LITERATURE SURVEY

A Novel Method for Handwritten Digit Recognition System

TEAM ID: PNT2022TMID42552

DONE BY:

YUVASRI.D

DHIVYA.P.V

ANAND KUMAR.S

SELVAMARI

GANESH.R

CASE STUDY I

Tittle: A novel approach of handwritten character recognition using positional featutre extraction

AUTHOR: SA Vaidya, BR Bombade

YEAR: 2013

ABSTRACT:

In this paper, we have presented a method of feature extraction for handwritten character recognition. We showed that our method, despite its simplicity, yields good classification results on handwritten characters. Normalization and binarization are the pre-processing techniques used for getting accurate results of classification process in handwritten character recognition. This method is based on positional properties of each pixel present in character image. First of all, add all the sample character image matrices and divide the resultant matrix by total number of matrices added, called as Avg_matrix. Then subtract it from each sample character image matrix, which results in unique features because of their positional properties of pixels present in that image. We used singular value decomposition technique to get projection vector matrix which is beneficial for more accuracy in results. Finally, we used generalized regression neural network for resulting feature vectors and obtain classification performance in the character recognition task. The proposed recognition scheme provided 82.89 percent and 85.62 percent accuracies on Devnagari and Kannada character databases respectively

CASE STUDY II

TITLE ::: A new database of Arabic historical handwriting and a novel approach for subwords extraction

AUTHOR : A Zoizou, A Zarghili, I Chaker

YEAR: 2022

ABSTRACT:

The digitalization of historical documents is vital to preserving their content and the historical memory of nations. Although, the results of historical Arabic handwritten text recognition and word spotting are still unsatisfactory. The increasing research efforts during the last few years are still not sufficient since handwriting recognition systems rely heavily on robust databases. In this paper, we present a new contour-based method of subword extraction from Arabic historical documents and a novel database of Arabic historical subwords MOJ-DB. The proposed method of subword extraction includes a process of touching components resolving. It proved high performance and consistency while tested on different databases and compared with other methods from the literature. The proposed database contains 560000 subwords distributed on 5600 different classes. It was built using 64 pages extracted from 10 books written in the 17th and 16th centuries. MOJ-DB database is divided into three sets; 70%,20%, and 10% for training, testing, and validation, respectively. Ground truth is established iteratively to guarantee minimal error. It includes information about the subword as of the sourcebook and page. We conducted several experiments to verify the robustness of the proposed database as well as the validity of the segmentation process. The database is freely available for the public research community. It can be used for word and subword recognition, word spotting, subword extraction, and database construction.

CASE STUDY III

TITLE : An application to artificial intelligence for biometric identification

AUTHOR : Z Akhtar, JW Lee, MA Khan, M Sharif

YEAR: 2020

ABSTRACT:

Purpose

In artificial intelligence, the optical character recognition (OCR) is an active research area based on famous applications such as automation and transformation of printed documents into machine-readable text document. The major purpose of OCR in academia and banks is to achieve a significant performance to save storage space.

Design/methodology/approach

A novel technique is proposed for automated OCR based on multi-properties features fusion and selection. The features are fused using serially formulation and output passed to partial least square (PLS) based selection method. The selection is done based on the entropy fitness function. The final features are classified by an ensemble classifier.

Findings

The presented method was extensively tested on two datasets such as the authors proposed and Chars74k benchmark and achieved an accuracy of 91.2 and 99.9%. Comparing the results with existing techniques, it is found that the proposed method gives improved performance.

Originality/value

The technique presented in this work will help for license plate recognition and text conversion from a printed document to machine-readable.

CASE STUDY IV

TITLE : Recognition of Devanagari handwritten numerals

AUTHOR : M Jangid, R Dhir, R Rani

YEAR: 2011

ABSTRACT:

Recognition of Indian languages scripts is challenging problems. In Optical Character Recognition [OCR], a character or symbol to be recognized can be machine printed or handwritten characters/numerals. There are several approaches that deal with problem of recognition of numerals/character depending on the type of feature extracted and different way of extracting them. In this paper an automatic recognition system for isolated Handwritten Devanagari Numerals has been proposed. The proposed system relies on a feature extraction technique based on recursive subdivision of the character image so that the resulting sub- images at each iteration have balanced numbers of foreground pixels as possible. Support Vector Machine (SVM) is used for classification. Accuracy of 98.98% has been obtained by using standard dataset provided by ISI (Indian Statistical Institute) Kolkata.

CASE STUDY V

TITLE :Recognition of similar shaped handwritten Marathi characters using artificial neural network.

AUTHOR: AP Jane, MA Pund

YEAR: 2012

ABSTRACT:

The growing need have handwritten Marathi character recognition in Indian offices such as passport, railways etc has made it vital area of a research. Similar shape characters are more prone to misclassification. In this paper a novel method is provided to recognize handwritten Marathi characters based on their features extraction and adaptive smoothing technique. Feature selections methods avoid unnecessary patterns in an image whereas adaptive smoothing technique form smooth shape of charecters. Combination of both these approaches leads to the better results. Previous study shows that, no one technique achieves 100% accuracy in handwritten character recognition area. This approach of combining both adaptive smooth.