Poornaprajna Institute of Management

Udupi - India

Micro Research Centre (MRC)

Centre for Computer Vision

Prof. Venugopala Rao A S, Assistant Professor

Name of Institution: Poornaprajna Institute of Management

1. Purpose of MRC:

The Center for Computer vision is set up with the objective of applying the computer vision and image processing technologies to cater the societal issues. It is also aimed to develop new technologies in computer vision

2. Objective of MRC:

The center for Computer Vision set up with the following objectives:

1. **Fundamental Research:** Investigating the underlying principles and theories of computer vision. This involves exploring new algorithms, models, and techniques for image and video analysis, recognition, and interpretation.
2. **Application Development**: Creating practical applications and systems that leverage computer vision technology. This can range from medical imaging and autonomous vehicles to surveillance systems and augmented reality.
3. **Interdisciplinary Collaboration**: Working with experts from other fields such as machine learning, robotics, Healthcare sector, and cognitive science to develop more comprehensive and effective computer vision solutions.
4. **Technology Transfer:** Bridging the gap between academic research and industry by developing technologies that can be commercialized. This includes working with industry partners to bring innovations from the lab to the market.

5. **Education and Training:** Providing education and training opportunities for students, researchers, and professionals. This often involves organizing workshops, seminars, and courses to disseminate knowledge and skills in computer vision.
6. **Publication and Dissemination**: Publishing research findings in leading journals and conferences, and presenting at academic and industry events to share insights and advancements with the broader community.

3. Description on Proposed Research:

There are various applications of computer vision and handwriting recognition is one of them. The handwriting recognition systems are classified into two basic types: online and offline. In our MRC we mainly focus on offline approach. We adopt the following methods and techniques for handwriting text recognition

Offline Character Recognition Algorithms

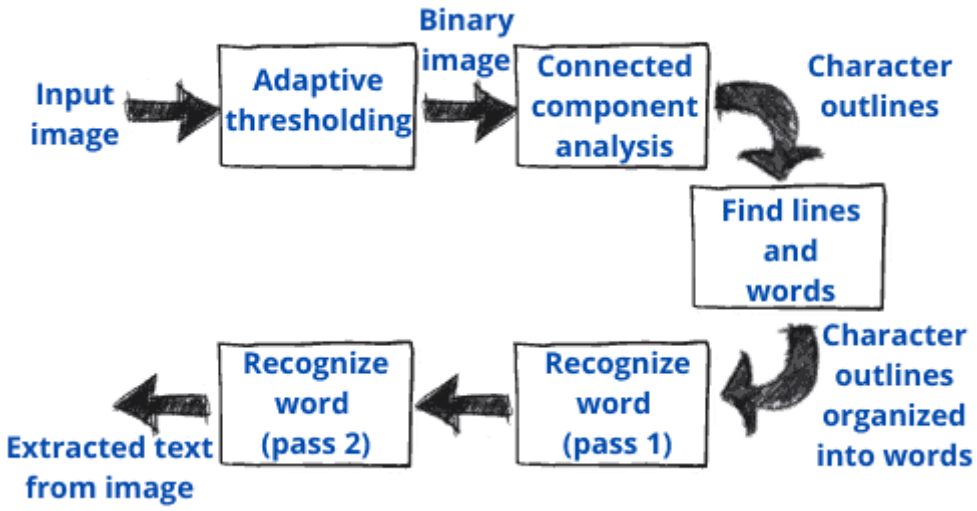
The offline approach involves the automatic conversion of the input into letter codes. Then, we use these letter codes in computer and text-processing programs. There are traditional and modern methods for the offline approach.

Traditional Methods

The traditional method of offline handwriting recognition consists of three steps:

- **Character extraction:** This implies that the individual characters in the scanned picture must be extracted. The most typical flaw is when connected characters are returned as a single sub-image comprising both characters. This creates a significant issue at the recognition step.
- **Character recognition:** Following the extraction of individual characters, we use a recognition engine to identify the corresponding computer character.
- **Feature extraction:** Feature extraction is comparable to neural network recognizers in its operation. However, we must choose which properties are relevant to the given context. As a result, feature extraction is not a completely automated operation.

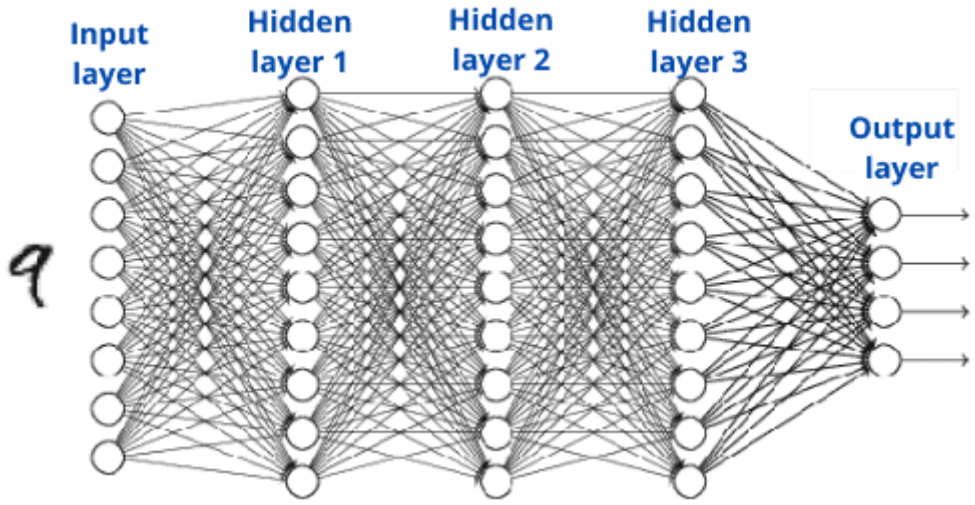
Here's an example of the classical processing for text recognition using an image as input:



3.2. Modern Methods

Whereas traditional approaches focus on identifying individual characters, modern approaches focus on segmenting all the characters in a segmented line of text. So, these approaches concentrate on machine learning algorithms that can learn visual characteristics rather than the previously employed limited feature engineering.

These modern approaches employ convolutional networks to extract visual information from many overlapping windows of an input. So, the feature extraction is automated and the convolutional network doesn't need feature engineering. Then, we use a recurrent neural network to generate character probabilities from the visual information. The image below represents an example of a deep learning algorithm:



Text Recognition Techniques

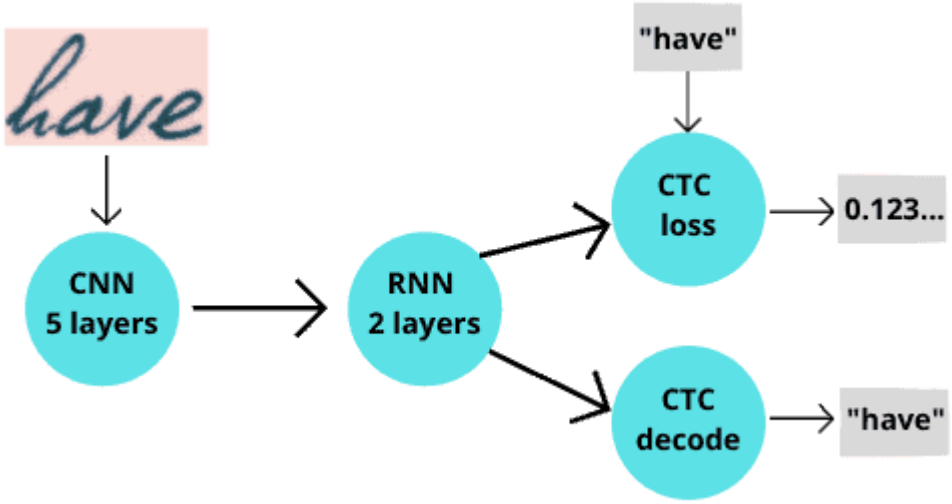
The handwriting recognition system requires handling a "stream of data" (a single input is mapped onto multiple outputs) since the number of symbols in output usually varies. For example, we have one input (for example an image as input) and we want to have the group of characters, words, and lines. This means that once we detect a character, the next cannot be a random character. In other words, we are dealing with sequential data.

So, if we try to rephrase this statement it's like we have one input and a lot of outputs. There are different solutions to develop a "stream of data" for Handwriting recognition:

- **Recurrent neural networks (RNNs)**: The main disadvantage of RNNs is that the previous state is always required for computing the current one. This means that the graphic processor unit (GPU) must focus the majority of its resources on a specific segment of the task rather than spreading to the largest possible scale. This means that training is not well parallelizable.
- **Transformers**: This method employs an essential notion known as "attention." As a result, training speed and inference are two of the primary benefits over RNN.

Model Overview

The neural network model for the HTR system consists of convolutional neural network (CNN) layers, recurrent neural network (RNN) layers, and a final Connectionist Temporal Classification (CTC) layer. Here's an overview of the HTR system:



Operation

The HTR system consists of three phases:

1. **Convolutional Neural Network (CNN) layers:** The purpose of this phase is to extract relevant features from the images. This phase consists of five layers and each layer consists of three operations. First, we apply a filter kernel of size 5X5 in the first two layers and 3X3 in the last three layers. Then, we use the ReLU as the activation function. Finally, we use the pooling layer to summarize the regions of each image.
2. **Recurrent Neural Network (RNN) layers:** We use the Long Short-Term Memory (LSTM) version of RNNs, which can propagate information over larger distances and has robust training characteristics.
3. **Connectionist Temporal Classification (CTC) layer:** Following the integration of CNN and RNN, the model can be trained with the loss developed by Alex Graves, known as CTC. The CTC receives only the matrix and decodes it into the final text.

4. Expected Outcome:

The proposed system is expected to identify the English handwritten text with an accuracy of at least 90 %

5. List of the Team Members :

1. Venugopala Rao A S
2. Priya K
3. Sneha Radhakrishnan

6. List of Working Papers:

NIL

7. List of related Published Papers in Journals, Proceedings, Book Chapters, Magazines by Coordinator & his/her Group members year-wise in APA format.

NIL

Name & Signature of Coordinator with date.