Grenze International Journal of Engineering and Technology, Jan Issue

An Intelligent Way to Recognize Digits using Convolutional Neural Networks (CNN) Algorithm

A.Virinchi Sai[1], M. Sai Dhanush[2], M.N. Akarshana[3], and Sandeep R[4]
[1-3]Methodist College of Engineering and Technology/Computer Science and Engineering, Hyderabad, India
Email: saidhanushccs@gmail.com, akarshanagoud2001@gmail.com
[4]Asst.Professor, Methodist College of Engineering and Technology/Computer Science and Engineering, Hyderabad, India
Email: rsandeep@methodist.edu.in

***Abstract*—The Handwritten digit recognition issue in pattern recognition systems is one of the more significant technical issues. Applications for digit recognition include data entry forms, bank check processing, postal mail sorting, and more. The capacity to create a practical algorithm that can recognize handwritten digits given by users via a scanner,tablet and other digital devices is at the core of the issue. The study's main goals were to create a machine learning algorithm and an optimization technique to improve the precision of handwritten digits recognition, as well as to analyze the performance of the proposed algorithm with test data set. This model offers a method for recognizing handwritten digits offline that is based on various machine learning techniques. There are several machine learning algorithms such as Support Vector Machine, Random Forest, Multilayer Perceptron, Convolutional Neural Network etc. In this project is aimed to use Convolutional Neural Network to complete the task. The MNIST dataset also used in this project. The main objective of this model is to ensure effective and reliable approaches for recognition of handwritten digits.**

***Index Terms*— CNN, Neural Network.**

I. INTRODUCTION

The sudden growth of new documents and multimedia news has created new challenges in pattern recognition and machine learning (Cecotti, 2016). Handwriting character recognition has become a common research area due to technological advances such as handwriting capture devices and powerful mobile computers. However, since handwriting focuses on the writer, building a highly reliable recognition system that recognizes any handwritten character input to an application can be challenging.

In this work, we deal with the recognition problem of handwritten numbers, i.e. the numbers 0-9. Handwritten number recognition is usually crucial in various practical applications such as management and economics (Niu & Suen, 2012). These industries require excellent recognition and the highest credibility. For example, unconstrained handwritten digit recognition has been applied to checks and hand-filled forms such as tax forms or postal codes on postcards with excellent results (Lauer, Suen, & Bloch, 2007). Constraint perception refers to the extent to which people believe that factors beyond their control limit their behaviour. In contrast, an unconstrained detection system can be divided into several parts: preprocessing, feature extraction, classification, evaluation, and verification.

Grenze ID: 01.GIJET.9.1.756
© Grenze Scientific Society, 2023

However, the detection accuracy of the tests is mostly around 95%. Since many classifiers cannot adequately process the original images or data, feature extraction is one of the preprocessing steps that seeks to abstract the relevant information and decrease the complexity of the data (Lauer, Suen, & Bloch, 2007).

A. Objective

The project's main objective on handwritten recognition is to calculate the accuracy of the digits written by the user and display the correct digits on the screen. The project aims to build an application program to recognize the digits written by the user and display them accurately.

II. Literature Survey

Numerous academics made their imprint on the digit recognition subject. Hanmandlu and Murthy built a character recognition system with varied priorities for particular aspects based on the accuracy of various features. In order to determine the character order in a handwritten script, a hidden Markov model utilising a recurrent neural network (RNN) has been utilised. This model was applied by Graves and Schmidhuber with 91 percent accuracy for the classification of handwritten Arabic words. Multilayer perceptron (MLP) was used by Pal and Singh to recognise handwritten English letters, and they were able to increase computation efficiency while achieving accuracy of up to 94 percent.

III. System Study

A. Input What is Convolution?

Convolution is an operation that transforms a function into another. For example, doing a convolution to transform the original function into a shape to get more information. Convolutions have long been used in image processing to blur, sharpen, and perform other operations on images, such as Image enhancement and embossing.

B. Convolution Process

Convolution is not a new concept. They are long being used in image and signal processing. However, convolution in machine learning is different from image processing.

Convolutional layers are the primary building blocks used in convolutional neural networks.

Convolution is a simple filter applied to an input that results in activation. Repeatedly applying the same filter to input will result in an activation map called an object map, which shows the position and strength of the detected input object, such as an image.

The innovation of the agglomeration neural network is the ability to automatically learn in parallel many specific filters for a training dataset under the constraints of a particular predictive model problem, such as classification. As a result, specific features are detected anywhere on the input image.

A Convolutional Neural Network, abbreviated as CNN, is a specialized neural network model designed to work with two-dimensional image data. However, it can be used with one- and three-dimensional data.

A convolutional layer is a layer that gives the network its name is the heart of the convolutional neural network. This class performs an operation called "convolution".

In the context of convolutional neural networks, convolution is a linear operation that involves multiplying a set of weights with the input, just like a traditional neural network. However, since this technique is designed for two-dimensional input, the multiplication is performed between an input data array and a two-dimensional array of weights, known as a filter or multiplication.

The dot product is an element-by-element multiplication between the input and the filter's filter size patch, which is then added together, always resulting in a single value. Here it results in a single value, and the operation is often referred to as the "dot product".

Using a filter less than input is intentional, as it allows the same filter (weight set) to be multiplied multiple times by the input array at different points in the input. Specifically, the filter is applied consistently to each filter-sized overlap or patch of the input data, left to right, top to bottom.

The systematic application of the same filter on an image is a powerful idea. For example, suppose a filter is designed to detect a particular type of feature in the input. In that case, applying that filter consistently across the entire input image will allow the filter to detect that feature at any time, anywhere in the image. This capability is often referred to as translation invariance, e.g. a general concern about whether the item is present rather than its location.

Convolution is an ordered procedure in which two sources of information are interleaved, and it is an operation that converts a function into something else. Convolutions are used for a long time in image processing to blur, sharpen images, and perform other operations. (e.g. edge enhancement and embossing) CNN's apply a model of local connectivity between neurons in adjacent layers

C. Convolution Layers

Convolutional neural networks use filters to detect features such as edges throughout the image.
There are four primary operations in CNN:
1)Convolution
2) Non-Linearity (ReLU)
3) Pooling
4) Classification
In Convolutional Neural Network, the first layer is the Convolutional Layer. These layers apply a convolutional operation to input, passing its result to the next layer. A convolution converts entire pixels in its receptive image field into a single value. For example, suppose you would apply convolution to an image. In this case, you would reduce the size of the image and pack all the field information into a single pixel. The final output of the convolution layer is a vector. We can use different complexes depending on the problem we need to solve and the type of feature we are looking at.

The 2D Convolution Layer

The most commonly used convolution type is the 2D convolutional layer, often abbreviated as 2D convolution. The filter or kernel in the collation layer "slides" on the 2D input data, performing element-by-element multiplication. As a result, it will aggregate the result into a single output pixel. The core will perform the same operation for each position it slides over, turning the 2D matrix into a different 2D matrix of entities.

The Dilated or Atrous Convolution

It will enlarge the window size without increasing the number of weights by adding null values into the convolution kernels. As a result, scaling or flickering combinations can be used in real-time applications and applications with lower processing power because less intensive RAM is required.

Separable Convolutions

There are two main types of separable convolution: spatial separable convolution and depth separable convolution. Spatial separable convolution mainly deals with the image's and kernel's structural dimensions: width and height. Compared to spatially separable convolutions, dimensional separable convolutions work with kernels that cannot be "computed" into two smaller kernels.

Transposed Convolutions

These types of convolutions are also known as deconvolutions or fractionally stridden convolutions. A transposed convolutional layer carries a regular convolution but reverts its spatial transformation. browser.

IV. SCREEN SHOTS

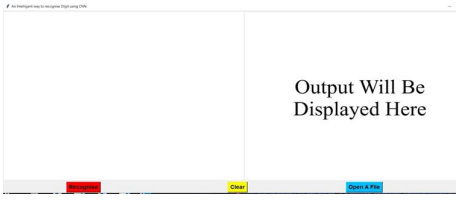

Figure 1. Main Menu

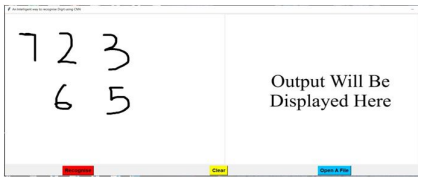

Figure 2. Handwritten digits on the screen using mouse

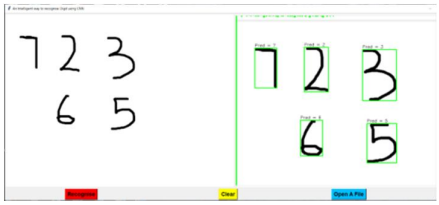

Figure 3. Output of the written digits

Figure 4. Output of the uploaded image

V. Conclusion

We have developed a handwritten digit recognition Python deep learning project with success. We created and trained a convolutional neural network that is exceptionally good at classifying images. Later, we create the GUI, which recognises numbers on images and allows us to draw a digit on the canvas, classify it, and display the findings.

Future Scope

This model has a scope of improving it's areas on recognizing texts, objects and even human faces with that we can information across the world.

References

[1] "Handwriting recognition": https : //en.wikipedia.org/wiki/Handwritingrecognition
[2] "What can a digit recognizer be used for?": https : //www.quora.com/W hat−can−a−digit−recognizer− be − used − for
[3] Ravikanti, Sandeep, and Gadasu Preeti. "Future's Smart Objects in IOT, Based on Big-Data and Cloud Computing Technologies." International Journal of Innovative Research in Computer and Communication Engineering 3.7 (2015): 6808-6817.
[4] Ravikanti, Sandeep, and Gadasu Preeti. "Evaluating the performance of reactive unicast routing protocols with OPNET simulator in MANETS under VOIP." International Journal of Innovative Research in Science, Engineering and Technology 4.7 (2015).
[5] Thirupathi, L., Rajesh, A., Sandeep, R. (2023). Internet of Lighting for Smart Cities. In: Tiwari, M., Ismail, Y., Verma, K., Garg, A.K. (eds) Optical and Wireless Technologies. OWT 2021. Lecture Notes in Electrical Engineering, vol 892. Springer, Singapore. https://doi.org/10.1007/978-981-19-1645-8_41
[6] Ravikanti, Sandeep, and Dheeraj Ganesh. "Smart Farming: A Techno Agriculture Advancement Powered by Machine Learning."
[7] Social Media: To Deal Crisis Circumstances Thirupathi, Lingala, and Ravikanti Sandeep. "Social Media: To Deal Crisis Circumstances." International Journal of Innovations & Advancement in Computer Science (IJIACS) 6.9 (2017).