

IMAGE CLASSIFICATION USING REGION WITH CONVOLUTIONAL NEURAL NETWORKS

Mr.G.Ravi¹, P.Revathi², Y.Sharanya reddy³, G.Vernika⁴

^{1*} assistant professor, Department of Electronics and Communication Engineering, St. Peter's Engineering College, Hyderabad, Telangana, India

^{2,3,4*} UG Student, Department of Electronics and Communication Engineering, St. Peter's Engineering College, Hyderabad, Telangana, India

ABSTRACT- In recent years, computer vision systems have placed a strong emphasis on the detection and categorization of the variety of materials that may be found in our immediate surroundings.

A complex procedure that has made it difficult to identify components in different photos using current developments in neural networks, which have the ability to train architectures to extract features for this difficult job. Modern Support Vector Machine (SVM) classifiers and Convolutional Neural Network (CNN) methods are used in this study to categorize the materials findings. on several commonly used material datasets gathered, a variety of CNN architecture are assessed to determine which is the best method to extract features in order to get exceptional outcomes for the animals, wood, plastic, metal.

KEYWORDS: Neural Networks, CNN, GPU. Deep Learning, CNN, Inception model, Faster RCNN, and Soft Max

INTRODUCTION

Our research on managing species in the environment will be impacted by having current information about various creatures. Manually identifying animals and their characteristics is still a costly, time-consuming operation. Therefore, we suggest that such identification and Deep learning neural network techniques may be used to do classification with the highest degree of accuracy.

Data mining, a method that combines machine learning. The emphasis is on removing relevant information from huge data sets and translating it into an understandable format for later use. Data pre-processing, classification, and categorization are its key uses. networks capable of training through unsupervised learning are used in machine learning.

Based on the unstructured data. Considering that deep learning is one of the most effective methods for

There are just a few techniques that can efficiently and automatically learn a variety of features from big data sets overcoming the difficulties of feature extraction.

A recognition system typically extracts characteristics from an input image when it is given. The Network employs these in conjunction with the training data to train itself and classifies the data into relevant groups. The learned information from training is applied to forecast test data based on characteristics and classifies them accordingly.

AIM AND OBJECTIVE

Deep CNN will be used in the study to identify different animal species. Manually recognizing them might be challenging due to the diversity of the species .A new approach employing CNN dubbed Lightweight Machine Convolution Network for Animal Recognition is proposed in this stage of the project to address limitations caused by data mining. The computational cost will be low, and the accuracy will be good.

NETWORK FUNCTIONING

There are two parts to how the network operates. Brief definitions of the theory relating to CNN are provided in Section 2.1. The characteristics in Section 2.2, along with the function of layers.

BACKGROUND INFORMATION

A long time ago, McCulloch and Pitts originally suggested a computational model of a neural network, as seen in [3].

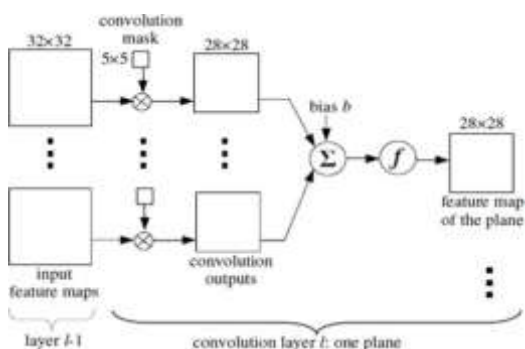
The layers that make up a neural network are interconnected to create the network, which is composed of several layers. It is possible to conceptualise a feed-forward neural network (FFNN) in terms of neural activity and the degree of connectivity between each pair of neurons [4].

We refer to the layer that lies between these two as the concealed layer. With the goal of minimizing error between the output and input, learning happens through weight modification.

Multi-Layer Perceptron (MLP) derivatives called Convolutional Neural Networks (CNN) are based on biological inspiration. Due to their limited nature in input space, these filters are better equipped to take advantage of the strong spatially local correlation found in natural pictures [5]. For processing two-dimensional (2-D) images convolution neural networks are used [6]. In this project a CNN architecture that is specified in [7] is employed. Three distinct types of layers a convolution layer a subsampling layer and an output layer make up the network.

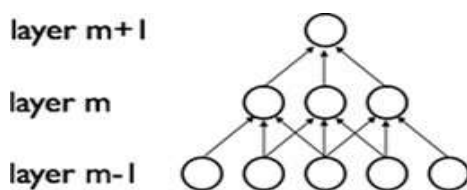
CNN (CONVOLUTION NETWORK)

Deep artificial neural networks called convolutional neural networks are typically used to categorized pictures, group them according to similarities, and recognize animals in scenarios. CNN classifies the animal picture and then determines the species of the animal in this module based on the attributes that the user has extracted from the animal image. The CNN design shown in Fig. 2 has three layers: a fully linked layer, a pooling layer, and two convolution layers that are sequentially applied. Convolution layer employs a number of filters that are going to be applied to the input picture to produce various activation characteristics. Similar to convolution, pooling layer.



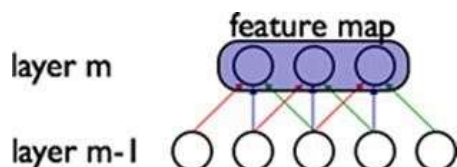
WORKING OF CNN ALGORITHM

The algorithm's operation is briefly described in this section. The explanation in [7] is comprehensive. The network receives an input of a 2D picture. The network is made up of intermediary layers, or "hidden layers," an input layer that accepts images, and an output layer from which we derive the learnt output. Multiple convolutional and subsampling layers make up the network, as was already mentioned. When the layers are combined, the incoming picture data is an approximation. In order to take advantage of spatially local correlation, CNNs enforce the local connection arrangement between the neurons in neighbouring layers [8]. Consider a scenario where only a tiny percentage of layer (m-1) neurons are connected to layer m neurons.



Layers are connected in a graphically shown flow

Between layers [4]



Layers' graphic flow demonstrating weight distribution 4

Each sparse filter is repeated over the full visual field using the CNN algorithm. The feature maps that result from these units share a weight vector and bias. Figure (2b) shows three concealed units on the same feature map. The weights of the same Color are shared and must thus be the same.

The gradient of shared weights is the accumulation of the gradients of the shared parameters. This replication enables the recognition of features independent of their location in the visual field. Additionally, weight sharing enables a reduction in the number of free learning parameters. CNN can more reliably generally their results on visual difficulties because to this control.

CNN also uses a non-linear down-sampling technique called max-pooling. Using this method, the input image is divided into non-overlapping rectangles. Each sub region's output represents its greatest value.

EXISTING SYSTEM

The safety of humans and animals will be enhanced by the use of animal detection to assist minimize animal-vehicle accidents. The vehicle will be warned by audio and visual indicators before any large animals cross the road. This also aids in preventing pests from damaging agricultural crops. Edge-based matching, skeleton extraction, and object matching are some of the object detection techniques that are being studied in this research, as well as techniques for detecting items as being those of animals. The best animal detection technique is selected following a survey, and its efficacy is assessed..

The proposed system has a low rate of false positives and false negatives.

The backdrop assumption used in previous background subtraction research was a static background.

A foundation frame with only backdrop is used by several models for background segmentation.

PROPOSED SYSTEM

The proposed technique aims to perform animal classification by combining two deep learning algorithms. demonstrating the value of deep learning technology in classification issues while learning neural network frameworks like CNN and Faster RCNN capturing images .CNN and Faster RCNN algorithm training on sliced data sets and classification using these techniques





SPECIFICATIONS FOR HARDWARE

Processor: Any processor with a frequency greater than 500 MHz.

Ram: 4 GB

250 GB hard drive, regular keyboard and mouse, and a webcam are used as input devices.

High-definition monitor, the output device.

SPECIFICATIONS FOR SOFTWARE

Windows 7 or a newer operating system

programming: libraries for Python 3.6

MACHINE LEARNING IS NECESSARY

Because of our capacity for thought, analysis, and problem-solving, humans are now the most intellectual and developed species on the planet. However, artificial intelligence (AI) is still in its infancy and hasn't fully overtaken human intellect in many domains. So the next query is, why is machine learning necessary? Making Judgment with efficiency and scalability on the basis of data is the best justification for doing this.

In recent years, businesses have made significant investments in cutting-edge Using technologies like artificial intelligence (AI), machine learning, and deep learning, we can extract the crucial information from data, carry out a range of useful tasks, and address problems.

Dataflow and differentiable programming may be done across a variety of applications using the free and open source software or Flow. It is a library of symbolic mathematics that is utilized by neural networks and other machine learning techniques. Google Brain's second-generation technology is known as Tensor Flow. Tensor Flow may operate on a range of CPUs and GPUs in contrast to the conventional approach, which only supports a single device. It also supports the SYCL and CUDA extensions for GPU processing that is used generally. Linux, MacOS Windows, and Tensor Flow are all 64-bit operating systems.. If we want to automate the process, we might refer to it as machine-driven decisions. Alternatives include using these data-driven judgement.

TENSOR FLOW

A complete open source platform for machine learning is called Tensor Flow. Its extensive, adaptable ecosystem of tools, libraries, and community resources enables developers to quickly create and deploy ML-powered apps while also enabling academics to advance the state-of-the-art in machine learning.

A free and open source program called Tensor Flow may be used in a variety of applications for dataflow and differentiable programming. It is a symbolic mathematics library that is utilized by machine learning algorithms like neural networks. The second-generation Google Brain technology is called Tensor Flow. Tensor Flow allows for the use of a multitude of CPUs and GPUs, as opposed to the old method, which only supports one device. For general-purpose GPU computation, it also supports the SYCL and CUDA extensions. Linux, macOS, Windows, and mobile operating systems like Android all provide 64-bit versions of Tensor Flow.

A variety of applications can utilize the free and open source software or Flow for dataflow and differentiable programming.

SOFTWARE BUILDING BLOCKS

- ❖ FLASK FRAME WORK
- ❖ My SQL
- ❖ HTML& CSS,JAVA SCRIPT
- ❖ PYTHON

FLASKFRAME WORK

Python was used to create the Flask framework for web applications. It was created by Armin Ronacher, the president of the Pocco organization a global organisation of Python fans. Flask is created by combining the WSGI tools of Werkzeug with the Jinja2 template engine. Projects for Pocco, all of them.

MYSQL

A benchmark for SQL, a condensed form of the structured query language that is often called S-Q-L or See-Quell, is the Image Net Large Scale Visual Recognition Challenge.

The main objective of this database language is to maintain relational database management systems for the data. Data experts use a specialist tool to process structured data—data that is stored in tables—for processing. It is also designed to be used for stream processing by RDSMS.

HTML, CSS, JAVA SCRIPT

Hyper Text Markup Language is what HTML stands for.

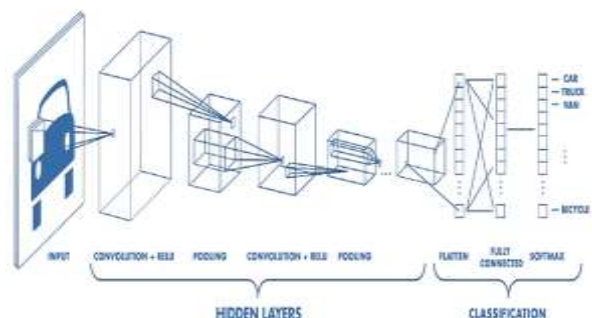
The fundamental framework of a website may be built using this very straightforward language by developers. HTML is at the heart of even the most complicated websites. A recent Stack Overflow study also found that it is the second-most popular programming language among developers.

CSS

If HTML acts as the foundation of a website, CSS gives those foundational elements form and enhancement. In order to indicate how various components of a webpage should seem to users, style sheets are employed, such as CSS. In other words, it's a technique to spruce up the HTML content you've previously created and add some extra formatting.

JAVA SCRIPT

The three front end languages mentioned in this article are all quite simple, but JavaScript is the most sophisticated since it builds on both HTML and CSS. Think of it like this while attempting to compare the languages: While HTML establishes the foundation of a website, CSS adds style to that foundation, and JavaScript takes all of that work and makes it interactive and more functionally complicated.



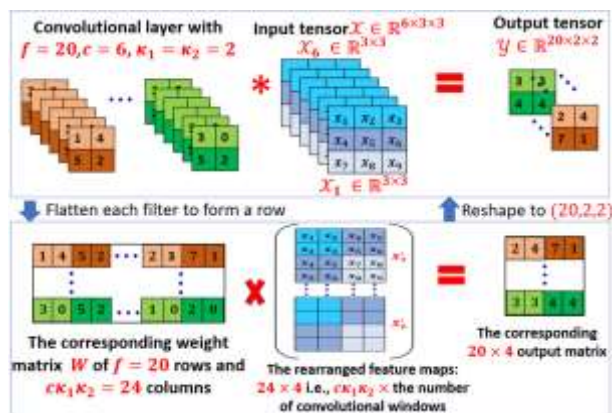
CNN ARCHITECTURE MOTION

SETTING UP THE DATABASE

The input is shown as the actual image. Since data information, not color information, is what the network needs, the pictures are changed from color to gray scale. The photos are also 32x32 resized. Pyramid reduction is performed to reduce the size of the photographs from data sets to 32x32. The image pyramid is a type of data structure created to provide effective scaled convolution with little picture representation. It is made up of a series of duplicates of an original image in which the sample density and resolution are gradually reduced To enable effective scaled convolution with little picture representation, a specific type of data structure known as a "image pyramid" was created. It consists of successively decreasing sample densities and resolution copies of the original image.

Images themselves serve as the input. Since the network is more concerned with data information than color information, the pictures are changed to gray scale. The pictures have also been scaled to 32x32. Pyramid reduction is used to reduce the pictures from data sets to 32x32 pixels since they are bigger. In order to enable efficient scaled convolution with reduced image representation, the picture pyramid data structure was developed. It comprises of a series of replicas of the original image, each with decreasing sample density and resolution [9].

CONVOLUTIONAL ANALYSIS



Three channels should be included in the filter when the picture is RGB. This is due to the fact that an RGB picture comprises 3 colour channels, necessitating the employment of 3-channel filters for the computations. Here, between the picture portion and the filter, each corresponding channel is where the computation takes place. All of the outputs from each channel's calculations are added to produce the result in question. The feature map lacks a third dimension as a result.

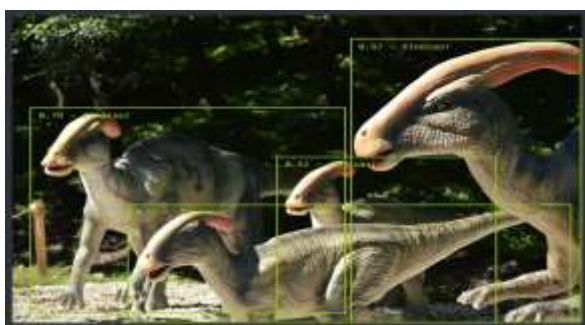
$$(0*0 + 3*1 + 0*1) + (2*0 + 0*1 + 1*0) + (0*1 + 1*0 + 3*0) = 3$$

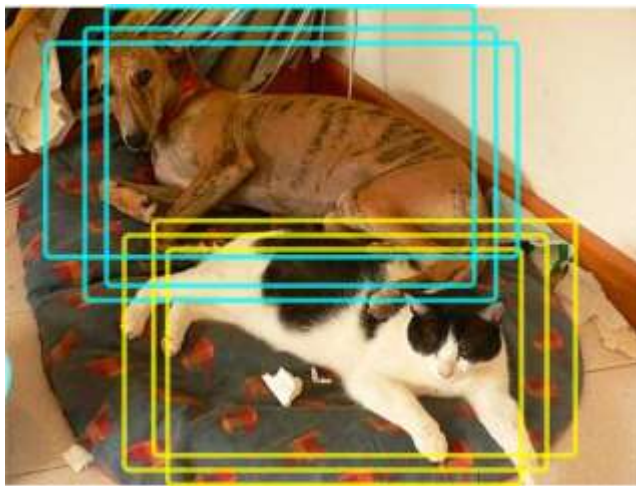
TRAINING DATA SET WITH IMAGE NET

A benchmark in object categorization and detection, the Image Net Large Scale Visual Recognition Challenge has millions of images and hundreds of item classes. that person Google Net provided the best outcome to yet, reducing classification error to 0.06656 and raising the mean average precision of item recognition to 0.439329. Its network is implemented with over 30 layers. Convolutional neural networks performed almost as well as people on the Image Net tests. With the aid of Google Net since its start, we collect animal photographs as training datasets for our system because to the enormous number of animal species.

RESULT

The classification accuracy results from several standard datasets using the CNN method are shown in this section. The outcomes are displayed using the classification accuracy expressed as a percentage both the test data and the within-train data separately. There is also an MSE (Mean Squared Error) graph displayed in addition to the classification accuracy % figures. The simplest and most used quality statistic is MSE. It is the squared difference between the trained approximation and the initial approximation.





CONCLUSION

Using the most recent deep learning neural network frameworks, we proposed in this study to automatically categorize photos of animals in order to save time and money. We assessed the neural networks' propensity for very accurate picture classification. The findings emphasize the significance of deep learning neural network technology in picture categorization. As a result, these network frameworks may be utilized to extract information about wildlife while also reducing human Labour and expense, which will help to preserve and manage the zoological system.

FUTURE SCOPE

Networks only use Color pictures in the system that was just suggested. In order to improve the grey scale photos that the system receives as input, we intend to work on them in the future. In order for the anticipated test image to be extremely accurate, we would like to extend the system. Only single labels may be classified using the approach above. Due to this, we want to strive towards multi-label categorization in the future.

ACKNOWLEDGEMENT

We want to express our appreciation to assistant professor Mr.G.Ravi sir from the dept of ECE

We wish to convey our appreciation to Dr. K. Prasanna Kumar the head of the ECE Department

REFERENCES

[1] SONALI JADHAV, AAKANKSHA CHAUDHARI, PRACHI CHANDANSHIVE, and PARINITA BADRE. Deep Learning for Automatic Animal Recognition. March 2022 issue of the International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE), Volume 6, Issue 3. ISSN (Online): 2320-9801 ISSN (Print): 2320-9798.

[2] RISHI KUMAR, CHRIS ROWEN. Using Convolution Neural Networks for Image Recognition. IP Group, Cadence. 2021

[3] Y, WEI L, ZHANG F, LI H (2015) Deep Convolution neural networks for Hyper spectral image classification. Journal of Sensors 2019.

[4] ZISSERMAN, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2018)

[5] REN, S., HE, K., GIRSHICK, R., SUN, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: NIPS. pp. 91–99 (2017)