AN OVERVIEW OF DEEP LEARNING FOR OBJECT DETECTION

Rasika Sachin Golhar
Research Scholar
Department of Computer Science & Engineering
Sunrise University, Alwar, Rajasthan.
rasikarode222@gmail.com

Dr. Pawan Kumar Pareek
Research Guide
Department of Computer Science & Engineering
Sunrise University, Alwar, Rajasthan.

Abstract

Theoretical, academic, and practical applications of computer vision continue to heavily rely on object detection. The main source of inspiration for conventional object identification methods was machine learning. This entailed creating features to describe the properties of the item, then integrating those features with classifiers. Deep learning applications, and more especially Convolutional Neural Networks, have generated significant innovation and hopeful results in recent years, and as a result, have attracted a lot of interest on the international arena of computer vision research. This article surveys some of the most significant and recent advancements and contributions made to the field of deep learning's application to object identification. Furthermore, as amply shown, the results of multiple research point to the fact that deep learning applications in object recognition much outperform traditional methods that center on manually created and taught features.

***Keywords**:-Deep learning; Object detection; Convolutional neural networks; Machine learning*

INTRODUCTION

Multiple processing levels in the brain may be seen by closely examining it. Every level may theoretically learn features or representations at increasing levels of abstraction. For instance, according to the general architecture of the visual cortex, the brain first extracts edges, then patches, then surfaces, then objects, and so on. One of the basic ways that the brain processes vision is in this manner. Deep learning, a kind of machine learning that seeks to mimic and copy a comparable computer architecture, is motivated by this discovery.

Deep learning and machine learning have shown to be quite useful in computer vision research. In comparison to other deep learning models, deep convolutional neural networks have outperformed them in a variety of computer vision tasks, including object identification, scene reconstruction, object posture estimation, learning, event tracking, and picture categorization. The key to Convolutional Neural Networks' (CNNs') success in object identification is their capacity to learn important mid-level visual properties, as opposed to the manually created low-level representations that are often used in certain techniques to image categorization. But this raises the question: What exactly is object detection?

In this area, an object is characterized by its primary characteristics, such as its shape, size, color, texture, and other aspects. To identify such an item, it would be necessary for a picture to not only show the object's existence but also its position. Consequently, object detection may be characterized as a technique for finding instances of actual things in pictures. Given that detection requires determining the existence and location of a certain item in an image, it is strongly connected to categorization. A number of items may be identified in a picture, including cars, people walking, buildings, street signs, and human faces.

Deep learning methodologies, such as deep neural networks, region-based convolutional neural networks, and deeply convolutional neural networks, can improve object detection precision, robustness, and effectiveness. This may result in more reliable surveillance and protection systems that can detect moving objects from video[5]. This is crucial for locating abandoned objects that could be anomalies in a scene, such as bombs or explosives, tracking robbery vehicles, and researching and keeping an eye on suspicious behaviors that frequently result in criminal situations in our society. Additionally, intelligent visual cameras may be used to monitor animal behavior in protected areas, either for ethology or the preservation of our natural environment. These cameras get their inspiration from deep learning in object identification. The identification of malignant cells in the human body and image processing in the medical industry are two key applications of deep learning algorithms for object detection.

One of the computer vision tasks that has benefitted from Deep Learning approaches in various research studies is object detection. The Deep Learning algorithms and methods used in object identification in both the fixed-image and video domains are reviewed in this work. It includes a thorough analysis of deep learning methods and how they are used in the area of image recognition. It also clearly illustrates the specific function of deep neural networks in object identification and their superiority to conventional machine learning methods. In addition, it covers common deep learning approaches for image and object identification research and summarizes the most noteworthy results to yet.

A variety of noteworthy and cutting-edge methods have been used recently to increase the detection accuracy of deep learning models and resolve challenging issues that arose during the training and testing of deep learning object recognition models. Modifying the activation function of deep CNNs, Transfer learning, and clever methods for selecting the activation function and the optimization system together for the proposed deep learning model are some of these novel strategies.

This paper is organized in the following manner:

Section 2 introduces object detection and classical machine learning. Section 3 introduces relevant DL models for demanding computer vision object identification problems and describes standard DL techniques and computational procedures for object detection. Section 4 briefly discusses novel ways to improve and optimize deep learning models and handle training and testing concerns. Section 5 ends the survey.

BACKGROUND REVIEW

Object Detection

Color communicates the surroundings. Images show colorful objects. Colors divide scenes. Direct categorization of pixels into "objects and background" is most usual. Colors usually match. The backdrop's colors may reflect the image's values. This strategy separates one item's colors from the backdrop and other things.

Cyganek designed "Road Signs Detection" using Direct Pixel Classification. Classifying pixels decreases dimensionality and optimizes image pre-processing. Use non-color characteristics.

Two pixel-based methods recognize road signs. Sample traffic photographs manually. Many category-specific object identification techniques exist. Most pedestrian detectors use Histogram of Oriented Gradients-based multi-scale sliding window systems.

Filtered channels increase pedestrian recognition. Modern pedestrian detectors were carefully tested to identify common mistakes and training data effects. These research explored cutting-edge technology including filtered channel features and R-CNN detectors and showed advantages over the standard technique.

Color histograms create fuzzy classifiers. Traffic light color histograms were created from thousands of color samples. It's swift. Fuzzy methods frequently provide false positives, lowering accuracy.

SVM uses pixels too. Vapnik introduced SRM-based SVMs. SVMs generalize and regularize. Binary large classifiers. SVMs classify exceptional unobserved data points.

SVM optimization addresses training speed, memory, and variable accuracy. Optimization jobs are large and expensive, and the single-threaded technique cannot keep up with the sophisticated learning process. Even during testing, SVMs may be computationally expensive.

Low-level computer vision detects lines, circles, polygons, and others. Parameterize integers. Form detection employs Hough's line recognition voting method. Ballard added arbitrary figure detection. Calculate frequent forms. The tensor enhanced common form recognition. Local phase and coherence of the tensor may provide rapid and accurate information. the orientation-based Hough transform.

Pictures are not immediately segmented. The structural tensor of a point determines its edge status by determining its local phase and structure.

After rearranging, the lower and upper indices yield the following transform:

McLaughlin and Alder's orientation-based Up Write method for identifying circles, ellipses, and lines is similar. The local orientations are computed to match the main eigenvector's phase to the image data's covariance matrix. Points that pass through successive mean points of neighboring pixel blobs with nearby orientations and follow the assumed curvature form a curve. In another interpretation, the inertia tensor of the pixel intensities determines a curve, and the points acquired may be molded to the image using the least-squares technique.

Figure detection adds to picture identification. Objects are distinguished by their distinctive characteristics. Sparse image coding inspired this concept. The definition seeks an item's most noise-resistant and geometrically stable properties. HOG, SIFT, OpponentSIFT, and PCA-SIFT are popular point descriptors. Mikolajczyk and Schmid studied sparse descriptors and found that the Gradient Location and Orientation Histogram (GLOH) often outperformed the SIFT technique. These methods only capture figure edge detail. High-level representations like object components and mid-level information like edge intersections are harder to build. This is evident even when extracting image attributes from scale-invariant key-points.

Machine Learning methods

Sparse linear regression

Scaled sparse linear regression estimates linear model noise and regression coefficients. It determines an equilibrium using sparse regression by continuously predicting the noise level using the mean residual square and boosting the penalty in proportion to the projected noise level. The iterative method is more costly than sparse regression estimator route or grid computation for penalty rates over a threshold.

Bayesian linear regression

Linear regression using Bayesian inference. When the regression model incorporates errors with a regular distribution and a specified prior distribution is employed, posterior probability distributions of the model's parameters may be calculated. Linear regression as a Bayesian model may automatically calculate the prior variance and noise variance and provide calibrated predictions. Bayesian linear regression aids complicated probabilistic models.

Bayesian logistic regression

Bayesian logistic regression is used to classify statistical items into two categories. This model is linear since the result is determined by the dot product of weight and feature vectors. A hyperplane may describe the classification boundary. The normal linear-followed-by-sigmoid neural network structure is a prominent model. Bayesian logistic regression is a logistic regression variant. Logistic regression uses a logistic function as the dependent variable.

Object detection machine learning models and algorithms are based on complicated statistical and mathematical equations that are interconnected. Learning and training these feature-based approaches frequently takes a lot of data. Below are the most popular object detection machine learning approaches.

1) Deformable Models

This statistical model maps instability while constructing an item's real instance using a prior distribution based on template deformation. Generators and inter-generator bonds describe the model. Template deformation variables identify generators and bonds. A template deformation generates an image data statistical model. This kind

Support Vector Machines

Vapnik invented SVMs, pattern recognition algorithms. Structural Risk Minimization constrains generalization faults in this statistical model. SVMs trigger an exceptional unobserved data point classifier. Due to training speed and memory constraints, SVMs need optimization. SVMs are computationally expensive.

Summary

Most traditional machine learning techniques for object detection struggled to detect complex objects like people, vehicles, and others because they required a lot of prior knowledge and domain knowledge. Another issue with these models is that most need visual coding to capture structural similarities between object class instances. Other data-driven learning methodologies may look better, but they need optimization and processing power.

DEEP LEARNING TECHNIQUES

Deep learning methods

Deep learning uses several methods to improve object recognition. Deep learning Belief networks are RBM bands. RBMs are two-layer, bipartite, undirected graphical models with complete interaction between observable units in each hidden layer. Two-layer neural network stack Stack Auto-Encoders (SAEs) train by lowering reconstruction error. SAE reworks inputs. Convolutional neural networks link neurons more than other neural networks. This model regularizes without an algorithm. CNNs use data-driven filters to extract features to characterize inputs. Deep learning uses feature-based methods to address issues. Using supervised and unsupervised methods, visual quality scores arise spontaneously from data.

Three sorts of deep learning object identification systems are common. The first category, unsupervised feature learning, extracts features using deep learning theory and concepts. These attributes will be passed to basic machine learning algorithms for classification, detection, or tracking, depending on the function. In the second category of supervised learning algorithms, end-to-end learning optimizes the model's feature extractor and classifier units when substantial labelled data is provided. Hybrid deep networks employ generative feature learning models to enhance optimization techniques for deep neural networks and other deep supervised feature-learning methods.

Deep Boltzmann Machine

DBMs are generative feature learning models. DBM layers have hidden variables. Level-specific variables do not interact. A stochastic instrument supports a binary, symmetrically linked Boltzmann Machine (BM). BMs' core learning algorithmic rule makes learning and training them difficult and time-consuming. Each layer of a Deep Boltzmann machine extracts high-order parallels that appear in the layer below during training of hidden attributes. These DBMs may develop complex internal properties, which helps with object recognition and other computer vision concerns. High-level traits may be created from several unlabeled sensory inputs. Then, a very small supply of labeled data may calibrate and improve the network for the purpose. Figure 2 illustrates DBM setup.

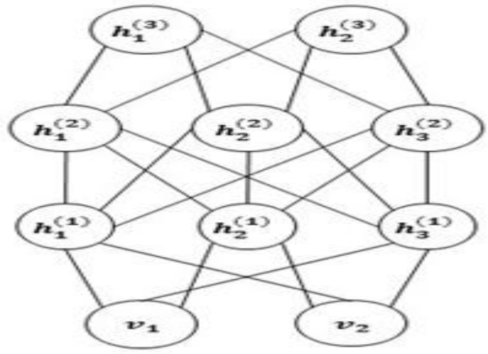

Fig. 2. Deep Boltzmann Machine

Restricted Boltzmann Machine

Restricted Boltzmann Machines are exceptions to the DBM rule of one hidden layer. RBMs and DBMs lack hidden-to-hidden and visible-to-visible links [38]. RBMs learn veiled layers well since one RBM's feature activation is utilized to teach the next. This is the RBM's greatest trait. Figure 3 displays the conventional RBM's architecture.

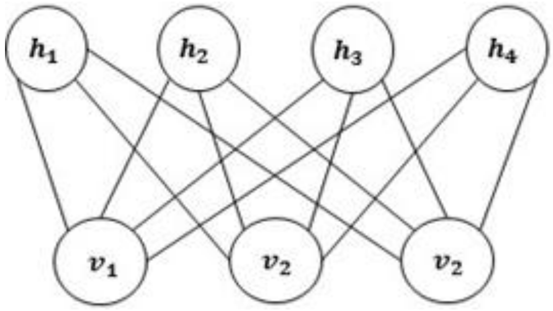

Fig:- Restricted Boltzmann Machine

Convolutional Neural Networks

Convolutional neural networks process input using a grid-like design. Time-series data is a 1-D matrix that tests at regular intervals. Pixel data is another example. Practical convolutional neural networks have shown excellent results. The "convolutional neural network" conducts convolution, a linear mathematical process.

CNN has three main stages. The first layer conducts numerous convolutions to obtain a set number of linear activations. The detector stage, where each linear activation function is applied through a non-linear activation operation like the rectified linear activation function, is the second stage. Pooling functions change layer output in the third stage. Pooling layers replace system yields with outline measurements of neighboring yields.

Deep Neural Networks

Deep neural networks are one discriminative feature learning method. This simple neural network augmentation improves model capabilities and introduces new training issues. Due to their construction, deep neural networks feature complex architectures and components. DNN models have two complicated architectures. First, evaluate its breadth or number of neurons per layer. Second, neuron depth. Deep neural networks may fit data with fewer parameters than conventional neural networks since they can utilize more layers for a more accurate and effective representation. Shallow neural network models require many more parameters to operate.

Deep Belief Networks

A Deep Belief Network (DBN) is a generative deep network that may be utilized as the starting network of a DNN for supervised learning while keeping its network architecture. Discriminately train or fine-tune the DBN using the target labels. Deep Belief Networks have multiple RBM layers of hidden units. Stacked RBM hidden layers are trained using back-propagation. Thus, although layer units are not intraconnected, the DBN structure's connection devices link each layer unit to each other. DBNs are multi-layered RBMs. Figure 6 displays the Deep Belief Network's architecture.

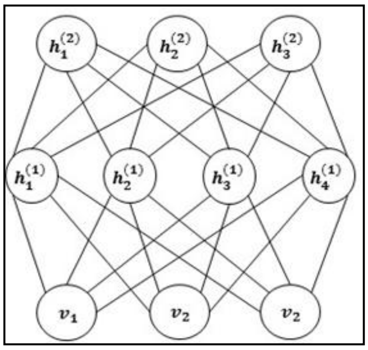

Fig:- Deep Belief Networks

Deep Learning Techniques in Object detection

Deep Learning methods identify objects.

In 2016, Hong Kong University researchers introduced DeepID-Net, a customizable deep Convolutional Neural Network Model for item identification. DeepID-Net learns feature representation and part deformation for many objects. It pre-trains deep CNN models. Pre-trained ImageNet image categorization and localization. Fine-tuned Image Net/PASCAL-VOC item recognition. This model improved RCNN's ImageNet advanced object identification from 31.0% to 50.3%. ILSVRC2014 excels.

2017 writers and researchers proposed a Boosted Convolutional Neural Network for pedestrian recognition. This study suggests a way to locate difficult CNN samples and improve model performance. The Fast-RCNN model can enhance any CNN architecture. Instead of illustrating, this model pictures. Regional ROI-pooling layer concepts pool conv features locally. Smooth loss predicts bounding boxes, whereas soft max trained score class. Image Net pre-trains VGG-16. NVIDIA K40 GPUs train Caffe's networks.

In June 2017, researchers created a deep learning framework integrating depth information and local patterns with a stereo vision system to accurately discriminate humans and cars in various driving conditions. The system includes unsupervised, supervised, and pre-processed local patterns. Auto-encoders and convolutional neural networks power it. Supervised training uses soft-max regression. This paper's robust object identification technique uses Nguyen's disparity methodology to locate, breadth, and height objects. Time-saving Tesla K40 GPUs process the suggested model. The GPU processes the 136 x 136 input picture in 70ms, whereas a PC (Core i7 4.0 GHz, 8.0 GHz RAM) takes 5.5s.

2016 Hangzhou Dianzi University students developed RCNN-based pedestrian detection. Edge Boxes algorithm: ROI from edge boxes. Second, two totally connected 227×227 deep convolution neural network layers receive region suggestions. Finally, linear SVMs classify CNN features and regions. Standard algorithms miss 23% more pedestrian identification. Viola-Jones, HOG, and Selective Search region proposals are tested. Viola-Jones misses 72%, HOG 46%, Selective Search 24%.

Sub-category-aware convolution neural networks were demonstrated in 2017. Traditional object recognition systems use sub-category information to identify two CNN-based objects. The system cannot manage object size, occlusion, or truncation. Second, difficulty reproducing 2D segmentation border, 3-dimensional posture, and occlusion link. The detection approach beats Fast R-CNN. Pyramids scale. The feature extrapolation layer scales the convolution feature pyramid after the previous convolution layer for feature extraction. The developed RPN outperforms the baseline RPN detection network in all KITTI val set item, pedestrian, and cyclist detection metrics.

2016 authors presented "You Only Search Once: Single Real-time Object Detection". Convolutional neural networks identify things. Image features project every bounding box. Every picture has these boxes. Real-time detection completes model training. A grid cell may identify an item if its center is in the input image's S×S grid. Every grid cell projects B boundary boxes and confidence ratings. The model's confidence ratings represent its ability to forecast the box's contents and attributes. Regression detection is easy. The neural network must discover an unusual picture. TitanX GPUs execute the principal network at 45 fps without batching. Bounding box predictions are geographically limited by YOLO since each grid cell speculates two boxes and one class. Space limits make the neural network model predict fewer neighbors. The model fights birds.

June 2017 witnessed Faster R-CNNs with region proposal networks for real-time item identification. This research develops Region Proposal Networks (RPNs) using convolution feature maps, region-based feature detectors like the huge R-CNN, and a few additional convolution neural networks to rapidly regress region boundaries and objectness scores at every point on a regular grid. Fast R-CNN uses deep convolution neural networks. Fast-CNN-RPN scores 59.9% on PASCAL VOC dataset with 300 suggestions. Shared convolution calculations make RPN quicker than SS and EB. Less concepts imply cheaper region-wise fully-connected layers.

Zhang Ren and Sun suggested Image Recognition Deep Residual Learning in 2016. Deep residual learning is matched to stacked model layers in this research.

Residual network shortcuts implement F(x)+x. VGG neural networks influenced plain network architecture [48]. Most convolution layers utilize 3 × 3 filters and these two design guidelines: When output feature map size is halved, filters double to maintain layer complexity. Faster-RCNN detection [5]. Optimizing deep residual networks is easy. The network's depth improves classification and detection over VGG-16.

Several authors proposed SSD:

Single-shot multibox detectors identified objects in 2017. A convolutional feed-forward network produces scores of a specified size for object class instances within bounding boxes in the SSD approach. Next, during the final detection cycle, is non-maximum suppression. Classifying high-quality photos uses the base network, which is limited before any image classification layers.

The network then receives an auxiliary structure for detection with preset key attributes. Hard negative mining stabilizes training and speeds optimization. After matching, sort the negative default box training instances. SSD rapidly adjusts bounding box size. It performs worse on little items. The upper layers of microscopic things often lack information.

In 2017, K. Kang, H. Li, and others proposed a Tubelet Proposal Network to efficiently manufacture movie tubelets. Based on single-frame static region proposals, the first sub-network accumulates visual data over time. The network has two primary parts. Since CNNs have large receptive (RF) fields, feature map pooling may be done at the same bounding box locations across time to extract moving object visual attributes.

The second component, based on integrated visual input, is a regression layer that approximates bounding box temporal displacements to infer tubelets. TPN develops candidate object tubelets, and CNN-LSTM classifies each object tube bounding box into various object classes. Tubelet proposals may provide useful temporal data to increase detection accuracy.

INNOVATION OF DEEP LEARNING MODELS

Computer vision scientists and engineers have long used advanced deep learning models. Benchmarks and training have improved.

DNN training requires computation. Training uses the stochastic gradient descent approach, which is hard to parallelize. This hinders large-scale learning. A GPU computer can train DNN-based object detectors with many huge image training datasets. Scaling this technique with numerous long training data sets gets mind-boggling.

Deep Stacking Neural Networks (DSNs) decrease computational errors during DNN training The DSN stacks capacity or classifier modules to learn complex functions or classifiers. Basic modules employ supervised information for stacking.

Superior stacked classifiers add features by concatenating classifier output from lower modules with raw input features. Conditional random fields stack easily. Supervisory information stacks essential links in the DSN. Figure 7 depicts Deep Stacking Network.

Transfer learning (TL) has helped engineers and academics construct deep learning models that can detect items in complex environments with many objects. Transfer learning is applying knowledge from one problem to another. The final deep learning model is retrained for a new item identification circumstance using a pre-trained deep CNN classification model like Alexnet, VGG-16, or VGG-19. Pretrained CNN models perform better.

Changing CNN architecture performance may also improve object recognition deep learning models. Performance underpins activation. CNNs employ Sigmoid, Tanh, and ReLU nonlinear functions. The sigmoid function ranges real values. When activation is 0 or 1, the function gradient is zero.

The neuron layer provides a nonzero mean signal input because the average function output value is not 0. It enhances neuron input. Weight increases. These difficulties slow parameter convergence, training efficiency, and model recognition. Despite being a sigmoid function, tanh(x) = 2sig moid(2x) − 1 translates a specific input into the range. Gradient saturation and constant positive weight plague sigmoid and tan h functions.

The ReLU function has the following properties: Unsaturated gradients have I (x>0) This eliminates gradient dispersion in reverse propagation, allowing fast modification of the neural network's initial layer. Simple computation: The ReLU layer's thresholds are 0 if x = 0 and x if x > 0. ReLU units may "die" or lose validity, which is tragic. A 'dead' ReLU's output is constant. Changing the activation function may reduce gradient saturation and positive weight in deep learning models. This may accelerate convergence.

A creative combination of activation function, regularization approach, weight update method, and variable step size may revolutionize deep learning models. The tan *h* activation function, dropout approach, gradient descent with momentum, and variable step size may enhance a deep CNN model on numerous fronts.

Other improvements are being made to deep learning CNN models for object recognition training and testing. Some deep learning models are good at micro item identification and category-specific object recognition. Thus, new deep learning models must be tested on more datasets.

CONCLUSION

This paper covers deep learning algorithm advances in object detection. Recent tests and studies reveal that deep learning systems can identify objects.

Convolutional Neural Networks, Deep Neural Networks, and Region-based Convolutional Networks have been used as the baseline for many robust detection systems and have achieved contemporary performance on various datasets in many experiments.

Deep learning has showed promise in object recognition, but further study on larger datasets with different categories is needed to confirm its efficacy. Due to the processing needed to train and evaluate deep learning models, more experiments on multiple platforms are needed to create the most user-friendly deep learning computing platform.

REFERENCES

1. *M. Ranzato, F.J. Huang, Y. Boureau and Y. LeCun, "Unsupervised Learning of Invariant Feature Hierarchies with Applications to Object Recognition, Proc. of Computer Vision and Pattern Recognition Conference" (CVPR 2007), (Minneapolis 2007)*
2. *M. Zeiler., Hierarchical Convolutional Deep Learning in Computer Vision (New York University, Jan. 2014)*
3. *Goodfellow, Y. Bengio and A. Courville, Deep Learning (2013), pages 654-720.*
4. *P. Dollár, R. Appel, S. Belongie and P.*

Perona, Fast feature pyramids for object detection, Transactions on Pattern Analysis and Machine Intelligence, Vol. 36, No. 8 (2014), pp. 1532–1545.

5. S. Ren, K. He, R. Girshick and J. Sun: Faster R-CNN: Towards Real- Time Object Detection with Region Proposal Networks, in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 6 (Jun. 2017), pp. 1137-1149.

6. H.C. Shin, Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning, in IEEE Transactions on Medical Imaging, vol. 35, no. 5 (May, 2016), pp. 1285-1298.

7. D. Hu, X. Zhou and X. Yu, Deep learning and its applications in Object Recognition and Tracking, Computer Science and Engineering Technology (CSET2015), Medical Science and Biological Engineering (MSBE 2015), pp. 87-92.

8. W. Nam, P. Dollár and J. Han, Local decorrelation for improved pedestrian detection, In Advances in Neural Information Processing Systems (2014), pages 424–432.

9. J. Yang and G. Yang, Modified Convolutional Neural Network Based on Dropout and the Stochastic Gradient Descent Optimizer, Key Laboratory of Advanced Manufacturing Technology of Ministry of Education, Guizhou University,2018 (Guiyang, China, 2018)

10. S. Li, W. Liu and G. Xiao, "Detection of Srew Nut Images Based on Deep Transfer Learning Network," 2019 Chinese Automation Congress (CAC), Hangzhou, China, 2019, pp. 951-955.

11. H. Yang, S. Jiao and P. Sun, "Bayesian-Convolutional Neural Network Model Transfer Learning for Image Detection of Concrete Water-Binder Ratio," in IEEE Access, vol. 8, pp. 35350-35367, 2020.

12. K. L. Masita, A. N. Hasan and S. Paul, "Pedestrian Detection Using R- CNN Object Detector," 2018 IEEE Latin American Conference on Computational Intelligence (LA-CCI), Gudalajara, Mexico, 2018, pp. 1-6.

13. Y. Ouyang, K. Wang and S. Wu, "SAR Image Ground Object Recognition Detection Method based on Optimized and Improved CNN," 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 2019, pp. 1727-1731.

14. K.E.A. van de Sande, T. Gevers, C.G.M. Snoek, Evaluating Color Descriptors for Object and Scene Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32 (2010), No. 9, pp. 1582–1596

15. B. Cyganek, Color Image Segmentation With Support Vector Machines: Applications To Road Signs Detection, International Journal of Neural Systems, Vol. 18, No. 4, World Scientific Publishing Company (2008), pp. 339–345.

16. S. Zhang, R. Benenson and B. Schiele, Filtered channel features for pedestrian detection, In CVPR (2015), pages 1751–1760.

17. C.P. Papageorgiou, M. Oren and T. Poggio, A general framework for object detection. In Computer vision, 1998. Sixth international conference (1998), IEEE pages 555–562.

18. N. Dalal, and B. Triggs, Histograms of Oriented Gradients for Human Detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1 (2005), pp. 886–893.

19. H. Wu, Q. Chen and M. Yachida, Face Detection From Color Images Using a Fuzzy Pattern Matching Method, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 6 (1999), pp. 557– 563.

20. V. Vapnik, Statistical Learning Theory, Wiley New York, Inc. (1998).

21. N. Lopes and B. Ribeiro, Support Vector Machines (SVMs), In Machine Learning for Adaptive Many-Core Machines - A Practical Approach. Studies in Big Data, Vol. 7, Springer, Cham (2015)

22. P.V.C. Hough, Machine Analysis of Bubble Chamber Pictures, Proceedings of International Conference on High Energy Accelerators and Instrumentation, CERN, (1959)

23. D.H. Ballard, Generalizing the Hough transform to detect arbitrary shapes, Pattern Recognition, Elsevier, Vol. 13, No. 2 (1981), pp. 111– 122.

24. B. Jähne, Digital Image Processing. 6th edition, Springer-Verlag (2005)

25. B. Cyganek, Orientation-based Hough transform and the UpWrite method, figure, Object detection and Recognition in Digital Images: Theory and Practice, May (2016), pp. 365.

26. R.A McLaughlin, and M.D. Alder, The Hough Transform Versus the UpWrite, IEEE Transactions On Pattern Analysis And Machine

Intelligence, Vol. 20, No. 4 (1998), pp. 396–400.

27. *J. Bigun, G.H. Granlund and J. Wiklund, Multidimensional Orientation Estimation with Applications to Texture Analysis and Optical Flow, IEEE PAMI, Vol.13, No.8 (1991), pp. 775–790.*

28. *T. Hazan, S. Polak and A. Shashua, Sparse Image Coding using a 3D Non-negative Tensor Factorization, ICCV 2005 10th IEEE International Conference on Computer Vision, Vol. 1 (2005), pp. 50– 57.*

29. *Y. Ke, and R. Sukthankar, PCA-SIFT: A More Distinctive Representation for Local Image Descriptors, Computer Vision and Pattern Recognition, Vol. 2 (2004), pp. 506–513*

30. *K. Mikolajczyk and C. Schmid, A Performance Evaluation Of Local Descriptors, Vol. 27, No. 10 (2005), pp. 1615-1630.*