AVOID TRESPASSING IN PLANTATION FARMS USING CONVOLUTION NEURAL NETWORK

Ronak Agarwal
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
ronak.agarwal2016@vitstudent.ac.in

Shubham Gupta
School of Computer Science & Engineering
Vellore Institute of Technology
Vellore, India
shubhamgupta.2016@vitstudent.ac.in

Parmeet Singh
School of Computer Science & Engineering
Vellore Institute of Technology
Vellore, India
parmeet.singh2016@vitstudent.ac.in

Sharmilla A
School of Computer Science and Engineering
Vellore Institute of Technology
Vellore, India
asharmila@vit.ac.in

- ***Abstract*---We aimed to deploy Convolution Neural Networks to overcome the issue of Object detection at Plantation farms. The system is designed here will ensure detection of objects and informing the concerned authorities, so that the necessary action can be taken. This system will reduce the efforts of human and make plantation field more secured in order to save the farm field from especially animals that cause large destruction of plantation farms at nights. There are lots of issues that taken into account for developing an efficient object detection algorithm.**

Keyword: *Convolutional Neural Networks, Max pooling, Dropouts, Flattening, Regression, KNN*

I. INTRODUCTION

Animal detection is a very important and emerging area due to a large number of real life applications. Various animal detection methods and warning systems are used for indicating the presence of animals on the plantation farms or roads. Applications which are very important in real life are preventing plantation farms from animals that cause large destruction of planation fields. Knowingall these applications can be narrowed down to two areas namely detection and identification of animals.

Researchers have tried to find out whether the presence of animal in the scene or image will affect the power spectral of the image or not which can be defined as the amplitude of the signal in the frequency domain. The power spectrum can be constructed by transforming images from s (spatial) domain to frequency domain with the help of the transformation function like Fourier transform. Work carried out in [9] shows that this approach is not suitable if a person wants quick result or wants to detect the animals very quickly as this approach takes more time.

For monitoring or observing the locomotive behavior of animals and their interaction with the surroundings, we are deploying Convolution Neural Networks to Overcome the issue of animals in plantation fields. The system will be designed in such a way that ensures the Detection of objects and classify with categories of animals such that necessary action can be taken without harming the animals. Simultaneously it reduce the Human's effort and make more secured the plantation field.

There are lots of issues that need to be taken into account for developing an efficient animal detection algorithm. Animals come from nowhere so you can't predict their presence and also the speed of the animals can't be monitored or detected. There is the lighting problem also, wherein a sudden change of lighting effect can affect the effectiveness in detecting the presence of animal intrusion. Also each animal has its own characteristics and behavior with the surroundings which leads to a problem in identification of correct animals. Apart from that moving background, such as trees or leaves caused by wind might be regarded as foreground image and some stationary animal which remain still for a long time can be mistakenly interpreted as background image by the algorithms.

II. LITERATURE REVIEW

Convolutional Networks (ConvNets) are a biologically-inspired trainable architecture that can learn invariant features. While ConvNets have been successfully deployed in many commercial applications from OCR to video surveillance, they require large amounts of labelled training samples [1]. ConvNets are biologically inspired multi-stage architectures that automatically learn hierarchies of invariant features. While many popular vision approaches use handcrafted features such as HOG or SIFT, ConvNets learn features at every level from data that are tuned to the task at hand[2]. Unlike many popular vision approaches that are hand designed, ConvNets can automatically learn a unique set of features optimized for a given task. We augmented the traditional ConvNet architecture by learning multi-stage features and by using pooling [3].

We propose a method that uses a multiscale convolutional network trained from raw pixels to extract dense feature vectors that encode regions of multiple sizes centered on each pixel. The method alleviates the need for engineered features, and produces a powerful representation that captures texture, shape, and contextual information. We report results using multiple post processing methods to produce the final labeling [4].A system utilizing such a neural network can greatly benefit a smart city by providing real time localized visibility data across all highways and roads by utilizing a dense network of traffic and security cameras that exist in most developed urban areas [5].

Towards this aim, a colorful image of the similar content as the grayscale image is taken, as an input source image by means of different image retrieval techniques. Then, the best matching source pixel is determined using luminance matching technique, for each pixel of the target grayscale image. Once a best matching source pixel is found, its chromaticity values are assigned to the target pixel while the original luminance value of the target pixel is retained [6].Convolutional neural networks, which are specifically designed to deal with the variability of 2D shapes, are shown to outperform all other techniques [7].Rectified activation units (rectifiers) are essential for state-of-the-art neural networks. In this work, they study rectifier neural networks for image classification from two aspects. First, they propose a Parametric Rectified Linear Unit (PReLU) that generalizes the traditional rectified unit.

PReLU improves model fitting with nearly zero extra computational cost and little overfitting risk. Second, they derive a robust initialization method that particularly considers the rectifier nonlinearities. This method enables us to train extremely deep rectified models directly from scratch and to investigate deeper or wider network architectures[8].Learning to store information over extendedtime intervals via recurrent back propagation takes a very long time, mostly due to Insufficient, decaying error back. They briefly review Hoch Reiter's 1991 analysis of this problem, then address it by introducing a novel, efficient, gradient-based method called "Long Short-Term Memory"(LSTM) [9].To address this, they further introduce a dedicated comic style CNN, which is trained for classifying comic images and photos. This new network is effective in capturing various comic styles and thus helps to produce better comic stylization results.

Even with a grayscale style image, Gates's method can still produce colored output, which is not desirable for comics. They develop a modified optimization framework such that a grayscale image is guaranteed to be synthesized. To avoid converging to poor local minima, they further initialize the output image using grayscale version of the content image [10].Rendering the semantic content of an image in different styles is a difficult image processing task.

Arguably, a major limiting factor for previous approaches has been the lack of image representations that explicitly represent semantic information and, thus, allow to separate image content from style. Here we use image representations derived from Convolutional Neural Networks optimized for object recognition, which make high level image information explicit.

III. DATASET DESCRIPTION

The problem of automatically identifying objects in Plantation farms is difficult because of the near infinite number of permutations of objects, positions, lighting and so on. It's a really hard problem. This is a well-studied problem in computer vision and more recently an important demonstration of the capability of deep learning. A standard computer vision and deep learning dataset for this problem was developed by the Canadian Institute for Advanced Research (CIFAR).

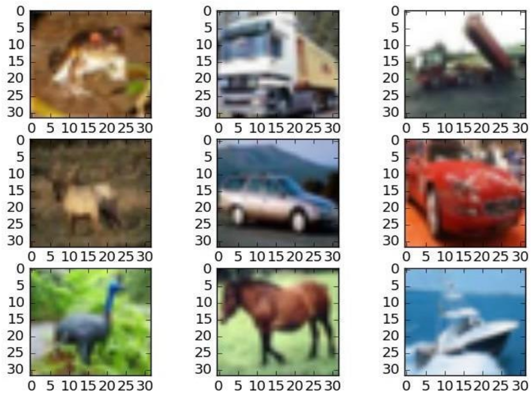

Figure1: Dataset Description.

The CIFAR-10 dataset consists of 60,000 photos divided into 10 classes (hence the name CIFAR-10). Classes include common objects such as airplanes, automobiles, birds, cats and so on. The dataset is split in a standard way, where 50,000 images are used for training a model and the remaining 10,000 for evaluating its performance.

The photos are in color with red, green and blue components, but are small measuring 32 by 32 pixel squares. State of the art results are achieved using very large Convolutional Neural networks. You can learn about state of their results on CIFAR-10 on Rodrigo Berenson's webpage. Model performance is reported in classification accuracy, with very good performance above 90% with human performance on the problem at 94% and state-of-the-art results at 96% at the time of writing.

IV. METHODOLOGY

We will use a structure with two convolutional layers followed by max pooling and a flattening out of the network to fully connected layers to make predictions. Our baseline network structure can be summarized as follows:

[1] Convolutional input layer, 32 feature maps with a size of 3×3, a rectifier activation function and a weight constraint of max norm set to 3.

[2] Dropout set to 20%.

[3] Convolutional layer, 32 feature maps with a size of 3×3, a rectifier activation function and a weight constraint of max norm set to 3.

[4] Max Pool layer with size 2×2.

[5] Flatten layer.

[6] Fully connected layer with 512 units and a rectifier activation function.

[7] Dropout set to 50%.Fully connected output layer with 10 units and a softmax activation function.

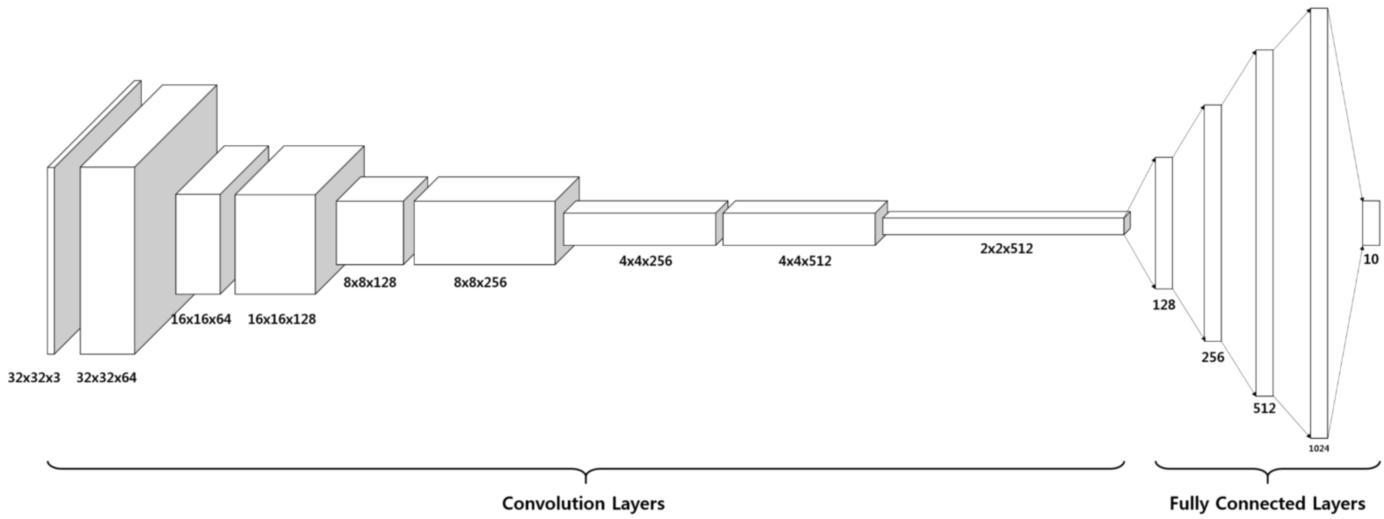

Figure2. Classification Model.

V. ASSOCIATED PARAMETERS

- Sigmoid function

$$y = \frac{1}{1+e^{-x}}$$

- Euclidean formula

$$d(p,q) = d(q,p)$$
$$= \sqrt{(q_1 - p_1)^2 + (q_3 - p_3)^2 + \cdots + (q_n - p_n)^2}$$
$$= \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$$

* Convolution function:

$$\boxed{f(x) * g(x) = \int_{-\infty}^{\infty} f(\tau) \cdot g(x-\tau)\, d\tau}$$

VI. CNN ALGORITHM

CNNs are suitable for any type of data where the order of features is important, or put in more technical terms, for any spatially ordered data. Convolution Neural Network is a type of feed forward neural network that is generally used for Image Recognition and Image Classification tasks. A CNN accepts arrays of pixel values as input to the network. The hidden layer consists of several different layers which carry out feature extraction. There is a fully connected layer that recognizes the objects in the image. Convolution operation forms the core of every convolution neural network. There are 4 layers in a CNN. These are Convolution layer, ReLU layer, Pooling layer and Fully Connected Layer.

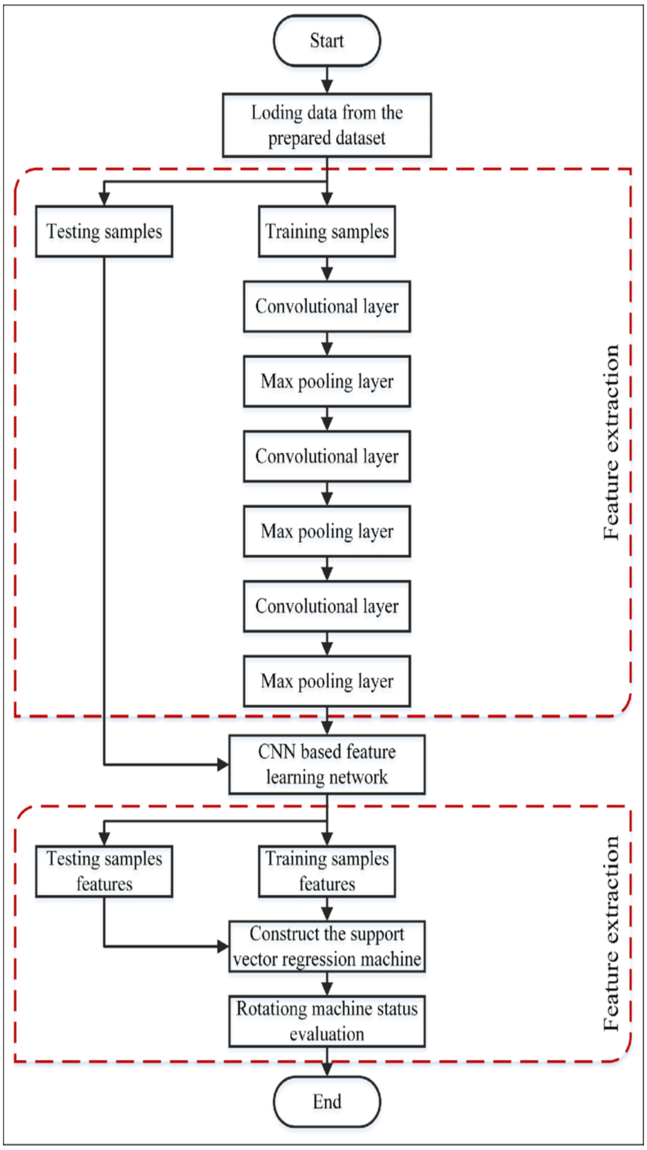

Figure 3: CNN flowchart

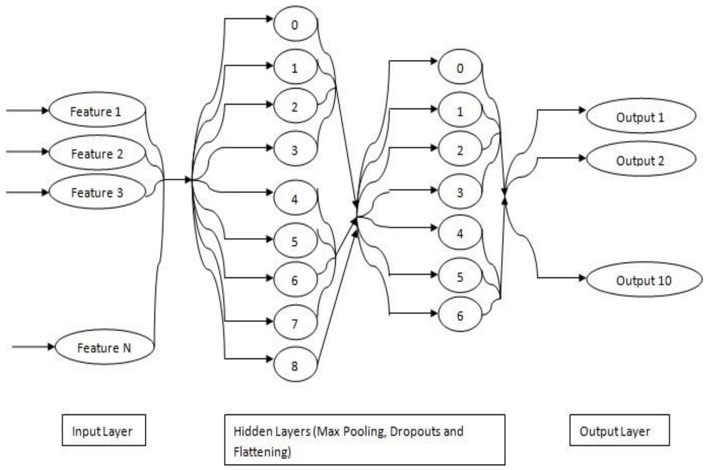

Fig4. CNN Process Flow.

VII. LOGISTIC REGRESSION

Logistic regression is generally used where the dependent variable is Binary or Dichotomous. That means the dependent variable can take only two possible values such as "Yes or No", "Default or No Default", "Living or Dead", "Responder or Non Responder", "Yes or No" etc. Independent factors or variables can be categorical or numerical variables.

The algorithm of Maximum Likelihood Estimation (MLE) determines the regression coefficient for the model that accurately predicts the probability of the binary dependent variable. The algorithm stops when the convergence criterion is met or maximum number of iterations are reached. Since the probability of any event lies between 0 and 1 (or 0% to 100%), when we plot the probability of dependent variable by independent factors, it will demonstrate an 'S' shape curve.

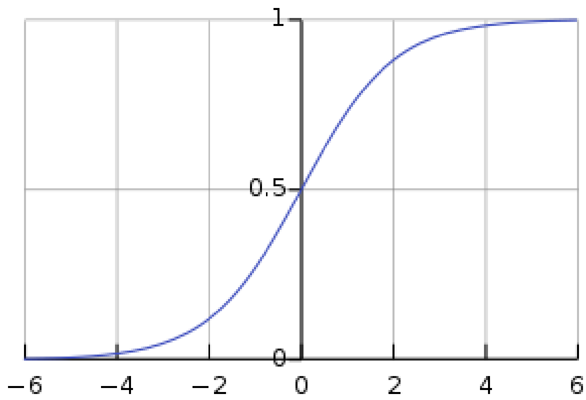

Figure 5. S shape Curve

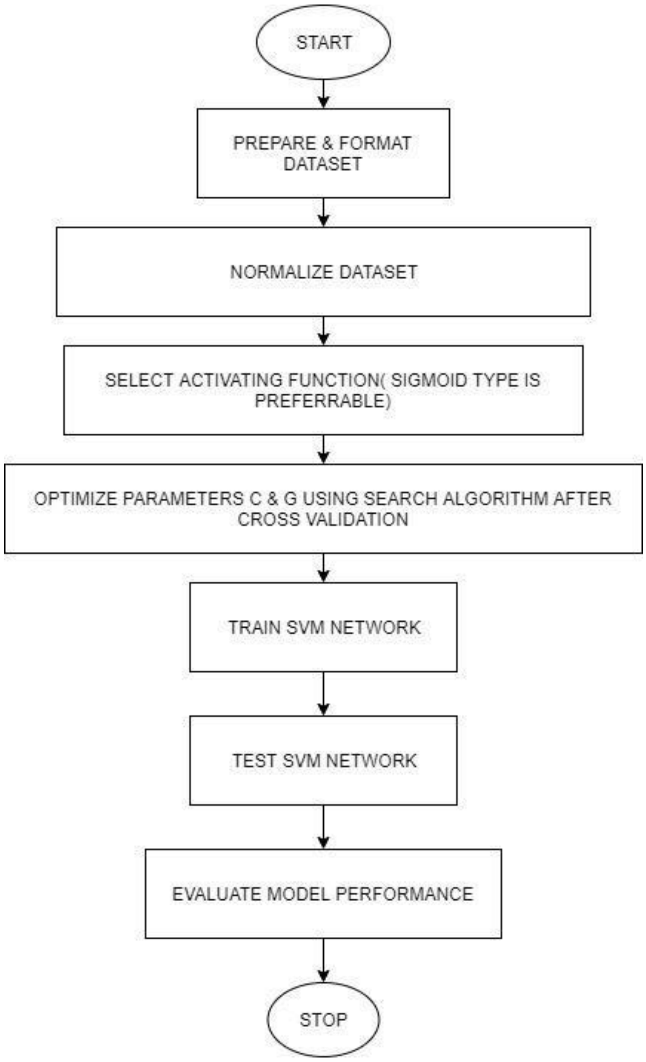

Figure 6: Logistic Flow chart

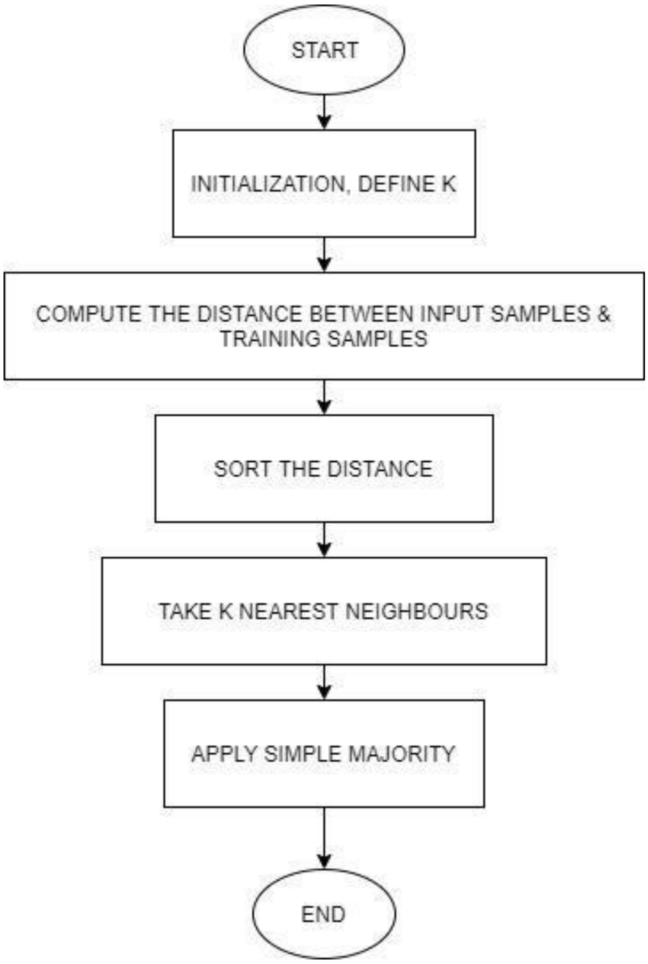

Figure7. KNN Flow chart

VIII. KNN ALGORITHM

KNN can be used for both classification and regression predictive problems. However, it is more widely used in classification problems in the industry. We can implement a KNN model by following the below steps:

1. Load the data
2. Initialise the value of k
3. For getting the predicted class, iterate from 1 to total number of training data points
    - Calculate the distance between test data and each row of training data. Here we will use Euclidean distance as our distance metric since it's the most popular method. The other metrics that can be used are Chebyshev, cosine, etc.
    - Sort the calculated distances in ascending order based on distance values
    - Get top k rows from the sorted array
    - Get the most frequent class of these rows
    - Return the predicted class

The Object detection in the plantation field is done using PCA(principal component analysis) recognize objects by deciding representative features of objects in the model image, extracting feature vectors from objects in an image and measuring the distance between them and object representation. Given frequent recognition problems associated with the use of point-to-point distance approach, this study adopted **the K-nearest neighbor technique** (class-to-class) in which a group of object models of the same class is used as recognition unit for the images inputted on a continual input image. However, we propose the object recognition technique new PCA analysis method that discriminates an object in database even in the case that the variation of illumination in training images exists. Object recognition algorithm proposed here represents more enhanced recognition rate to change of illumination than existing methods.

IX. RESULTS.

S.No	ALGORITHM	NO OF EPOCHS	TIME ELAPSED	ACCURACY
1.	**Logistic Regression**	**30**	**1674.2 min**	**46.33%**
2.	**K-Nearest Regression**	**30**	**1263.5 min**	**54.33%**
3.	**Convolutional Neural Network**	**30**	**575.63 min**	**82.54%**

X. CONCLUSION

Comparative study of the all the algorithms is done to detect objects in the field. The A logistic regression is analysis of data yields a accuracy of around 46.33% due to multi-class classification. The accuracy is improved with the help of the KNN algorithm which yields 54.33% accuracy which is way lesser than the modern CNN which increases the efficiency to the level of 82.54%. This is due to the self-learning and feature extraction done by the Convolution Neural Network on its own. The proposed algorithm is faster, reliable and more accurate.

X1. ACKNOWLEDGEMENT

This research was partially supported by one of the professors' at VIT **Prof Sharmila A.** We thank her for providing us the insight and expertise that greatly assisted the research of this paper. We are also grateful to Vellore Institute of Technology for giving us the Opportunity and making it possible by providing us labs and support whenever needed.

X11. REFERENCES.

1. Farabet- Convolutional networks and applications in vision- Published in Proceedings of IEEE International Symposium on Circuits and Systems DOI:10.1109/ISCAS.2010.5537907
2. Dominik Scherer, Andreas Muller, and Sven Behnke- Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition-20th International Conference on Artificial Neural Networks (ICANN), Thessaloniki, Greece, September 2010
3. Pierre Sermanet and Yann LeCun - Traffic Sign Recognition with Multi-Scale Convolutional Networks - Courant Institute of Mathematical Sciences, New York University
4. Xavier Glorot,Antoine Bordes and Yoshua Bengio - Deep Sparse Rectifier Neural Networks - DIRO, Universit´e de Montr´eal Montr´eal, QC, Canada
5. Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton- ImageNet Classification with Deep Convolutional Neural Networks - University of Toronto
6. Pierre Sermanet, Soumith Chintala, Yann LeCun - Convolutional neural networks applied to house numbers digit classification - Published in Proceedings of the 21st International Conference Pattern Recognition (ICPR2012)
7. Clement Farabet, Camille Couprie, Laurent Najman, Yann LeCun - Learning Hierarchical Features for Scene Labeling - Courant Institute of Mathematical Sciences, New York University (New York, NY 10003, USA)
8. Saluky Saluky and Fetty Lubis - Moving Image Interpretation Models to Support City Analysis - 2018 International Conference on ICT for Smart Society (ICISS) DOI: 10.1109/ICTSS.2018.8550012
9. Divyang Patel and Shanker Parmer - Image retrieval based automatic grayscale image colorization - Mumbai 2013
10. Razvan Pascanu, Tomas Mikolov and Yoshua Bengio - On the difficulty of training recurrent neural networks - Universit´e de Montr´eal, 2920, chemin de la Tour, Montr´eal, Qu´ebec, Canada 2012
11. Dominik Scherer, Andreas Müller and Sven Behnke - Evaluation of pooling operations in convolutional architectures for object recognition - University of Bonn, Institute of Computer Science VI, Bonn, Germany 2010
12. Yann LeCun, Leon Botton, Yoshua Bengio and Patrick Haffner - Gradient Based Learning

Applied to Document Recognition - AT&T Shanon Lab 2012.
13. Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun - Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification - Microsoft Research 2015
14. Sepp Hochreiter and Jurgen Schmidhuber - LONG SHORT-TERM MEMORY - Fakultat fur Informatik Technische Universitat Munchen 80290 Munchen, Germany 1997
15. Razvan Pascanu, Tomas Mikolov and Yoshua Bengio - On the difficulty of training recurrent neural networks - Université de Montréal, Montréal, Québec, Canada
16. Zezhou Cheng, Qingxiong Yang and Bin Sheng - Deep Colorization - 2010
17. Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun - Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification - Microsoft Research 2015
18. Yang Chen, Yu-Kun Lai, Yong-Jin Liu - Transforming photos to comics using convolutional neural networks -Published in IEEE International Conference on Image Processing (ICIP) 2017 DOI:10.1109/ICIP.2017.8296634
19. Gantugs Atarsaikhan, Brian Kenji Iwana, Seiichi Uchida - Contained Neural Style Transfer for Decorated Logo Generation - Published in 13th IAPR International Workshop on Document Analysis Systems (DAS) 2018 DOI:10.1109/DAS.2018.78
20. Ayumu Shinya, Nguyen Duc Tung and Tomohiro Harada - Object-Specific Style Transfer Based on Feature Map Selection Using CNNs - DOI: 10.1109/NICOInt.2017.39 Conference: 2017 Nicograph International (NicoInt)
21. Leon A. Gatys, Alexander S. Ecker and Matthias Bethge - Image Style Transfer Using Convolutional Neural Networks - Centre for Integrative Neuroscience, University of Tubingen, Germany
22. Haozhi Huang, Hao Wang, Wenhan Luo, Lin Ma, Wenhao Jiang, Xiaolong Zhu, Zhifeng Li, Wei Liu - Real-Time Neural Style.
23. Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun - Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification - Microsoft Research 2015
24. Sepp Hochreiter and Jurgen Schmidhuber - LONG SHORT-TERM MEMORY - Fakultat fur Informatik Technische Universitat Munchen 80290 Munchen, Germany 1997
25. Razvan Pascanu, Tomas Mikolov and Yoshua Bengio - On the difficulty of training recurrent neural networks - Université de Montréal, Montréal, Québec, Canada
26. Zezhou Cheng, Qingxiong Yang and Bin Sheng - Deep Colorization - 2010
27. Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun - Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification - Microsoft Research 2015
28. Yang Chen, Yu-Kun Lai, Yong-Jin Liu - Transforming photos to comics using convolutional neural networks -Published in IEEE International Conference on Image Processing (ICIP) 2017 DOI:10.1109/ICIP.2017.8296634
29. Gantugs Atarsaikhan, Brian Kenji Iwana, Seiichi Uchida - Contained Neural Style Transfer for Decorated Logo Generation - Published in 13th IAPR International Workshop on Document Analysis Systems (DAS) 2018 DOI:10.1109/DAS.2018.78
30. Ayumu Shinya, Nguyen Duc Tung and Tomohiro Harada - Object-Specific Style Transfer Based on Feature Map Selection Using CNNs - DOI: 10.1109/NICOInt.2017.39 Conference: 2017