As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 7(11): 104-108, 2021

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0711018
Available online at: http://www.ijmtst.com/vol7issue11.html

Improved Object Detection in Video Surveillance Using Deep Convolutional Neural Network Learning

Dr. G. Kiruthiga[1] | Dr. N. Yuvaraj [2]

[1]Department of Computer Science and Engineering, IES College of Engineering, Thrissur, Kerala, India
[2]Deputy Managerof Training and Research, ICT Academy, Chennai, India

To Cite this Article
Dr. G. Kiruthiga and Dr. N. Yuvaraj. Improved Object Detection in Video Surveillance Using Deep Convolutional Neural Network Learning. *International Journal for Modern Trends in Science and Technology* 2021, 7, pp. 104-108. https://doi.org/10.46501/IJMTST0711018.

Article Info
Received: 05 October 2021; Accepted: 07 November 2021; Published: 11 November 2021

ABSTRACT

In this paper, a convolutional neural network is used to improve the functionality of object detection using probabilistic neural network in the analysis of surveillance images. We will next start to examine the fundamental principles and some of the underlying theories of perceptual and neural networks that are often missed. This helps you understand why, in many areas of usage, deep learning has increased. The processing of surveillance photos is obviously one of the areas most influenced by this exponential progress, particularly in the identification and detection of images. The simulation result shows that the CNN-PNN obtained improved achievements in simulation for the optimal detection of objects in video streaming or video surveillance systems. The results show that the proposed method achieves higher rate of accuracy, f-measure and reduced percentage error than other methods.

KEYWORDS: Deep Learning, Object detection, video surveillance, CNN

INTRODUCTION

Object identification aims mainly at identifying and locating one or more efficient targets for silent images or videos. In such fields, it offers extensive potential applications, such as the avoidance of accidents [1], warnings of harmful items in factories, military control of limited regions and advanced interaction between humans and computers [2, 3]. Since multi-target detection scenarios are typically complicated and varied in the real world, it is a difficult challenge to balance the ratio between precision and computation cost.

Traditionally, the detection procedure is based on manually deriving functional models using the common features of classical gray-scale methods [4] [5]. These classical extraction feature models can only identify low-level information such as contour and texture information and are restricted by their inadequate overall performance in recognising many targets in a complex image [6]. However, the detection of objects based on features of deep learning, which is more familiar [7].

In addition to extracting detailed texture features from convolutional networks, deep learning convolutional neural network (CNN) models also allow for higher level information from the convolution layer after the level [8]. After the typical CNN process, the R-CNN series employs the enumeration approach to presuppose the target region of the candidate that gradually finalises position data and optimises the classification and recognition position of the item [9]. In contrast, the bounding box and categorization are predicted simultaneously by multiple object detection models directly within the feature map. The CNN models have two operating phases for greater detection precision, while the SSD and YOLO can directly detect categorization and location information, and sensing speed is improved [10].

In order to tackle the compromise of accurate and fast object detection, we offer a new multi-scale deformable convolution network model [11]-[13]. For imaging functionality that is more sensitive to information on object deformation, the CNN employs a convolution with two compensations. In addition, it improves the detection capability of objects with geometric deformations. Second, multi-scale function maps for final detection perform feature fusion operations. In order to anticipate the categorization and position information, image information from multiple scale maps is simultaneously utilised. This change ensures detection speed, increases the information target for small objects and improves the precise detection of objects.

In this paper, a convolutional neural network is used to improve the functionality of object detection using probabilistic neural network in the analysis of surveillance images. The simulation result is conducted with state-of-art deep learning and machine learning models to check the optimal detection efficiency in video streaming or video surveillance systems.

II. RELATED WORKS

In order to categorise area proposals for object identification, Girshick et al [14] devised a multi-stage pipeline, known as the regions with CNN. It breaks the detection problem down into multiple stages, including the proposal for binding boxes, CNN pre-training, CNN finishing, SVM training and the reverse binding boxes. This was a high performance framework and a lot of other work has been taken on.

The Fast R-CNN [15] was proposed to expedite the R-CNN training. Each image patch will no longer be wrapped in a predefined format before being fed into the CNN. Rather, the corresponding characteristics are cut out of the last convolutional layer output characteristic maps.

The Regional Opportunities Proposition Network for the Faster CNN pipeline [16] can therefore be formed on a final basis. The unique candidate framework concept is highly accurate, but it places pressure on the detection speed.

In [17], the authors also suggested the use of a combination of a region proposal approach based on the CNN and the super pixel method of category and position information for the segmentation of objects and image redundancy. Although improved models such as CNN have been developed to accelerate object detection breakthroughs, the process of producing candidate frame areas still adds an unavoidable amount of running time.

YOLO [18] separates the image into grids and forecasts the rating scores for each grid concurrently. In order to anticipate categorization scores for a further regression, SSD [19] constructs anchor boxes at each location. The absence of candidate region will significantly enhance the detection rate, but the basic estimate of the item position ignores information on many small items and dense objects, which will reduce the overall accuracy of detection.

III. PROPOSED METHOD

In this section, we provide the details of how a probabilistic neural network is used for optimizing the CNN hidden layers to improve the process of detection of objects.

A. Probabilistic neural network for CNN optimisation

In the PNN, the fundamental structure is shows in Figure 2. The introduction of feature vectors is managed through the input layer of the network. The number of the pattern layers is equal to the count of the training set. The matching relationship is calculated on this layer between the vector of input and the mode of each training batch. Using the summation layer, probabilities of the same pattern are added. Neurons of the output level are competing neurons, selecting neurons of the highest posterior likelihood density as the whole system output in all patterns. Training sets as an input are provided for the network and trained. The data forecast is made with a neural network training. In this back propagation neural network, a major advantage of PNN [18] – [20] is that complex back calculations are not needed.

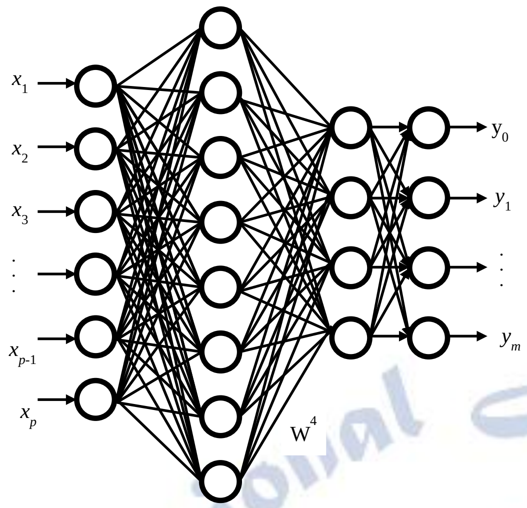

Figure 1: WPNN structure

The input Data X will be transferred from the input layer straight to every unit in the pattern layer when data X is entered into the network. Effective reception of X as a weighted input and pattern layer is obtained by a weighted function of Euclidean distance. This produces a distance expression from Euclidean distance as below:

$$D(i,1) = dist(IW, X) = sum\left(\left(IW(i,:) - X\right)^2\right)^{0.5}$$

where,

i =1,2,...,N

$IW = Q\times R$ - weighted matrix during training,

$X = N\times 1$ - dimension of input matrix, where $X = [x_1, x_2, \ldots, x_N]$

$D = Q\times 1$ - dimension of output matrix.

After the calculation of Euclidean distance, the output of the pattern layer is calculated using radial basis function. In radial function, the Gaussian function is considered.

radbas(n) = e^{-n^2}

Pattern layer output is hence given as,

$$\hat{f}_c^{\,p}(X) = \exp\left[-\frac{\left(X - X^{(P)}\right)^T \left(X - X^{(P)}\right)}{2\delta^2}\right]$$

where,

$\hat{f}_c^{\,p}$, c =1,2,...,C

C - Data categories count

δ - Smoothing factor

The PNN is weighted layered one, where the weights are introduced between the pattern and summation layer, and c^{th} class weights are used to reflect the pattern p specific importance. The summation layer output is modified as below:

$$\hat{f}_c(X) = \sum_{P=1}^{P_c} w_c^{(P)} \exp\left[-\frac{\left(X - X^{(P)}\right)^T \left(X - X^{(P)}\right)}{2\delta^2}\right]$$

where

$w_c^{(P)}$ - weight.

The output of the Pattern Layer is accumulated using the summation layer of PNN and not the output of individual neurons. The output of the pattern layer in WPNN is added to a weighted approach and differences in the output of the pattern layer are taken into account. This procedure is used to make more appropriate data transactions in the prediction capacity and weighted summation. A weights are multiplied by a pattern layer output and the multiplied data is delivered as a summation layer input for the data addition. The weighted summation methodology of WPNN and PNN and this approach are different.

IV. RESULTS AND DISCUSSIONS

In the section, we validate the CNN-PNN models in terms of various performance metrics. This model either predict image deformation explicitly or utilize enhanced learning procedures to make it an excellent problem to govern the registration of a model. Deep learning networks that are either unregulated or use the registration metric themselves gain additional advantages.

The CNN automatically concluded networks are not as suitable for intricate diagnostics since the evidence is difficult to grasp. Therefore, approaches which link observations to facts are important in order to create a path of reasoning for a decision. The authors are firmly sure that the strategy that we are currently taking will have a big influence on the diagnosis of computer aids only until such evidence is reached.

Therefore, these methodologies are extremely crucial, especially in interventions where real-time processing is needed. Initial approaches are available, but further developments continue to exist. In particular, precision learning and variable networks seem to be suited as they ensure prediction results. We feel that there are numerous recent developments, in particular in the field of video surveillance. Reconstruction on the basis of evidence-based methods results in good results (Fig. 2 – Fig. 7)

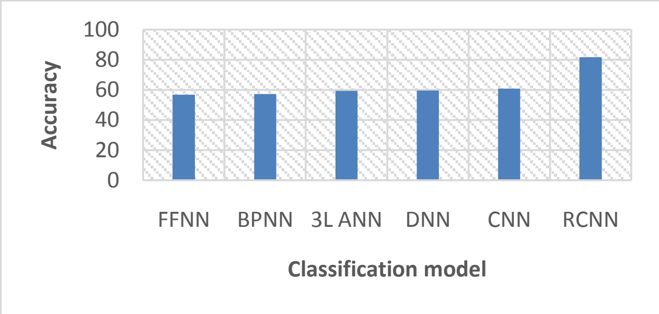

Fig.2.Accuracy

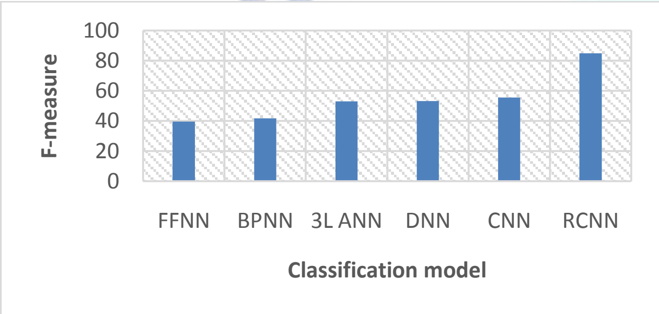

Fig.3 F-measure

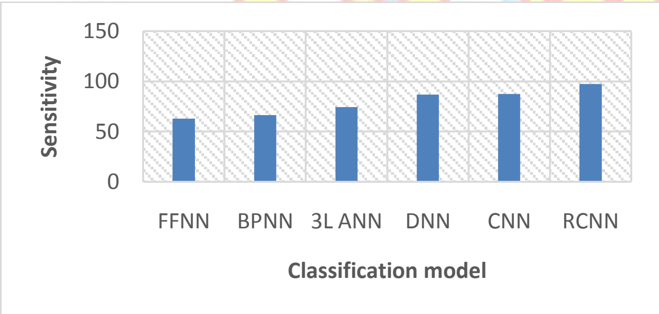

Fig.4. Sensitivity

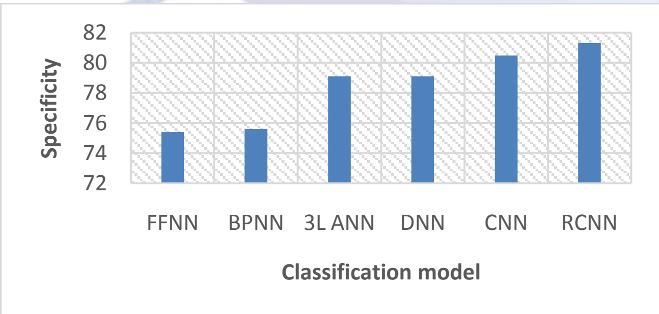

Fig.5. Specificity

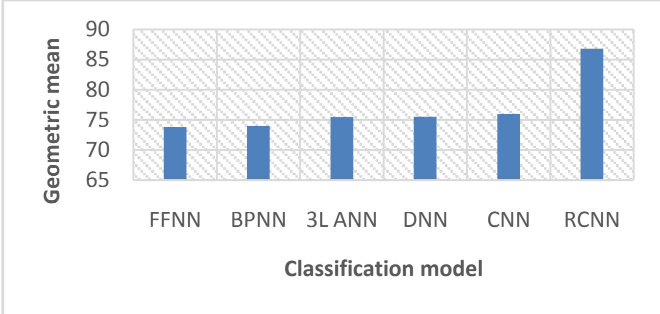

Fig.6. Geometric mean

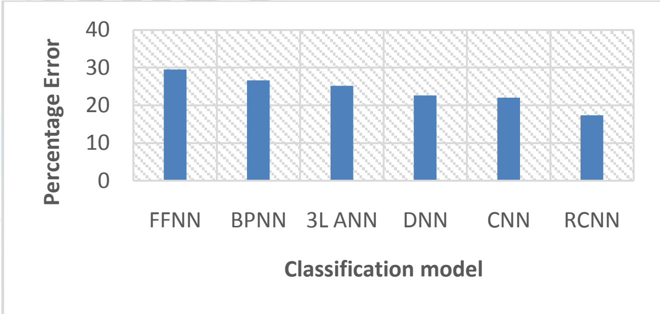

Fig. 7. Percentage error

The way data-driven can best support and preserve the image naturally and securely read is unknown. One of the great advantages of every deep-seated learning technology is the implicit consistency of many classical methodologies. In the future, various new ideas will be generated by this union. Preparation via network level fusion is achievable by precision training. End-to-end algorithms can be developed. In any operation of the hybrid network for this deep fusion, there is just one gradient or subgradient for optimization.

In general, we have noticed how deep learning produces highly efficient CNN structures. Networks can find solutions that match or exceed multiple powerful techniques. Their computer expenditure, however, is normally considerably lower in the traditional areas of medical imaging, when identification, segmentation, registry, restoration and physical simulation are frequently significantly cheaper than state-of-the-art techniques. These improvements are achieved at considerable computational cost during training that can take days, even on GPU clusters. We can exploit this impact to minimise run time at the expense of more training to provide an appropriate problem field and preparation setup.

V. Conclusion

In this paper, CNN-PNN is used in detection of objects in surveillance videos and the study intend to achieve two objectives concurrently. The CNN-PNN effectively confers to the detection of objects in dynamic environment than other methods. The results of simulation shows that the proposed model achieves higher level of accuracy in detecting the objects than other existing methods. Further, the elimination of backgrounds well improved the detection probability using the proposed model than other models.

Acknowledgment

This paper and the research behind it would not have been possible without the exceptional support of my supervisor, Dr. N. Yuvaraj. I am also grateful for the insightful comments offered by my colleagues allowed me to continue my research with the book much longer than I could have hoped. Finally I would like to thank my college and university for the unconditional support for completing this work.

References

[1] Chandan, G., Jain, A., & Jain, H. (2018, July). Real time object detection and tracking using Deep Learning and OpenCV. In *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 1305-1308). IEEE.

[2] Han, J., Zhang, D., Cheng, G., Liu, N., & Xu, D. (2018). Advanced deep-learning techniques for salient and category-specific object detection: a survey. *IEEE Signal Processing Magazine*, *35*(1), 84-100.

[3] Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep learning for generic object detection: A survey. *International journal of computer vision*, *128*(2), 261-318.

[4] Moniruzzaman, M., Islam, S. M. S., Bennamoun, M., & Lavery, P. (2017, September). Deep learning on underwater marine object detection: A survey. In *International Conference on Advanced Concepts for Intelligent Vision Systems* (pp. 150-160). Springer, Cham.

[5] Pathak, A. R., Pandey, M., & Rautaray, S. (2018). Application of deep learning for object detection. *Procedia computer science*, *132*, 1706-1717.

[6] Schneider, S., Taylor, G. W., & Kremer, S. (2018, May). Deep learning object detection methods for ecological camera trap data. In *2018 15th Conference on computer and robot vision (CRV)* (pp. 321-328). IEEE.

[7] Tang, C., Feng, Y., Yang, X., Zheng, C., & Zhou, Y. (2017, July). The object detection based on deep learning. In *2017 4th International Conference on Information Science and Control Engineering (ICISCE)* (pp. 723-728). IEEE.

[8] Wang, W., Lai, Q., Fu, H., Shen, J., Ling, H., & Yang, R. (2021). Salient object detection in the deep learning era: An in-depth survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[9] Wu, X., Sahoo, D., & Hoi, S. C. (2020). Recent advances in deep learning for object detection. *Neurocomputing*, *396*, 39-64.

[10] Wu, X., Sahoo, D., & Hoi, S. C. (2020). Recent advances in deep learning for object detection. *Neurocomputing*, *396*, 39-64.

[11] Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, *30*(11), 3212-3232.

[12] Zhou, X., Gong, W., Fu, W., & Du, F. (2017, May). Application of deep learning in object detection. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)* (pp. 631-634). IEEE.

[13] Jeon, Y., & Kim, J. (2017). Active convolution: Learning the shape of convolution for image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4201-4209).

[14] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).

[15] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).

[16] Jinbo, C. H. E. N., Zhiheng, W. A. N. G., & Hengyu, L. I. (2018). Real-time object segmentation based on convolutional neural network with saliency optimization for picking. *Journal of Systems Engineering and Electronics*, *29*(6), 1300-1307.

[17] Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, *39*(6), 1137-1149.

[18] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

[19] Sun, A., Li, Y., Huang, Y. M., Li, Q., & Lu, G. (2018). Facial expression recognition using optimized active regions. *Human-centric Computing and Information Sciences*, *8*(1), 1-24.