

TTO - IPM Cell



Industrial Consultancy & Sponsored Research (IC&SR)

A SYSTEM AND METHOD FOR RECOGNITION OF HANDWRITTEN TELUGU **CHARACTERS**

IITM Technology Available for Licensing

PROBLEM STATEMENT

- Generally, Handwritten character recognition can be performed either online or offline.
 - Based on prior arts survey, a Hidden Markov Model based system is used for online character recognition in Telugu & report a top-I accuracy of 91.6%, but said system operates at symbol level, not at character level.
- Further there are other issues including accuracy to identify characters both in online and offline mode.
- Hence, there is a need to mitigate above challenges, and present invention provides the solution in efficient manner.

INTELLECTUAL PROPERTY

IITM IDF Ref. 1023; IN Patent No: 387357

TECHNOLOGY CATEGORY/ MARKET

Technology: Method for recognition

handwritten Telegu Character;

Industry: Banking Sector, Security; Applications: Banking Sector & others;

Market: The global **optical** recognition market is projected to grow at a

CAGR of 14.8% during 2023-2027.

TRL (TECHNOLOGY READINESS LEVEL)

TRL-3/4, Proof of Concept ready, tested in lab

TECHNOLOGY

- The present invention describes a **method &** system for recognition of handwritten characters based on Convolutional Neural Networks (CNN).
- Each network comprises a first, second, third, & a fourth hidden layers of neurons connected to each other.
- **method** of recognizing handwritten characters is depicted in the smart chart and figures.

- The handwritten characters are scanned into an input image & processed by Convolutional neural networks.
- **Principal component analysis** (PCA) system is to identify the output class to which the characters belong.
- The support vector mechanism (SVM) is configured to determine support vectors & identify the output class based on the determined support vectors, & also to train weight stage from the last hidden layer of the CNN to the output layer.
- The entire hierarchy of CNN layers from the input to the last hidden layer may then be considered as a kernel layer of the SVM.

RECEIVE INPUT IMAGE FOR PROCESSING BY A NEURAL NETWORK HAVING FOUR HIDDEN LAYERS 702

DETERMINE WEIGHT VECTOR OF FIRST AND THIRD HIDDEN LAYERS BY USING PRINCIPAL COMPONENT ANALYSIS TECHNIQUE & PERFORM SUB-SAMPLING ON THE TRAINED FIRST AND THIRD HIDDEN LAYERS TO GENERATE SECOND AND FOURTH HIDDEN LAYERS RESPECTIVELY 704

DETERMINE SUPPORT VECTORS OF THE FOURTH HIDDEN LAYER USING SUPPORT VECTOR MACHINE (SVM) TECHNIQUE 706

IDENTIFY OUTPUT CLASS OF THE INPUT IMAGE BASED ON THE DETERMINED SUPPORT VECTORS 708

Fig.1: Illustrates the claimed method for recognition of handwritten characters

RESEARCH LAB

Prof. Srinivasa Chakravarthy V, Department of Biotechnology

CONTACT US

Dr. Dara Ajay, Head Technology Transfer Office, IPM Cell- IC&SR, IIT Madras **IITM TTO Website:**

https://ipm.icsr.in/ipm/

Email: smipm-icsr@icsrpis.iitm.ac.in

sm-marketing@imail.iitm.ac.in

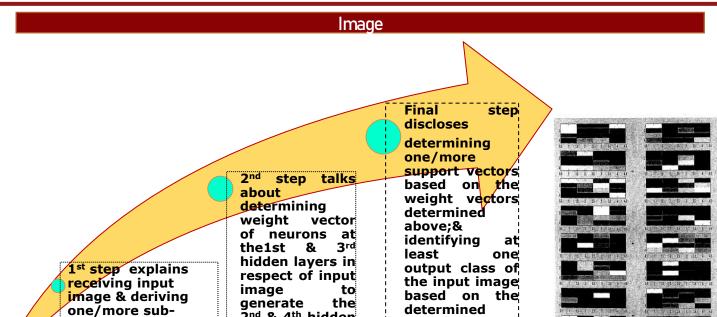
Phone: +91-44-2257 9756/ 9719



IIT MADRAS Technology Transfer Office TTO - IPM Cell



Industrial Consultancy & Sponsored Research (IC&SR)



support

vectors;

Smart Chart. A

2nd & 4th hidden

respectively.

lavers

KEY FEATURES / VALUE PROPOSITION

images from the

hidden layer

received input image

to generate the first

* Technical Perspective:

- Data collected splits into training & testing data wherein 90% data used for training & remaining for **testing**.
- **Training data** divides into two sets & used to train nine different networks.
- A classifier is created by combining the with least correlation networks performance on the test dataset. Further, a new classifier pair yielded higher accuracy than individual classifiers.
- Facilitates the feature of supporting recognition of handwritten not only Telegu character but also support other Indian regional language like Tamil, Kannada & Malayalam.

* Industrial Perspective:

- · User friendly with high accuracy & costeffective system & method.
- · Very fast process & consumes less time for recognition of Indian Language character (Telegu).

Flowchart

SELECT A STACK OF INPUT IMAGES AS TRAINING IMAGES 802

DETERMINE PRINCIPAL COMPONENTS OVER THE STACK OF SUB-IMAGES IN THE INPUT IMAGE CONNECTED TO NEURON AT LOCATION (I, J) IN ANY SUB-LAYER OF THE FIRST HIDDEN LAYER 804

FOR EACH NEURON AT LOCATION (I,J) IN EACH SUB-LAYER, ASSIGN THE KTH PRINCIPAL COMPONENT TO THE WEIGHT OF THE NEURON AT (I,J) IN KTH SUB-LAYER 806

GENERATE SECOND HIDDEN LAYER BY SUB-SAMPLING THE SUB-LAYERS OF THE FIRST HIDDEN LAYER AND COMPUTE THE RESPONSES OF THE SECOND HIDDEN LAYER TO ALL INPUT IMAGES IN THE TRAINING SET 808

DETERMINE PRINCIPAL COMPONENTS (125 DIMENSION VECTOR) OVER THE STACK OF SUB-IMAGES OF EACH OF THE SUB-LAYER BASED ON THE COMPUTED RESPONSES OF THE SECOND HIDDEN LAYER 810

FOR EACH NEURON AT LOCATION (I,J) IN EACH SUB-LAYER, ASSIGN THE KTH PRINCIPAL COMPONENT TO THE WEIGHT OF THE NEURON AT (I, I) IN KT LAYER 812

Fig. 2: Illustrates a flowchart of a method of obtaining weights for neurons at the fIrst and third hidden layers using Principal Component Analysis technique

CONTACT US

Dr. Dara Ajay, Head Technology Transfer Office, IPM Cell- IC&SR, IIT Madras

Email: smipm-icsr@icsrpis.iitm.ac.in **IITM TTO Website:** sm-marketing@imail.iitm.ac.in https://ipm.icsr.in/ipm/

Phone: +91-44-2257 9756/ 9719