# Infrared Image Detection Using Convolutional Neural Network With Cloud Computing

**Nibedita Panda, Abhyarthana Bisoyi**
Department of Instrumentation and Engineering,
College of Engineering & Technology, Bhubaneswar, Odisha, India

**Abstract:**

Convolutional neural network is an active field of research and it provide a large platform for computer vision. It recognizes the images and can classify them as needed to provide better information. Hence an update can be made for the ways used in defense system for better surveillance at the border. Here the network is trained for airplane and tank. Any image of airplane or tank if appears then the network can detect that image and give information accordingly. By combining this CNN with cloud computing, work becomes easy and much less time consuming without human intervene. Even at night or at bad or foggy weather an eye can be kept for security as infrared camera is used which provide vision even in this adverse weather condition. This combination of technology if used can provide a better option to alert us from any upcoming urgency.

**Keywords:**

Soft-max pooling, flattening, back propagation algorithm, deep learning, VGG-16.

## 1. INTRODUCTION

These years a great success can be witnessed in different fields by using convolutional neural network. Due to the high demand in the requirement of accuracy of visual recognition[5], the requirement for higher computational resources have increased a lot so to achieve the high demand accuracy and performance the cloud resources are made available which are cost effective, provide on demand service, we can pay as per the use, and is reliable[6]. The back-propagation algorithm used here is very accurate in infrared image recognition [4] but found to be slow for the task of discovering the airplanes hence here we combinedly used it with cloud resources [2] to provide better speed along the high accuracy [1]. Cloud provide various pre trained CNN network which we can use as a base and with some more further training to that pretrained network our work can easily be done, the base pre trained network we used here is VGG-16 where 16 means number of [3] hidden layers in convolutional neural network.

The problem statement has been discussed in section2, the flow chart is explained in section3, the training procedure, transfer learning and results are presented in section 4, 5 & 6 respectively. The paper has been concluded in section 7.

## 2. PROBLEM STATEMENT

Deep learning-based approach is advocated for automatic recognition of civilian targets in thermal infrared images. High variability of target signature and low contrast ratio of targets to background makes the task of target recognition in infrared images challenging, demanding robust adaptable methods capable of capturing these variations. As opposed to the traditional shallow learning approaches which rely on hand engineered feature extraction, deep learning-based approaches use environmental knowledge to learn and extract the features automatically. We present convolutional neural network (CNN) based deep learning framework for automatic recognition of civilian targets in infrared images.

## 3. PROPOSED WORK

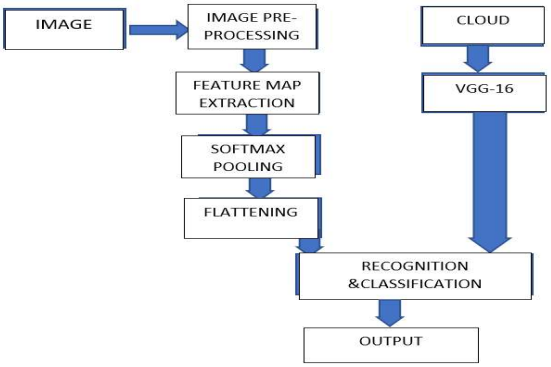



Fig. 1 Block Diagram of the system

Figure 1 shows the basic block diagram of the proposed work. The aim of the paper is to detect the match between the desired image among a whole scale of images. The detail process is described as follows:

Image pre-processing: The aim of pre-processing is an improvement of image data that suppresses unwanted distortions and enhances some features of the images for further processing. It is a process for image at lowest level of abstraction. Here the input and the output both are intensity images. There are 4 categories of image pre-processing according to the size of the pixel neighbourhood which are Geometric transformation, Pixel brightness transformation, Image restoration & Use of local neighbourhood of the processed pixel.

Feature map extraction: A feature map can roughly be created from the inputs given to the single feature kernel in a convolutional network. The output we see maybe the result of several neurons with the same weight or a single neuron run several times over the entire input.

Soft-max pooling: This is the new kind of pooling layer for faster and sharper convergence. This layer produce output by applying soft-max function as activation to the net input coming from the previous layer. The soft-max function squashes the output of each unit to be between 0 and 1, just like a sigmoid function. But it also divides each output such that the total sum of the outputs is equal to 1.

Flattening: Here we flatten our pooled feature map into a column. The purpose of doing this is to insert data into an artificial neural network later on.

Prediction: Neural networks are trained so that they can learn from examples only and after their learning, they will be able to predict hidden and strongly non-linear relations, even when there is a significant noise in the training set. But the error of prediction cannot be generally estimated.

## 4. TRAINING PROCEDURE

At first, all the data sets were collected and stored in the AWS cloud. By using python in build library these data transformed the infrared image into normal grayscale image and then the convolutional neural network was provided with or trained with these images. As the network is very complex and there are so many hidden layers present hence computation process becomes complex so we run the program hosting in AWS cloud [2] and the image we provided is detected.

The data present are divided into 3 groups:

Training (80%)

Validation (15%)

Test (5%)

Back propagation algorithm

Back propagation algorithm is a supervised learning algorithm which uses the technique of correcting the error with each iteration [1]. It compares the deviation in the actual output obtained with the desired output needed and fed the error back to the network by adjusting the weights each time until the network is trained as desired. The network learns from its error and every time gives better result than before thus training the neural network.

The training of a neural network by back propagation takes place in three stages. Feedforward of the input pattern. Calculation and Back propagation of the associated error.

Adjustments of the weights.

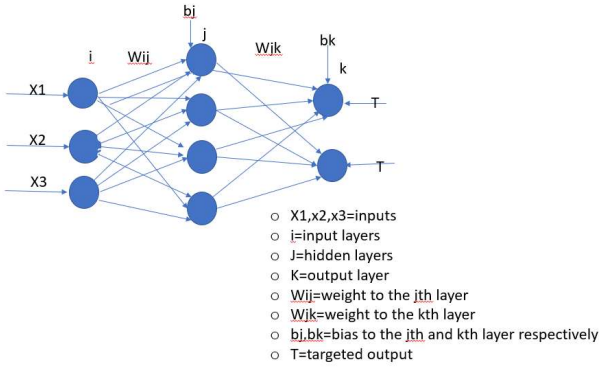



Fig. 2 Training Procedure with weights

$O_i = I_i \quad (I_i = x_1, x_2, x_3 \,\&\, O_i = output\ of\ the\ i\ layer)$

$I_j = \sum(W_{ij} * O_i) + b_j$

$O_j = {}^{1}/_{(1+e^{-I_j})}$

$I_k = \sum(W_{jk} * O_j) + b_k$

$O_k = {}^{1}/_{(1+e^{-I_k})}$

$E_k = O_j(1 - O_j)(T_j - O_j) \quad (T_j = targeted\ attribute\ \&\ E_K = error\ calculated\ in\ the\ k\ layer.$

$E_j = O_i(1 - O_i)\sum_k E_k * W_{jk}$

$$\Delta W_{ij} = l * E_j * O_j \qquad (l = known\ learning\ rate)$$
$$\Delta W_{jk} = l * E_k * O_j$$
$$W_{ijnew} = W_{ij} + \Delta W_{ij}$$
$$\Delta b_j = l.E_j$$
$$b_{jnew} = b_j + \Delta b_j$$

## 5. TRANSFER LEARNING

In this technique the network was trained (here VGG-16) [3] as a base and update it with further training and get result easily. Here 16 is the number of hidden layers present in the convolutional neural network. These networks are already having some source knowledge about a specific task and as per our requirement we add some other features by programming it and can get desired output. For example, if there is a source domain and a source learning task and a target domain and a target learning task, it helps to improve the target prediction using the source knowledge.

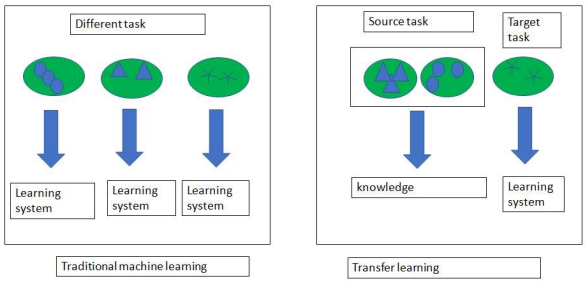
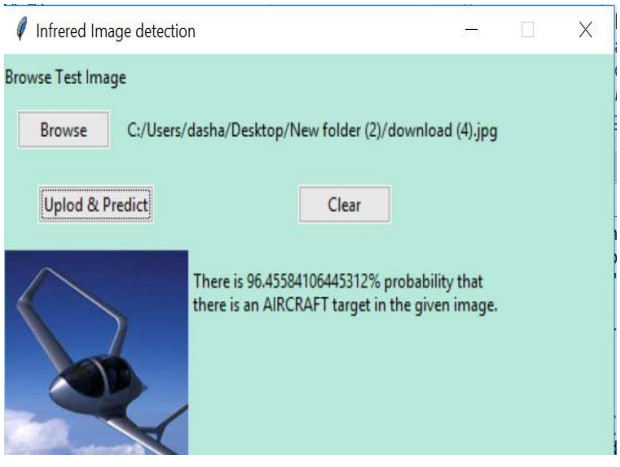


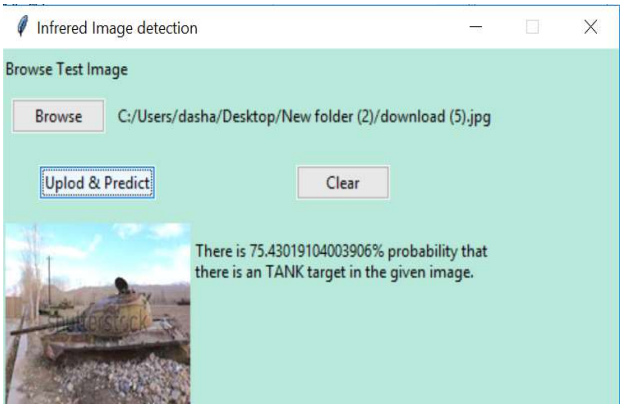

Fig. 3 The Basic Learning Process

## 6. RESULTS & DISCUSSIONS



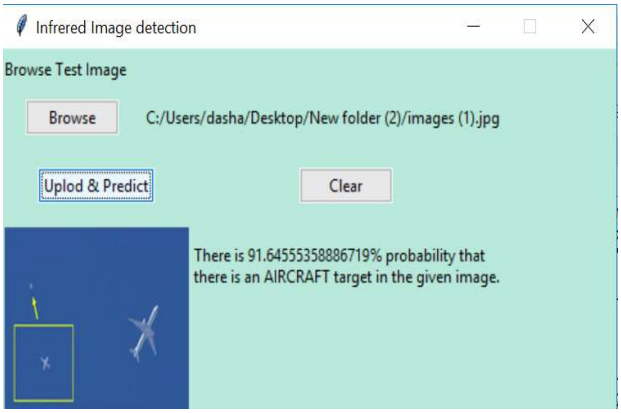

(a)

Here in (a), the device matching is 96.45%, as the total image consists of aeroplane similar to the given input image.



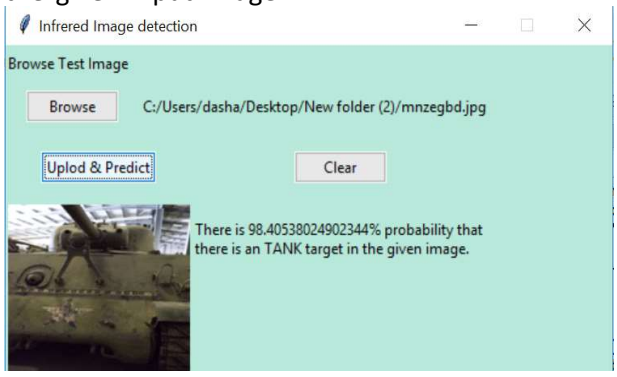

(b)

Here in (b), the device matching is 75.43%, as the total image consists of tanker and some surrounding picture. So, the similarity is less with respect to the given input image.



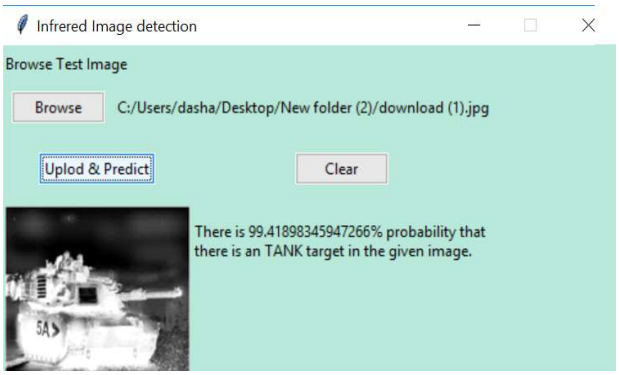

(c)

Here in (c), the device matching is 91.64%, as the total image consists of three aeroplanes similar to the given input image.




(d)

Here in (d), the device matching is 98.4%, as the total image consists of tanker similar to the given input image.




(e)

Fig. 4 The similarity results of various pictures

Here in (e), the device matching is 99.41%, as the total image consists of tanker similar to the given input image.

For training the CNN Google ML (machine learning) dataset & UCI ML library are used as source of the above thermal and true pictures. The cloud service used for running deep learning program is "aws (amazon web service) ec2 t2.big" and for algorithm is sagemaker. The loss function used is logarithimic with adam optimiser. The learning rate is 0.025. The CNN is trained for both normal (ground true) images as well as thermal (IR) images using Keras image augmentation. These figures show that we detected the images with a very high percentage of precision. The images provided here are the images of a broken airplane, damaged tank, a distant aircraft, partly showing tank, and a infrared image of a tank. In every case the network detected the object and classified it with a high percentage of accuracy and a very high speed. The VGG-16 network is fed almost 1GB of images and the iteration it took to train the network is 20. We had taken 4types of images for training i.e. aircraft, aeroplane, airplane and a tank. The infrared images taken (shown in Fig 5) is also detected clearly within fraction of seconds and high accuracy. This network works with the RBG as well as with the infrared image both.

## 7. CONCLUSION

It can be concluded that by using convolutional neural network in the technique of image detection, security system can be improved. The back-propagation algorithm gives very high accuracy in detection of normal, infrared as well as distant and broken images and the speed issue related to this back-propagation algorithm can be compensated by using this algorithm with cloud computing technique. The whole new combination of network formed become cost effective and reliable to use, very fast and even easy to design due to the presence of the pre trained basic CNN network in the cloud. This image detection using machine learning is a very emerging field for research. This paper can be extended by training the CNN to detect multiple targets using infrared imaginary.

## REFERENCES

[1] P. Y. Liu and H. X. Li. Efficient learning algorithms for three-layer regular feedforward fuzzy neural networks. IEEE Trans. Neural Network, 15(3): 545-558, 2004.

[2] C.C. Rao, M. Leelarani, and Y.R. Kumar, Cloud: Computing Services and Deployment Models. International Journal of Engineering and Computer Science, 2(12), pp. 3389-3390, 2013.

[3] M. Rezaee ,Y. Zhang ,R. Mishra , F. Tong, H. Tong, "Using a VGG-16 Network for Individual Tree Species Detection with an Object-Based Approach", 10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS), 10.1109/PRRS.2018.8486395, 2018.

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in ICLR, 2015.

[5] Tianmei Guo, Jiwen Dong, Henjian Li Yunxing Gao, "Simple Convolutional Neural Network on Image Classification", IEEE 2nd International Conference on Big Data, 978-1-5090-3619-6/17.

[6] A.M Chandrashekhar, Shashikumar, "Cloud Computing service and Deployment Models", IJRASET, Volume 5 Issue VI, 2017.