REVIEW OF RESEARCH
ISSN: 2249-894X
IMPACT FACTOR : 5.2331(UIF)
VOLUME - 7 | ISSUE - 5 | FEBRUARY - 2018

CHARACTER RECOGNITION USING DEEP LEARNING ALGORITHM

Sidagouda Patil
Principal
VSM BBA and BCA College, Nipani, Karnataka.

ABSTRACT

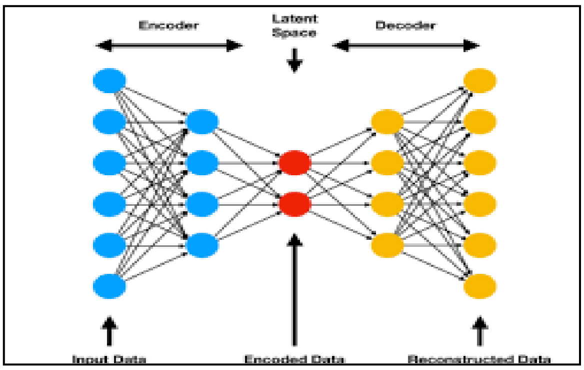
Deep learning character recognition algorithms are discussed in this paper. It is now easier to train detailed neural networks thanks to the massive amount of data and algorithms. Another application that makes use of alternative methods is character recognition. A character recognition technique based on CNN (Convolutional Neural Network) is presented in this paper. To see the results, train your dataset and classify it. Image segmentation was used in the design of a system for recognizing characters. Python is a programming language that aids in character recognition.

KEY WORDS: *Datasets, OCR, deep learning, neural networks, Python, and deep learning .*

1. INTRODUCTION

In today's highly technological world, huge data volumes and requirements necessitate a large data storage input into computers for all purposes. With the assistance of deep learning networks, various AI (artificial intelligence) applications are utilized worldwide. because of the high cost and complexity of the DNNS. Deep learning methods are widely used to improve AI systems' efficiency without sacrificing accuracy or increasing hardware costs. After learning on a large amount of data and processing it to extract features, the performance of DNNs comes from the raw data. However, in order to achieve greater accuracy, the DNNs have a high level of complexity. While graphical processing units are required for DNN processing, general-purpose computers are essential for DNN computation.

Characters and words are typically recognized by OCR technology, which then computes them to determine the information and transform them into computerized characters. because OCR integrates with the majority of research areas, including computer graphics, signal processing, and others. The process of converting images of various characters into machine-encoded text or text-imposed images is known as optical character recognition. Because OCR is widely used for a variety of applications, such as automatic data entry for passports, computerized receipts for a variety of purposes, business cards, and so on. Since earlier versions needed to be trained on a variety of images before using a single font, subsequent versions use a variety of image fonts. Nowadays, a wide range of image file format inputs are capable of high-level font classification. There are AS systems that can reproduce formatted output and approximate the original images. Telegraphy, which was used for optical character recognition, was an early form of technology. In 1914,



Emmanuel Goldberg developed a method for reading characters, which are then transformed into telegraph codes. In this project, we take a picture with characters like words and digits and process it further to create digital characters. As the project progressed, it combined the training of the neural network with an algorithm for segmenting the characters' images within a given image that is processed by the neural network. The full-featured model then enables the end user to convert various characters into digitized output by adding layers to take the form available to them. In order for the word segmentation to take place, the layers will need to be added. How we approach this problem Compared to image features or parts that have a complete word image, CNNs typically perform better with raw input pixels. To categorize and identify various images, make use of the deep learning methods that have been suggested.

2. ABOUT ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

2.1 Artificial Intelligence

John McCarthy coined the term "artificial intelligence" in 1956, describing it as "the science and engineering of creating intelligent machines." AI is a computerized recreation of human intelligence. Expert systems, natural language processing, and speech are just a few of AI's unique computer system applications. Machine vision and cognition. The simulation of various resources to imitate human behavior is known as AI. The human mind may also be referred to by the term, such as machine learning or problem-solving abilities. Machine learning, neural network expert systems, and deep learning, as well as robotics, are examples of AI technologies and algorithms. At the moment, AI technology is making a big difference all over the world in areas like medicine, space, robotics, and the military. Computational intelligence, or methods like neural networks, fuzzy systems, evolutionary computations, and other computational models, are used to circumvent some of the limitations of symbolic AI. Top-down and bottom-up are the two main methodologies in AI, despite the fact that the methods used vary. While the hierarchical methodology trusts in making PC programs as per the activity of the human cerebrum, and blueprints of the framework are laid out and determined, the granular perspective is acknowledged by giving comparative electronic replication to the brain organizations of the human mind.

2.2 Machine Learning

The subfield of artificial intelligence (AI) known as machine learning (ML) is able to learn and improve an application's explicit programming by running it automatically or on a regular basis (or) by improving over time through machine learning (ML) experiences. the study of algorithms in computers. data. It is being considered an artificial intelligence component. Without having to be explicitly programmed, machine learning algorithms construct models from training data and make predictions and judgments. Algorithms for machine learning, medicine, email filtering, computer vision, and other things are being used in a wide range of situations where it is either difficult or impossible to create conventional algorithms that can carry out the required tasks. a subfield of machine learning that is closely related to computer statistics and focuses on how computers make predictions. However, statistical learning is not the only type of machine learning. Methods, theories, and examples for machine learning are provided by mathematical optimization research. Data mining is a related field of study that uses machine learning and focuses on exploratory data analysis when applied to business issues.

The process of figuring out how computers can do things without having to be programmed is known as machine learning. Using the information provided, you can learn about your computer and carry out specific tasks here. You can program an algorithm that tells a computer how to perform all necessary steps to solve a problem for simple tasks assigned to it. The computer part doesn't have to teach you anything. The necessary algorithms can be created manually for more difficult tasks. Rather than letting a human programmer specify each necessary step, it is more efficient to assist with the machine's algorithm development. Available. Display some of the correct answers as valid if there are many possible answers. To enhance the algorithms that computers use to determine the correct answer, these data can be used as training data. The MNIST handwritten digit dataset, for instance, is frequently utilized for the training of single-task digital character recognition systems.

2.3 Deep Learning

Deep Learning is a feature of artificial intelligence (AI) that imitates the human brain's ability to process data and generate decision-supporting patterns. Also known as deep neural networks or deep chi learning. A subset of machine learning algorithms known as deep learning employs multiple layers to gradually extract high-level features from a primary input. For instance, picture handling can recognize lower layer edges and upper layers can distinguish human ideas like numbers, letters, and faces. When processing unstructured data, deep learning is extremely effective because it can handle a large number of features. Deep learning algorithms, on the other hand, are likely to be too much for simpler problems because they need to access a lot of data to be effective. Algorithms for deep learning attempt to extract high-level features from data. This is a crucial step in conventional machine learning and a characteristic of deep learning. As a result, deep learning can make it easier to create new feature extractors for every problem. Deep learning methods for putting deep neural networks into action are being driven by the rise in facilities for high-performance computing. Because it can handle multiple functions, deep learning has more power and flexibility when working with unstructured data. Algorithms for deep learning transmit data to multiple layers. Features can be passed to the next layer step by step from each layer. A complete expression is created by combining the features of subsequent layers and extracting low-level features in the first layer. Deep learning is still developing rapidly. Nonetheless, deep learning can be used to solve the issue at hand. Deep learning's structure remains a mystery, but it has the potential to improve computers' intelligence. You might even be more intelligent than we are. In order to make applications smarter and more intelligent, the current task is to develop deep learning models in conjunction with mobile. In the area of machine learning and pattern recognition, the following are some examples of OCR solutions. The technology behind optical character recognition can be broken down into two categories: conventional approaches, which include deep learning techniques and neural network models built from scratch. Homemade feature extraction and classification are used in the current methods. We are unable to guarantee a high recognition rate or accuracy due to the fragility of these homemade features. Due to the higher dimensions of these features, it is also computationally intensive. As a result, discrimination decreases. Convolutional neural networks (also known as CNNs), recurrent neural networks (also known as RNNs), and recurrent convolutional neural networks (also known as RCNNs) are powerful methods that can be used with deep learning algorithms to speed up training and improve models.

3. DIFFERENT CHARACTER RECOGNITION TECHNIQUES

The majority of the work involved in character recognition consists of a number of distinct approaches that involve a number of distinct steps. These approaches include checking images and determining the characters that are present in each text area. To identify words and characters in images, it makes use of a standard deep learning model. There are excellent learning models for character identification as deep learning models. As a result, special deep learning models have been developed to aid in image detection and localization. These are some of the most common approaches.

The RAM model and various models are utilized for character recognition. The RAM model, or Recurrent Attention Model, is designed so that when a new scene is shown to the human eye, a specific part of the image draws the line of sight. As the first model, other models can recognize characters by focusing on the information's "slip." As filtering and glyph vectors are created for the most important aspects of each cut version, model images are provided in a variety of sizes around a common center. A "glyph network" based on visual attention moves these glyph vectors through a flattening process. After that, the Glimpse vector is sent to the location network, where the RNN is used to predict the next area of the image that needs to be noted. For a closer look at the network, this is the next input. To ensure that the previous glimpse of information is sufficient to achieve a high level of accuracy each time the error station radio wave method is carried out, the model gradually moves additional parts of the image.

An OCR project for the problem of original image captions can be used with Tens or Flow in the second method via Attention. Caution: Include a decoder first, then a CRNN. To begin, the model extracts image features through a complex network layer. It uses a week mechanism borrowed from the Seq2Seq machine translation model and encodes these functions as a string before passing them to the RNN. Note that the prediction of the input image's text is made by the base decoder. In addition to being more effective, these two methods are also less accurate.

The Convolutional Recurrent Neural Network (CRNN), the final and most significant strategy. To identify a word, the CRNN method makes use of three fundamental steps. Processing an image with a convolutional neural network (CNN) is the first step. The image is divided into "feature columns" and features on the first layer. Later, deep bidirectional LSTM (long-term and short-term memory) cells receive these columns, which provide a sequence for determining character relationships. Finally, the LSTM cell output is sent to the transfer layer, where a stochastic method is used to organize the output and proceeds through a string with duplicate characters.

4. PROPOSED CNN ARCHITECTURE

There are two main parts to CNN architecture. First, a fully connected layer-previous step that uses the convolution tool and the output of the convolution process to separate and identify the features in an image, analyze them in a process called feature extraction, and then predict the class of the image based on the extracted features

A CNN is made up of three different kinds of layers that connect together. Pooling Layer, Full Connection (FC) Layer, and Convolutional Layer These layers are stacked up into a CNN structure. There are two crucial parameters, in addition to these three layers. Below are definitions for the activation function and dropout layer.

1. Convolutional Layer

The first layer used to extract features from the input image is this layer. The convolution process between the input image and a MxM filter of a particular size takes place in this layer. To determine the dot product (MxM) that exists between some of the filter's input images, slide the filter onto the image. The result is known as a utilitarian guide and gives data about edges and edge-like pictures. In order to learn additional input image features, the feature map is later moved to additional layers.

2. Pooling Layer

The majority of the time, a full layer continues after the convolution layer. The primary objective of this layer is to reduce the size of the convolution feature map, thereby lowering the cost of computation. By working independently on each functional map and reducing the connections between layers, this is accomplished. Depending on your approach, pooling operations can take many different forms. The function map is used to retrieve the largest element of max pooling. The average pool finds the average of the elements in a predetermined image section. The result of total pooling is the sum of the parts of a predetermined section. Typically, the pool layer serves as a link between the FC layer and the convolution layer.

3. Fully Connected Layer

Between the two layers of weight and flexion with neurons, a fully connected (FC) layer is used to connect neurons. The last few layers of the CNN architecture are made up by this layer, which typically comes before the output layer. The input image from the preceding layer is merged here before being transferred to the FC layer. The flattened vector is then passed through some FC layers, which are typically used for math function operations. At this point, the classification procedure begins.

4. Activation Functions

The activation function is, lastly, the CNN model's most crucial parameter. It is utilized for the estimation and learning of numerous intricate relationships between network variables. At the edge of the network, the model determines what is transferred and what is not. enhances the nonlinearity of your network. The Re LU, Soft Max, tanH, and Sigmoid functions are just a few of the activation functions that are frequently utilized. The purpose of each of these features is distinct. We recommend the sigmoid and Soft Max functions for multiple classifications in CNN models for binary classification. Soft Max is frequently utilized.

Proposed Work

The EMNIST dataset is expanded by including characters from other languages in the proposed system. The input image is first provided, then it is normalized and made into a grayscale image with the same resolution as the EMNIST dataset (28 x 28). CNNs are used to train categorizers that can outperform other machine learning algorithms and are trained using the EMNIST dataset. A trained Convolution Neural Network model is used to extract the features from the input image and produce the desired output. A character recognition architectural diagram is depicted in Figure. Pre-processing, feature extraction, normalization, and data classification of minimal Max Scaler-fitting images are just a few of the system's various modules.

PRE-PROCESSING: The input image undergoes pre-processing by being transformed into a grayscale image. Commonly referred to as RGB, a normal color image typically consists of three channels: a red channel, a green channel, and a blue channel. To remove any unnecessary noise from the image, the color image is then converted to a gray scale image with one black and white channel. When compared to images from image-trained convolutional neural networks, accurate predictions may be lost when input images are of varying sizes. As a result, the image is resized to match the EMNIST data set's resolution and placed as a blank image with a size of 28 x 28 pixels.

FFEATURE EXTRACTION: The process of converting input data into a set of features that accurately represent the input data is known as feature extraction. Dimensionality reduction is connected to feature extraction. The input data can be reduced to a smaller set of attributes (also known as attribute vectors) if it is too large to handle. The initial feature subset is assumed to be selected through feature selection. You can do what you want with this simplified representation rather than the full initial data because the selected function is expected to have information about the input data. The pixel values are taken in the form of a 1D array, ranging from 255 to 0 depending on the intensity of the pixels, after the image's size is changed.

IMAGE NORMALIZATION: The process of altering the pixel intensity values is known as normalization. It is also referred to as contrast stretching or histogram stretching. This input image's background pixels are removed, leaving only the characters in the running image for normalization. This can be accomplished by using any value, as long as the value of the background pixel is definitely lower than the value of the character's color pixel. The EMNIST data set's values are used as a basis for normalizing the images. After normalizing the image, the area where the letter "A" is written in this image has a pixel value greater than zero, and all other areas have a pixel value of zero.

CLASSIFICATION: A CNN is utilized as a grouping gadget to order penmanship from input pictures. A CNN comprises of an information layer, a result layer, and a few secret layers. A convolutional layer, a pooling layer, a fully connected layer, and a regularization layer make up the CNN's hidden layers. CNN is made up of three main parts: the pooling layer, the output layer, and the convolutional layer. Rectified Linear Unit, or ReLU, is a popular activation function in CNN. The convolutional layer computes the weighted dot product between the small regions connected in the input volume and the neuron's output that is connected to the local region of its input. A nonlinear down sampling form is the pooling layer. Maximum pooling, which divides the input image and outputs the maximum for each such subregion, is the most common method. A non-saturating activation function is used by ReLU. Without affecting the convolutional layer receptors, it enhances the network's determinant function and nonlinear properties as a whole. If the input is less than zero, a rectified linear device outputs 0 and, otherwise, a raw output. The following formula is used to determine its value. The SoftMax function is most frequently utilized in the final layer of neural network-network classifiers. f (x) = max (x, 0). Similar to the sigmoid function, the SoftMax function reduces each device's output between 0 and 1. Split each output, however, so that the total number of outputs equals 1. The categorical probability distribution is identical to the SoftMax function's output. As a result, the event probability distribution for "n" distinct events is calculated using the SoftMax function.

5. EXPERIMENTAL RESULTS

The CNN classifier receives the input after it has been scaled and normalized in the past. The CNN classifier can predict input characters because it is based on the EMNIST dataset. There were a variety of characters in the database we used for our experiments.

It can be divided equally into sets for training, testing, and validation. The system defines local geometry like endpoint cross-points. Statistics are used to retrieve the characteristics of the training set that have been extracted by the membership function. The project's outcomes are listed in the table below, along with an accuracy graph.

6. CONCLUSION

Back propagation algorithms are used to automatically design and adaptively learn the spatial layers of features in convolution neural networks, which are made up of multiple blocks like convolution layers, pool layers, and fully connected layers. I did. In recent years, handwriting recognition has been a challenging endeavor. However, computer vision's image recognition has significantly improved as a result of recent advances in the field of machine learning and the enormous amount of data generated every day. There are approximately 132,000 images of 47 people in the EMNIST dataset that can be identified as learning. The EMNIST datasets were trained using convolution neural networks for high accuracy. In order to predict characters, the input image is pre-processed, standardized, and provided with a categorizer.

7. REFERENCES

1. Rohan Vaidya , Darshan Trivedi , Sagar Satra.”Handwritten Character Recognition Using Deep-Learning” in Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018) IEEE Xplore Compliant
2. Wei Lu, ZhijianLi,Bingxue Shi . ”Handwritten Digits Recognition with Neural Networks and Fuzzy Logic” in IEEE International Conference on Neural Networks, 1995. Proceedings
3. B. V. S. Murthy. “Handwriting Recognition Using Supervised Neural Networks” in International Joint Conference on Neural Networks, 1999. IJCNN ’99.
4. J.Pradeep, E.Srinivasan, S.Himavathi. “Neural Network based Handwritten Character Recognition system without feature extraction” in International Conference on Computer, Communication and Electrical Technology ICCCET 2011
5. Chunpeng Wu, Wei Fan, Yuan He, Jun Sun, Satoshi Naoi.” Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network” in 2014 14th International Conference on Frontiers in Handwriting Recognition
6. Yu Weng1 &Chunlei Xia.” A New Deep Learning-Based Handwritten Character Recognition System on Mobile Computing Devices” at https://doi.org/10.1007/s11036-019- 01243-5 M
7. BatuhanBalci, Dan Saadati, Dan Shiferaw.” Handwritten Text Recognition using Deep Learning” Through research paper
8. Mori S (1992) Historical Review of OCR Research and Development. Proc IEEE 80(7):1029–1058

9. J.Pradeep, E.Srinivasan, S.Himavathi. ”Neural Network based Handwritten Character Recognition system without feature extraction” in International Conference on Computer, Communication and Electrical Technology ICCCET 2011.
10. Rahul R. Palekar , Sushant U. Parab , Dhrumil P. Parikh , Prof. Vijaya N. Kamble. ”Real Time License Plate Detection Using OpenCV and Tesseract” in International Conference on Communication and Signal Processing.
11. ”OpenCV” https://en.wikipedia.org/wiki/OpenCV through google
12. ”An open-source machine learning framework for everyone” https://www.tensorflow.org/,[Online] Available: https://www.tensorflow.org/.
13. ”NIST Special Database 19” https://www.nist.gov, [Online]. Available: https://www.nist.gov/srd/nist-special-database-19.
14. K.Fukushima, “Handwritten Alphanumeric Character Recognition by Neocognitron”, IEEE Trans. Neural Networks, V01.2, No.5, May, 1991.
15. ”A-Z Handwritten Alphabets in .csv format” https://www.kaggle.com/sachinpatel21/az handwritten-alphabets-in-csv-format
16. Md. Musfiqur Rahman Sazal; SujankumarBiswas;Md.Faijul Amin; Kazuyuki Murase .” Bangla handwritten character recognition using deep belief network” published in 2013 International Conference on Electrical Information and Communication Technology (EICT)
17. “Artificial Intelligence” Available: https://searchenterpriseai.techtarget.com/definition/AIArtificial-Intelligence.
18. G. Y. Chen, T. D. Bui and A. Krzyzak, “Contour-Based Handwritten Numeral Recognition Using Multiwavelets and Neural Networks,” Pattern Recognition, Vol. 36, No. 7, 2003, pp. 1597-1604. doi:10.1016/S0031-3203(02)00252-2