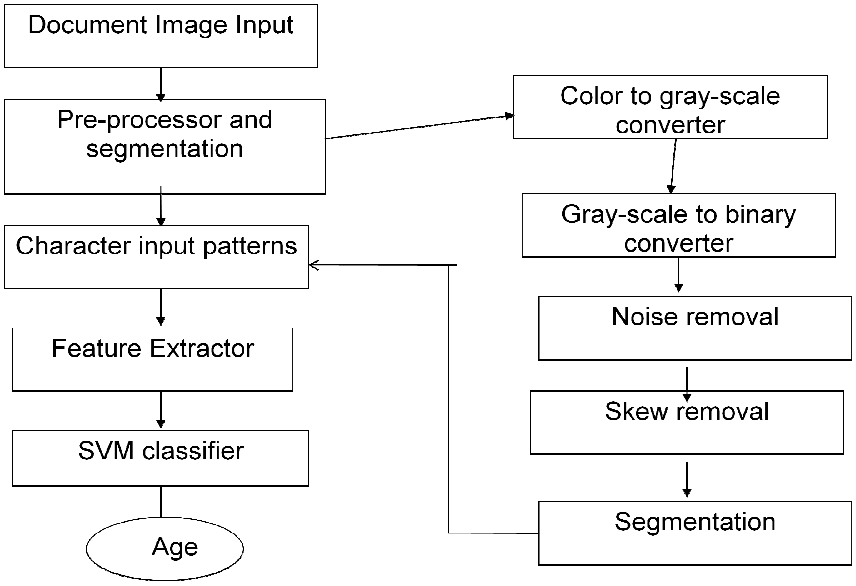


Figure 2. Age Identification System

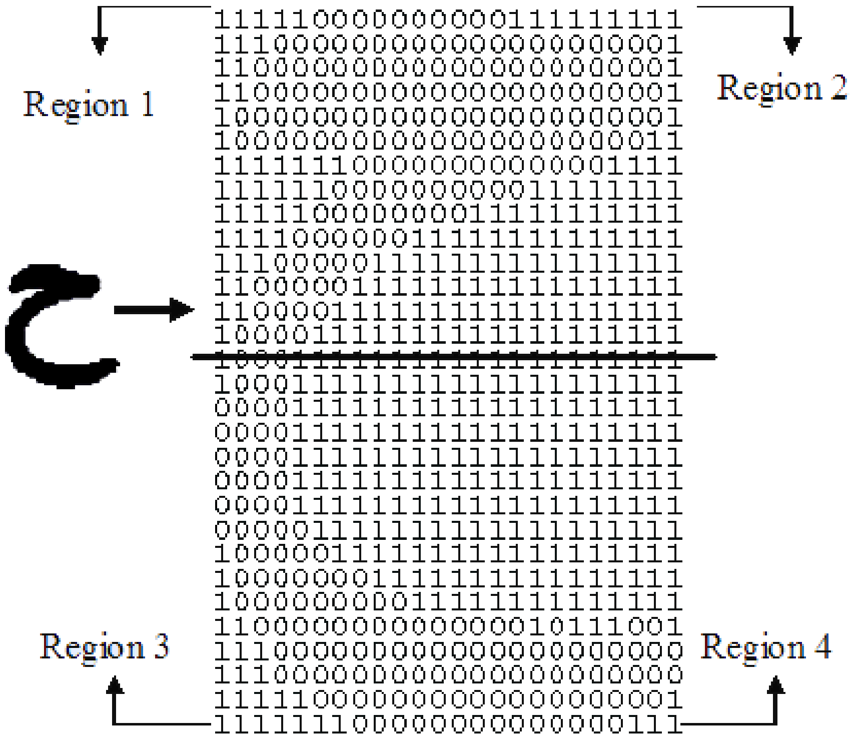
There are totally14 features extracted from the character of which 4 of them are for the whole image as listed below:

1) Height / Width
2) Number of black pixels / number of white pixels image
3) Number of horizontal transitions
4) Number of vertical transitions

In addition, the image is divided into four regions as shown in Figure 3 and the following features are extracted from these regions:

5) Black Pixels in Region 1/White Pixels in Region 1
6) Black Pixels in Region 2/White Pixels in Region 2
7) Black Pixels in Region 3/White Pixels in Region 3
8) Black Pixels in Region 4/White Pixels in Region 4
9) Black Pixels in Region 1/Black Pixels in Region 2
10) Black Pixels in Region 3/Black Pixels in Region 4
11) Black Pixels in Region 1/Black Pixels in Region 3
12) Black Pixels in Region 2/Black Pixels in Region 4
13) Black Pixels in Region 1/Black Pixels in Region 4
14) Black Pixels in Region 2/Black Pixels in Region 3

IV. The characters have evolved gradually over a course of time, leading to varying character sets from time to time. Hence the ***era to which an epigraphical script belongs to has been predicted*** so as to further enable automatic recognition of the epigraphical characters. For the classification of the script to their respective period, a **Support Vector Machine (SVM) Classifier using linear kernel function** is devised. The principle of SVM is to map the input data onto a higher dimensional feature space and determine a separating hyper plane with maximum margin between two classes in the feature space. We use SVM as a multiclass-classifier system which is trained on the distinct/unique characters in the Kannada script of various periods. Each vector is given to SVM which classifies it into a class in the vector space. The important criteria for deciding the final character are: the number of points belonging to a particular class and the minimum distance from a class. It represents the degree to which a particular snapshot resembles a character. We search for the unique (test) characters identified aprior, of various periods in the given ancient script. The number of matches of these unique characters of different era, with the characters of input document image is accounted. The period of input epigraphical document image of ancient Kannada script is the one with maximum number of matches of characters of input ancient script with unique (test) characters. Thus the input document image is classified to the period of any of these dynasties: Ashoka dynasty , Shatavahana dynasty , Kadamba dynasty ,Badami Chalukya dynasty , Rashtrakuta dynasty, Kalyan Chalukya dynasty, Hoyasala dynasty, Vijayanagar dynasty or Mysore Wodeyar dynasty.



```
111110000000000111111111
111000000000000000000001
110000000000000000000001
110000000000000000000001
100000000000000000000001
100000000000000000000011
111111100000000000001111
111111000000000011111111
111110000000011111111111
111100000011111111111111
111000001111111111111111
110000011111111111111111
110000111111111111111111
100001111111111111111111
100011111111111111111111
100011111111111111111111
000011111111111111111111
000011111111111111111111
000011111111111111111111
000011111111111111111111
000011111111111111111111
000001111111111111111111
100000111111111111111111
100000001111111111111111
100000000011111111111111
110000000000000010111001
111000000000000000000000
111000000000000000000000
111110000000000000000001
111111100000000000000111
```

Figure 3. Dividing the image to 4 regions and extracting features

V. **Display the era of the Script**

Thus the input epigraphical document image is classified to the period of any of the nine dynasties that prevailed in ancient Karnataka: Ashoka dynasty , Shatavahana dynasty , Kadamba dynasty ,Badami Chalukya dynasty , Rashtrakuta dynasty, Kalyan Chalukya dynasty, Hoyasala dynasty, Vijayanagar dynasty or Mysore Wodeyar dynasty

4. IMPLEMENTATION AND EXPERIMENTAL RESULTS

We have implemented the Age identification System on Window's Xp Operating System, using MATLAB version 7.8.0.347(R2009a) on a PENTIUM IV 2.6 GHz processor. The functionalities of Image processing toolbox is used for developing GUI, preprocessing and segmenting input document image.

In this section we provide sample experimental results out of several experiments conducted. We have tested the proposed method on more than 50 samples of ancient Kannada Epigraphical documents. Figure 4 to Figure 6 illustrates the snapshots of GUI developed, the output of segmentation and feature extraction, and output display of the classifier respectively. Figure 7 to Figure 9 depicts sample results of Classification and Display of the period for the given input epigraphical script.

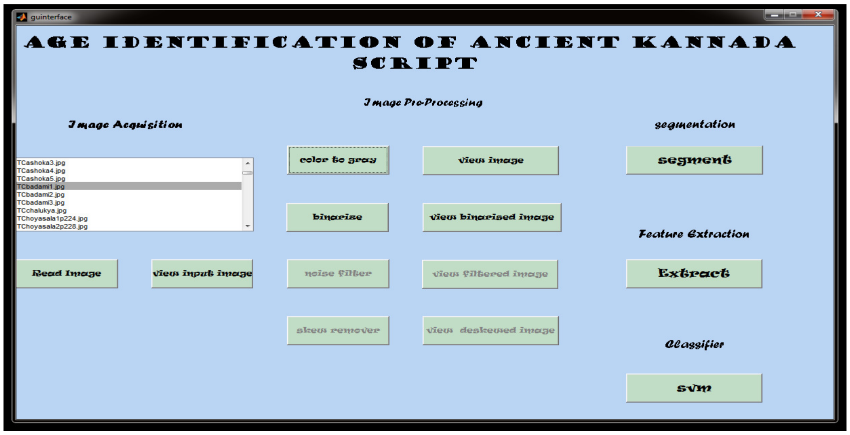
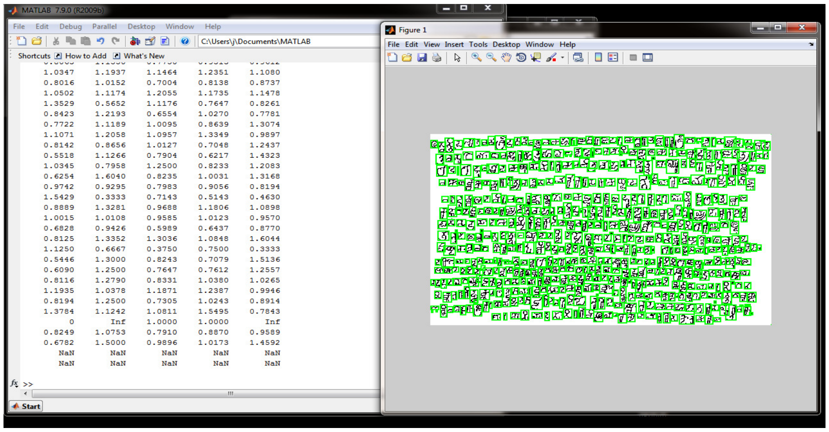


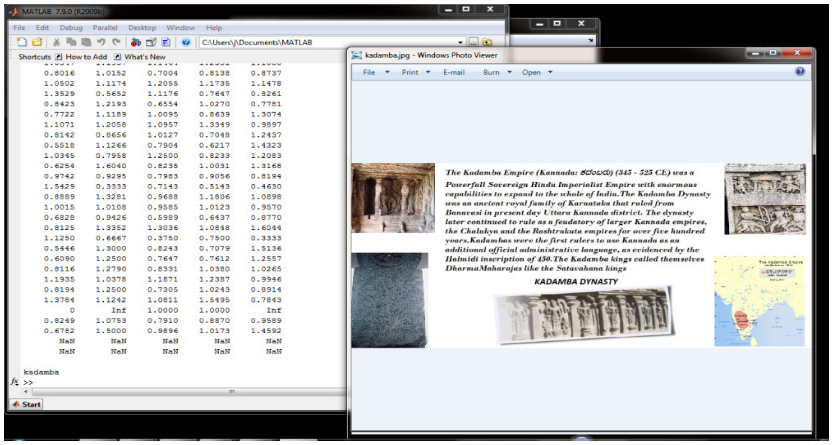
Figure 4. GUI of the Period prediction System for ancient Kannada Script


MATLAB 7.9.0 (R2009b)
File Edit Debug Parallel Desktop Window Help
C:\Users\j\Documents\MATLAB
Shortcuts How to Add What's New
Figure 1
File Edit View Insert Tools Desktop Window Help
Start

```
    1.0347    1.1937    1.1464    1.2351    1.1080
    0.8016    1.0152    0.7004    0.8138    0.8737
    1.0502    1.1174    1.2055    1.1735    1.1478
    1.3529    0.5652    1.1176    0.7647    0.8261
    0.8423    1.2193    0.6554    1.0270    0.7781
    0.7722    1.1189    1.0095    0.8639    1.3074
    1.1071    1.2058    1.0957    1.3349    0.9897
    0.8142    0.8656    1.0127    0.7048    1.2437
    0.5518    1.1266    0.7904    0.6217    1.4323
    1.0345    0.7958    1.2500    0.8233    1.2083
    0.6254    1.6040    0.8235    1.0031    1.3168
    0.9742    0.9295    0.7983    0.9056    0.8194
    1.5429    0.3333    0.7143    0.5143    0.4630
    0.8889    1.3281    0.9688    1.1806    1.0898
    1.0015    1.0108    0.9585    1.0123    0.9570
    0.6828    0.9426    0.5989    0.6437    0.8770
    0.8125    1.3352    1.3036    1.0848    1.6044
    1.1250    0.6667    0.3750    0.7500    0.3333
    0.5446    1.3000    0.8243    0.7079    1.5136
    0.6090    1.2500    0.7647    0.7612    1.2557
    0.8116    1.2790    0.8331    1.0380    1.0265
    1.1935    1.0378    1.1871    1.2387    0.9946
    0.8194    1.2500    0.7305    1.0243    0.8914
    1.3784    1.1242    1.0811    1.5495    0.7843
         0       Inf    1.0000    1.0000       Inf
    0.8249    1.0753    0.7910    0.8870    0.9589
    0.6782    1.5000    0.9896    1.0173    1.4592
       NaN       NaN       NaN       NaN       NaN
       NaN       NaN       NaN       NaN       NaN

>>
```

Figure 5. Snapshot of Segmentation and Feature extraction for input epigraphical document


MATLAB 7.9.0 (R2009b)
File Edit Debug Parallel Desktop Window Help
C:\Users\j\Documents\MATLAB
Shortcuts How to Add What's New
kadamba.jpg - Windows Photo Viewer
File Print E-mail Burn Open
Start

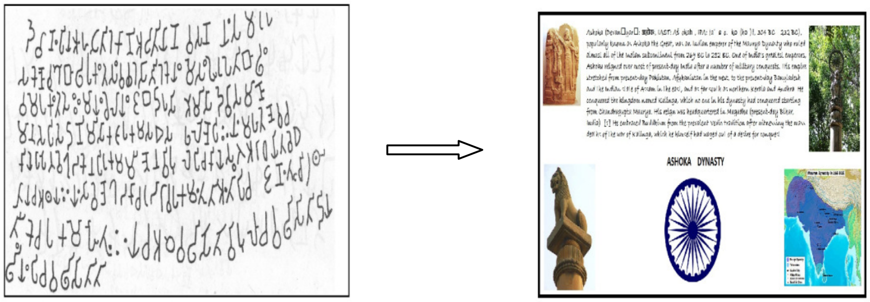
```
    0.8016    1.0152    0.7004    0.8138    0.8737
    1.0502    1.1174    1.2055    1.1735    1.1478
    1.3529    0.5652    1.1176    0.7647    0.8261
    0.8423    1.2193    0.6554    1.0270    0.7781
    0.7722    1.1189    1.0095    0.8639    1.3074
    1.1071    1.2058    1.0957    1.3349    0.9897
    0.8142    0.8656    1.0127    0.7048    1.2437
    0.5518    1.1266    0.7904    0.6217    1.4323
    1.0345    0.7958    1.2500    0.8233    1.2083
    0.6254    1.6040    0.8235    1.0031    1.3168
    0.9742    0.9295    0.7983    0.9056    0.8194
    1.5429    0.3333    0.7143    0.5143    0.4630
    0.8889    1.3281    0.9688    1.1806    1.0898
    1.0015    1.0108    0.9585    1.0123    0.9570
    0.6828    0.9426    0.5989    0.6437    0.8770
    0.8125    1.3352    1.3036    1.0848    1.6044
    1.1250    0.6667    0.3750    0.7500    0.3333
    0.5446    1.3000    0.8243    0.7079    1.5136
    0.6090    1.2500    0.7647    0.7612    1.2557
    0.8116    1.2790    0.8331    1.0380    1.0265
    1.1935    1.0378    1.1871    1.2387    0.9946
    0.8194    1.2500    0.7305    1.0243    0.8914
    1.3784    1.1242    1.0811    1.5495    0.7843
         0       Inf    1.0000    1.0000       Inf
    0.8249    1.0753    0.7910    0.8870    0.9589
    0.6782    1.5000    0.9896    1.0173    1.4592
       NaN       NaN       NaN       NaN       NaN
       NaN       NaN       NaN       NaN       NaN

kadamba
>>
```

The Kadamba Empire (Kannada: ಕದಂಬರು) (345 - 525 CE) was a Powerfull Sovereign Hindu Imperialist Empire with enormous capabilities to expand to the whole of India.The Kadamba Dynasty was an ancient royal family of Karnataka that ruled from Banavasi in present day Uttara Kannada district. The dynasty later continued to rule as a feudatory of larger Kannada empires, the Chalukya and the Rashtrakuta empires for over five hundred years.Kadambas were the first rulers to use Kannada as an additional official administrative language, as evidenced by the Halmidi inscription of 450.The Kadamba kings called themselves DharmaMaharajas like the Satavahana kings

KADAMBA DYNASTY

Figure 6. Snapshot of classification of input epigraphical document into respective period and corresponding dynasty



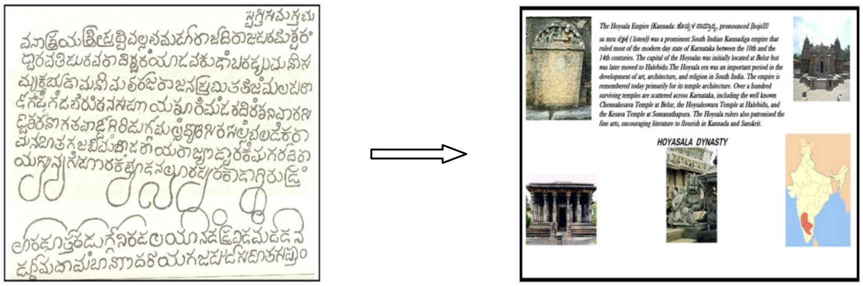
Ashoka (Devan□gar□: अशोक, IAST: Aś okah , IPA: [ɑˈ ɕ o. kɑ (kɑ)], 304 BC 232 BC), popularly known as Ashoka the Great, was an Indian emperor of the Maurya Dynasty who ruled almost all of the Indian subcontinent from 269 BC to 232 BC. One of India's greatest emperors, Ashoka reigned over most of present-day India after a number of military conquests. His empire stretched from present-day Pakistan, Afghanistan in the west, to the present-day Bangladesh and the Indian state of Assam in the east, and as far south as northern Kerala and Andhra. He conquered the kingdom named Kalinga, which no one in his dynasty had conquered starting from Chandragupta Maurya. His reign was headquartered in Magadha (present-day Bihar, India). [1] He embraced Buddhism from the prevalent Vedic tradition after witnessing the mass deaths of the war of Kalinga, which he himself had waged out of a desire for conquest.

ASHOKA DYNASTY

Input epigraphical document

Output image with details pertaining to the period predicted

Figure 7. Brahmi Script of Ashoka Dynasty being dated



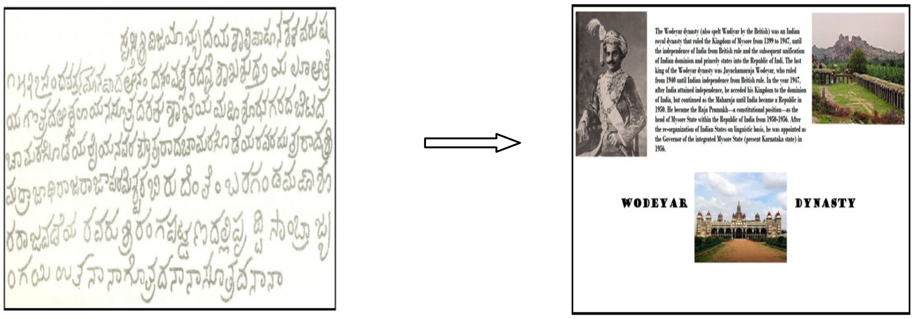
The Hoysala Empire (Kannada: ಹೊಯ್ಸಳ ಸಾಮ್ರಾಜ್ಯ, pronounced [hojsəɭə sa:mra:dʒjə] (listen)) was a prominent South Indian Kannadiga empire that ruled most of the modern day state of Karnataka between the 10th and the 14th centuries. The capital of the Hoysalas was initially located at Belur but was later moved to Halebidu.The Hoysala era was an important period in the development of art, architecture, and religion in South India. The empire is remembered today primarily for its temple architecture. Over a hundred surviving temples are scattered across Karnataka, including the well known Chennakesava Temple at Belur, the Hoysaleswara Temple at Halebidu, and the Kesava Temple at Somanathapura. The Hoysala rulers also patronised the fine arts, encouraging literature to flourish in Kannada and Sanskrit.

HOYASALA DYNASTY

Input epigraphical document

Output image with details pertaining to the period predicted

Figure 8. A Sample Kannada Script of Hoysala Dynasty being dated



The Wodeyar dynasty (also spelt Wadiyar by the British) was an Indian royal dynasty that ruled the Kingdom of Mysore from 1399 to 1947, until the independence of India from British rule and the subsequent unification of Indian dominion and princely states into the Republic of Indi. The last king of the Wodeyar dynasty was Jayachamaraja Wodeyar, who ruled from 1940 until Indian independence from British rule. In the year 1947, after India attained independence, he acceded his Kingdom to the dominion of India, but continued as the Maharaja until India became a Republic in 1950. He became the Raja Pramukh—a constitutional position—as the head of Mysore State within the Republic of India from 1950-1956. After the re-organization of Indian States on linguistic basis, he was appointed as the Governor of the integrated Mysore State (present Karnataka state) in 1956.

WODEYAR DYNASTY

Input epigraphical document

Output image with details pertaining to the period predicted

Figure 9. A Sample Kannada Script of Wodeyar Dynasty being dated

5. CONCLUSION & FUTURE ENHANCEMENT

Epigraphical sources reflect the importance during regime of several rulers. Script identification is a key step that arises in document image analysis, especially for the documents which belong to ancient days. The study of Epigraphical scripts is vital in knowing the civilized past and hence classification of character belonging to various periods is imperative before using the character bank of the particular period. The proposed system for the prediction of the era is being done by examining a few characters in Kannada script of various periods referred to as test characters. These test characters are sampled from the script automatically and matched with the characters available for different periods using machine intelligence. The proposed system here has various modules like Preprocessing and Segmentation (involves binarization, Noise removal, Skew removal & Segmentation), feature extraction and finally classification of input epigraphical documents into their respective eras using Support Vector Machine. This classifier is tested on more than 50 samples of Kannada epigraphical document images belonging to nine different periods. To sum up, the research issue taken here is to produce a computer perceivable image from a raw epigraphical script which are the inscriptions on rocks or pillars or plate and then predict the era of the ancient script. Prediction of period of ancient scripts is the first step in automatically deciphering epigraphical scripts. Automatic period identification for a given document image, of a script facilitates the selection of the script specific OCR in an environment where scripts of various periods are given as input. The work can be further extended to facilitate the computer dating of inscriptions of various other scripts that were prevalent in ancient India and other non Indian regions, during the regime of various rulers.

REFERENCES

1. D Dayalan *Computer Application in Indian Epigraphy,* Bharatiya Kala Prakashan publication (2005).
2. A.V.Narasimha Murthy - *'Kannada Lipiya Ugama Mattu Vikasa'*, Kannada Adhyayana Samsthe, Mysore University, Mysore, (1968).
3. Dr M G Manjunath, G K Devarajaswamy – '*Kannada Lipi Vikasa'*, Jagadhguru Sri Madhvacharya Trust, Sri RagavendraSwamy Matt, Mantralaya
4. Dr. Devarakonda Reddy : *Lipiya Huttu Mattu Belavanige* — Origin and Evolution of Script, Published by Kannada Pustaka Pradhikara (Kannada Book Authority), Bangalore
5. Anasuya Devi H.K, (2002), "*Automated Recognition of Ancient Indian Scripts*", Proceedings of National workshop on Computer Vision, Graphics and Image processing, WVGIP, pp 216-219.
6. K Harish Kashyap, Bansilal, P Arun Koushik – '*Hybrid Neural Network Architecture for Age Identification of Ancient Kannnada Scripts', 2003*
7. Haidar Almohri. John S. Gray and Hisham Alnajjar – '*A Real-time DSP-Based Optical Character Recognition System for Isolated Arabic characters using the TI TMS320C6416T'***,** Proceedings of the IAJC-IJME International Conference, 2008
8. Works on OCR for Indian Languages [online] http:// Indira Gandhi National Centre for the Arts (IGNCA's) Southern Regional Centre, Bangalore
9. Gopal Datt Joshi, Saurabh Garg, and Jayanthi Sivaswamy – '*Script Identification from Indian Documents',* H. Bunke and A.L. Spitz (Eds.): DAS 2006, LNCS 3872, pp. 255–267, 2006. c Springer-Verlag Berlin Heidelberg 2006
10. Manish Kumar – *'Degraded Text recognition of Gurumukhi Script',* Ph.D thesis, Thapur University, Patiala (Punjab), March 2008
11. Naveen Garg – '*Handwritten Gurumukhi Character Recognition using Neural networks*', M.E thesis, Thapur University, Patiala, June 2009
12. T V Ashwin and P S Sastry – ' *A Font and size-independent OCR system for printed Kannada documents using Support Vector Machines*', Sadhana, Vol 27, Part 1, February 2002

13. S Pletschacher, J Hu and A Antonacopoulos- ' *A New framework for Recognition of Heavily Degraded Characters in Historical Typewritten Documents Based on Semi-Supervised Clustering',* 10th International Conference on Document Analysis and Recognition, 2009
14. Sheikh Faisal Rashid, Faisal Shafait and Thomas M. Breuel – *'Connected Component level Multiscript Identification from Ancient Document Images*' **,** Copyright 2010 by the author(s)
15. P. Subashini, M. Krishnaveni and N. Sridevi – *' Period Prediction System for Tamil Epigraphical Scripts Based on Support Vector Machine'*, Information Systems for Indian Languages - Communications in Computer and Information Science, Volume 139, Part 1, 23-30, 2011.
16. Srikanta Murthy.K – **'***Transformation Of Epigraphical Objects Into Machine Recognizable Image Patterns'*, Ph.D Thesis, University Of Mysore, Mysore, December 2005
17. Srikanta Murthy.K - *'Prediction of Era of Epigraphical scripts based on texture features-A statistical approach'*, Journal of Analysis and Computation (JAC), New Delhi, 2006
18. Srikanta Murthy.K *"Invariant Moments Based Feature Extraction for the Classification of Epigraphical Characters"*, Journal of Statistics and Applications, 2005
19. Srikanta Murthy.K , *"Neural Network classifier for the prediction of an Epigraphical script",* Journal of Computer Society of India (CSI), 2004
20. Srikanta Murthy.K *"Prediction of Era of Epigraphical scripts based on texture features-A statistical approach"*, National conference on convergence of linguistics, E-Governance & I.T. for making kannada- A Tech-savvy Language, Shimoga 2006

Authors

Soumya A

Soumya A, Assistant Professor, Department of Computer Science and Engineering, R.V College of Engineering, Bangalore. She has obtained M.S degree in Computer Cognition Technology, University of Mysore, Mysore and B.E in Computer Science and Engineering, Bangalore University, Bangalore. Her areas of research are Artificial Intelligence, Soft Computing, Pattern Recognition and Image Processing. She has guided several Under- graduate Projects and 7 Post Graduate projects.

Dr G Hemantha Kumar

Dr G Hemantha Kumar, Professor & Chairman, Department of studies in Computer Science, University of Mysore, Mysore. Serving as Course - Coordinator of Chinese B.Tech Programme, University of Mysore, Mysore. He was awarded Ph.D. for his thesis titled : "On Automation of Text Production From Pitman Shorthand Notes" from University of Mysore, Mysore. His areas of research are: Image Processing, Pattern Recognition, Numerical Techniques, Bio-Metric. He has to his credits 31 publications in International / National Journals and 46 publications in International / National Conferences/Workshops. He has guided several Ph.D candidates and is presently guiding 5 Ph.D candidates.