International Journal of Latest Trends in Engineering and Technology
Special Issue April-2018, pp. 125-128
e-ISSN:2278-621X

SURVEY ON DEEP CONVOLUTIONAL NEURAL NETWORKS FOR BACKGROUND IMAGE DETECTION

Akula Monika[1], Rimo Mallick[2], C. Jerin Mahibha[3]

Abstract— Computer vision and object identification are two critical fields that assume real part in image recognition and detection. image discovery and detection has been utilized as a part of the considerable number of fields we know. It incorporates the techniques for getting, handling and breaking down the datasets that are given. Neural systems and datasets are the two critical variables for deep learning. As deep learning is a propelled condition of-workmanship which helps in accomplishing insight in computer vision with the assistance of creating models and calculations. It utilizes learning calculation, such as SVM(Support Vector Machines), ANN(Artificial Neural Network) and CNN(Convolution Neural Network).The essential advances associated with image discovery and recognition is pre-processing where the information picture is standardized, further includes extraction where imperative data is separated from the image and arrangement which utilizes learning calculation algorithm.
Keywords—Convolutional Neural Network, Deep Learning, Image Detection, Background Subtraction, Feature Extraction

1. INTRODUCTION

Deep learning is a learning method that processes target classification. It requires less human interventions and provides better accuracy and faster results. The deep learning system automatically learns the features and representations that can be applied to object recognition. Deep learning techniques are widely used in artificial intelligence. Using automated systems for discovering hidden features, it has achieved improved performance and efficiency. The most popular deep learning architecture is convolutional neural networks (CNNs). The input images and respective annotations are provided, and a CNN model is designed to learn and generate predictive data representations. These representations are used for target classification of testing image. Unsupervised learning is also sometimes applied to neural networks for data representation learning. The advantage of using unsupervised feature learning is that it does not involve human annotations. Deep learning, which learns feature representations and pattern recognition, takes advantage of large-scale high dimensional image data to discover hidden structures for better image detection. These days, deep learning is the major technique among the best solutions in image detection. It holds great potential for the field. In this paper, we emphasis on deep learning in image detection, which conceals several topics, such as pre-processing, feature extractions followed by training and testing the systems. Finally, we discuss the efficiency of the results produced by various deep leaning systems. This survey intent to help other investigators to catch a hint of the state-of-the-art methods in the ground of image detection.

The rest of the paper is organized as follows. Section 2 deals with detection of objects in computer vision. Section 3 involves various background subtraction algorithms. Section 4 involves the concept of deep learning and further ventures regarding them. The Section 5 deals with Convolutional Neural Network which helps in computational model. Finally we conclude the work.

2. OBJECT DETECTION

This technology is related to image processing. This is used to track semantic objects. This technology is also implemented in computer vision and in face recognition. This uses the objects feature to track them

In [1] the focus is on designing discriminative features and training powerful object detectors for judging whether there exists an object in each candidate region. When these detectors are used in real surveillance scene where the view changes rapidly then the performance and efficiency drops to overcome this here this paper uses an adaptive method to track the objects.

In [2] Object detection is developed from the single object recognition to the multi-object recognition. The meaning of the first is just from an image to identify a single object, it can be said that it is a problem of classification, and the meaning of the later is not only can identify all the objects in an image, including the exact location of the objects. It uses COCO(Common Objects In Context) which was introduced by Microsoft for image recognition and segmentation datasets.

[1] UG Student,Department of CSE, Meenakshi Sundararajan Engineering College, Chennai, India
[2] UG Student, Department of CSE, Meenakshi Sundararajan Engineering College, Chennai, India
[3] Assistant Professor, Department of CSE, Meenakshi Sundararajan Engineering College, Chennai, India

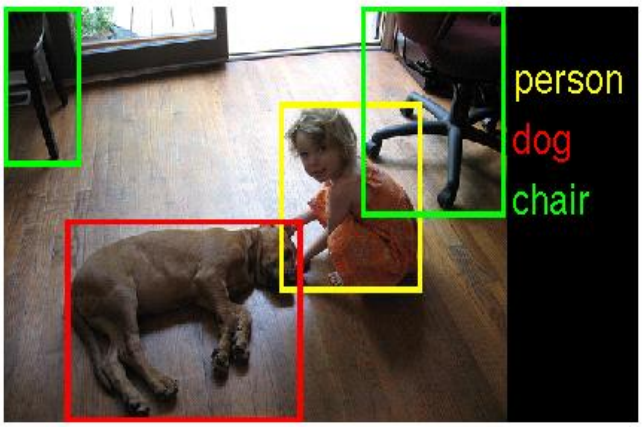

Fig 1.Object Identification and detection

In [3] the object detection is bit complex as it does not need to detect non stationary background objects like falling leaves or rain it needs to eliminate these objects and needs to detect only the moving objects in outdoor environments. The problems encountered in background subtraction are merging of object, distortion of object shape, missing object and classification of background as foreground

3. BACKGROUND SUBTRACTION

As the name suggest background subtraction is a technology used in the fields of computer vision and image processing which eliminates the background and detects only the foreground

In [3] the aim is to create a less-complex background subtraction algorithm in video application to detect moving object for outdoor environment by using MoG (Mixture of Gaussian) background subtraction method. This is done by evaluating each parameter in the MoG approach. Mixture of Gaussian (MoG) do not maintain a buffer for background estimation but updated a single background based on input frame. It maintains a single background model that is updated with each new video frame that makes this technique have minimal memory requirements.

- SOBS

The Self-Organizing Background Subtraction (SOBS) algorithm implements an approach to moving object detection based on the neural background model automatically generated by a self-organizing method, without prior knowledge about the involved patterns.

In [5] a variant of SOBS algorithm is used which is spatially coherent SOBS(SC-SOBS) multi modal background model. A background model is initially constructed and maintained in the algorithm. It is based on the idea of building an image sequence neural background model by learning in a self-organizing manner image sequence variations, seen as trajectories of pixels in time. Therefore, the neural background model well adapts to scene changes, capturing the most persisting features of the image sequence.

4. DEEP LEARNING

It is a subset field of machine learning. It performs unsupervised learning with unstructured data. Deep learning is just a field in artificial intelligence where the training data sets are collected altogether and then various machine learning algorithms are performed on these sets of data.

In [2] the deep learning is done by two factors neural network and datasets as the quality of number of dataset will affect the accuracy of neural network output the commonly used datasets in general computer vision fields is ImageNet, PASCAL VOC and COCO which was introduced by Microsoft. The input dataset to the deep learning is VOC dataset which is small, compared to the ImageNet dataset, which is very suitable for researchers to test network programs.

In [7] the paper uses the deep learning to show great potential in the field of remote sensing classification. A CNN-based spatial feature extraction method was proposed for the HSI classification. The extracted Deep Learning spatial features are more robust and effective than the traditional handcrafted ones.

In [9] and in [10] the active deep learning method is proposed which combines both deep learning and active learning it is mainly used for exploiting the property where the merit of a deep model lies in its ability to learn a discriminating set of features for a given task

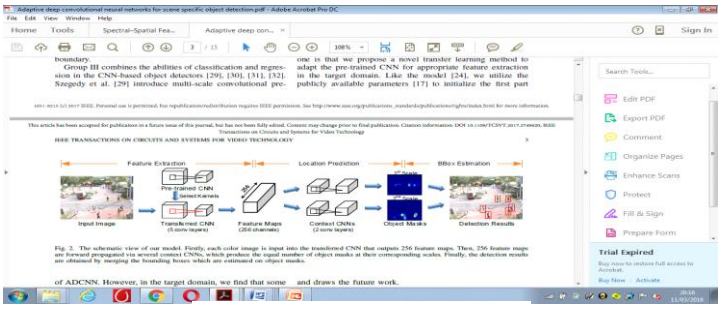

Fig.2. Feature Extraction Steps

4.1, Convolutional Neural Network

Convolutional Neural Network is mainly used for computer vision imagery. CNNs use a variation of multilayer perceptrons designed to require minimal preprocessing. It is a type of Artificial Neural Network where ANN is a computational model based on the structure and functions of biological neural networks

4.2 Deep Convolutional Neural Network

In [4] here, we investigate the application of supervised CNN, which is one of the deep models. It is challenging to apply deep learning to HSI since its data structure is complex and the number of training samples is limited. Three deep feature extraction architectures based on a CNN are proposed to extract the spectral, spatial, and spectral–spatial features of HSI. The designed 3-D CNN can extract the spectral–spatial features effectively, which leads to better classification performance. here training of samples in done using virtual datasets this tackles problem of limited samples and avoids regularization of CNN. 2-D CNN is used for spatial feature while 1-D and 3-D CNN is used for spectral classification and spatial-spectral classification. The design of proper deep CNN models is the first important issue to overcome this the paper proposed a 3-D deep CNN has been demonstrated to provide excellent classification performance under the condition of limited training sample

4.3 Text Cnn

in [8] In this work, a new system for scene text detection was proposed a novel Text-Attentional Convolutional Neural Network (Text-CNN) that particularly focuses on extracting text-related regions and features from the image components. We develop a new learning mechanism to train the Text-CNN with multi-level and rich supervised information, including text region mask, character label, and binary text/non text information. The rich supervision information enables the Text-CNN with a strong capability for discriminating ambiguous texts, and also increases its robustness against complicated background components.

4.4 R- Cnn

In [2] Regions with Convolutional Neural Network is One of the most noteworthy points of this paper is that the CNN is applied to the candidate box to extract the feature vector, and the second is to propose a way to effectively train large CNNs. SPP-Nets is an improvement compared to R-CNN which is faster. SPP-Net proposed a spatial pyramid pooling (SPP) layer that removes restrictions on network fixed size. Fast R-CNN improves the shortcomings of both R-CNN and SPP-Net and the improvements were the following: higher detection quality (mAP) than R-CNN and SPP-Net; It writes the loss function of multiple tasks together to achieve single-level training process; in the training can update all the layers; do not need to store features in the disk. Fast R-CNN can improve the speed of training deeper neural networks

Table 1. Comparative Analysis

CNN Types	Parameter		
	Speed	Performance	Dataset used
R-CNN	Average	Average	ILSVRC
SPP-NET	26-64 times faster than R-CNN	Better compared to R-CNN	PASCAL VOC
Fast R-CNN	10 times faster than SPP-NET	Accuracy rate increases compared to SPP-NET	VOC 2007
Faster R-CNN	Fastest	Greater precision near to 80 %	VOC 2007

5. CONCLUSION

In the survey, we have reviewed different papers that uses CNN to analyse image data. Since the power and efficiency of a system depends on how well it has been processed and trained, each method discussed in this paper gives different results based on the type of techniques adapted by it. CNN has proved to produced high accuracy and improved efficiency. The efficiency of any system can be improved by adjusting the number of training samples, the number of epochs and iterations used, and the number of hidden layers modelled in the CNN. It can also be combined with other deep learning architectures to create a highly robust and powerful system.

6. REFERENCES

[1] Xinyi Zhou,Wei Gong,WenLong Fu "Application Of Deep Learning In Object Detection" Computer and information Science(ICIS),2017 IEEE/ACIS 16th International Conference

[2] Xudong Li,Mao ye "Adaptive Deep Convolutional Neural Networks For Scene-Specific Object Detection" IEEE Transactions on Circuits and Systems for Video Technology (Volume: PP, Issue: 99)

[3] Shahrizat Shaik Mohamed, Nooritawati Md Tahir, Ramli Adnan" Background Modelling and Background Subtraction Performance for Object Detection" 2010 6th International Colloquium on Signal Processing & Its Applications (CSPA)

[4] Yushi Chen" Deep feature extraction and classification of hyperspectral images based on convolutional networks" IEEE Transactions on Geo Science and Remote Sensing (Volume 54 No.10, October 2016).

[5] Lucia Maddalena . Alfredo Petrosino "Extracting a Background Image by a Multi-modal Scene Background Model" 2016 23rd International Conference on Pattern Recognition (ICPR) Cancún Center, Cancún, México, December 4-8, 2016.

[6] Yingfeng Cai, Hai Wang, Zhengyang Zheng, And Xiaoqiang Sun" Scene-Adaptive Vehicle Detection Algorithm Based on a Composite Deep Structure"IEEE Access September 26, 2017

[7] Wenzhi Zhao and Shihong Du "Spectral–Spatial Feature Extraction for Hyperspectral Image Classification: A Dimension Reduction and Deep Learning Approach" IEEE Transactions On Geoscience And Remotesensing.

[8] Tong He, Weilin Huang, Yu Qiao, and Jian Yao "Text-Attentional Convolutional Neural Network for Scene Text Detection" IEEE Transactions on Image Processing (Volume: 25, Issue: 6, June 2016)

[9] Hiranmayi Ranganathan, Hemanth Venkateswara, Shayok Chakraborty, Sethuraman Panchanathan"Deep Active Learning For Image Classification" 2017 IEEE International Conference on Image Processing (ICIP)

[10] Sebastian Stabinger; Antonio Rodríguez-Sánchez "Evaluation of Deep Learning on an Abstract Image Classification Dataset" 2017 IEEE International Conference on Computer Vision Workshops (ICCVW)