International Journal of Future Generation Communication and Networking
Vol. 13, No. 3s, (2020), pp. 88–93

Inspection of metal parts based on CNN using YOLO Network.

Prashant S. Gandhile[1], Sanika Y. Gandhe[2], Ram R. Mankar[3], Sapna R. Nikam[4], Prof. Bhagyashree D. Shendkar[5]

Computer Engineering, Sinhgad Institute Of Technology & Science, Narhe, Pune-41

[1]prashantgandhile123@gmail.com

[2]gandhesanika@gmail.com

[3]rammankar003@gmail.com

[4]sapnanikam4321@gmail.com

[5]bdshendkar_sits@sinhgad.edu

Abstract

The machine vision is a latest approach which provides automation in inspection of systems. Automatic inspection using machine learning and Convolutional Neural Network (CNN) is a popular trend for detection of metal parts on automobile infrastructure in order to speed up the process and improve accuracy of results. Deep CNN are able to perform both the feature extraction and classification. This proposed model is used to detect, label and classify metal parts and inspect them based on image analytics using deep learning. This is used to eliminate manual errors and provide better, fast and accurate result.

Keywords—Image Analytics, CNN, Object detection, Classification

I. INTRODUCTION

Automatic inspection and defect detection using image processing is an area of machine vision that is being widely adopted in many industries. It is used for high through put quality control in production systems such as the detection of flaws on manufactured surfaces, e. g. metallic rails or steel surfaces. The proposed model is to design autonomous devices that automatically detect and examine specific visual patterns from images and videos in order to overcome the limitations and improve the performance of the traditional inspection systems that depend heavily on human inspectors. They are generally composed of a pipeline of several steps; each one introduces a set of challenges. First, the acquisition step requires efficient calibration of various types of sensors such as cameras and lighting systems. The quality of the images produced by the acquisition systems impacts directly the performance of the subsequent analysis steps. Different layers of CNN contribute in the processing of images. This automation in the process helps in automation of inspection of parts and hence classifies good and bad parts.

II. PROPOSED APPROACH

The first objective is collection of data from samples of good and bad metal parts. A VGA camera is mounted on the table. The setup is made which is able to capture images of parts for training data. About 50-100 images of each metal part sample are captured. This is served as an input to the next objective.

ISSN: 2233-7857 IJFGCN

The second objective receives input of images from the first task of data collection. This objective is based on the CNN model which carries out image processing for the metal sheet images. In this way the object detection and labeling are done as a part of image analytics. This detects the components embedded on the metal sheets and labels its types. This is stored as training dataset. The third objective is capturing the test image and identifying it as good or bad part. In this the comparison of training data and test data is done and the part is classified as good or bad part. In this way, the planned objectives are carried out and the machine learning and CNN model is implemented.

A. To collect data for and from inspection: A hardware setup is assembled, a camera is mounted on a tripod and this camera is connected to the computer. Images of defective parts and undefective parts are captured using VLC player feature. Images of both the samples are captured using a VGA camera of 5 Megapixels (Up to average 50 images in PNG format are captured) and software VLC player using snapshot feature.

B. To detect parts and label them: Author Ananthu Raj et.al [2] proposed an algorithm for defect detection in PCB boards using Image Processing. The board is used for holding various electronic components. Various defects can arise in these boards during manufacturing process. Similarly, scenario occurs while assembling metal sheets, brackets, punching holes of different size and mounting different components. Object detection is locating the object to be processed in the image. The aim is to detect metal sheet, the nuts, bolts punched holes. Labeling is naming the detected objects according to the trained data set. Labeling of types of nuts, bolts viz. m4, m8. Labeling missing components.

C. To classify parts as good and bad parts: Comparison of test data with training dataset Classification of good and bad parts according to the labels. If any missing component is found, classify as bad, else good part. Display the result. Thomas Konrad et.al [1] proposed detection of surface defects using machine learning and CNN approach. This methodology also inspected the parts through CNN as defect free or defect less. The advantages of this paper are CNN's for automatic defect detection in infrastructure inspection is presented and computational cost is minimized.

III. SYSTEM ARCHITECTURE

In the figure 1.1, system architecture is specified which describes the overall functioning of the system. Different phases and levels of the working of the system is described in this architecture. This system architecture consists of system components and the sub-systems developed, that will work together to implement. the overall system. The input and output at each phase are shown in detail.

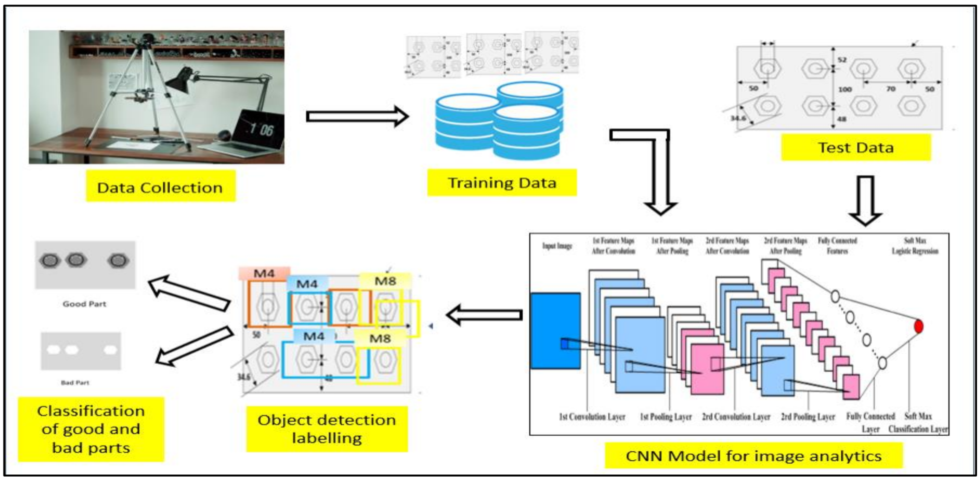

Figure 1.1 - System Architecture

ISSN: 2233-7857 IJFGCN

Different phases and levels of the working of the system is described in this architecture. This system architecture consists of system components and the sub-systems developed, that will work together to implement the overall system. The phases describe different phases which are data collection, managing of training and test data, building CNN model, object detection and labelling and classification of good and bad parts. The first phase is data collection in which training data is collected. The samples of good and bad parts are stored which undergoes training and analysis is done. This training data is then further passed to CNN model for image acquisition. Here, feature extraction and image analysis are carried out. The CNN model use Python libraries and different layers of CNN are worked upon. This model helps in serving the objectives which include object detection, object labelling and classification of data. Later the image to be tested is passed on to the CNN model, where the comparison with training data set is carried out. In this way, the test image is classified as good part or bad part.

IV. YOLO FRAMEWORK FOR OBJECT DETECTION

Object detection in images means not only identify what kind of object is included, but also localize it inside the image(obtain the coordinates of the "bounding box" containing the object). In other words, **detection = classification + localization.**

YOLO (You Only Look Once) uses deep learning and convolutional neural networks (CNN) for object detection. It takes the entire image in a single instance and predicts the bounding box coordinates and class probabilities for these boxes. The biggest advantage of using YOLO is its superb speed – it's incredibly fast and can process 45 frames per second. YOLO also understands generalized object representation.

Different phases of object detection are Bounding Box Prediction, Class Prediction, Predictions Across Scales, Feature Extractor.

- Bounding Box Prediction: It is the same as YOLOv2 [9]. The top most **coordinates (**tx, ty) and the image width tw and image height are predicted. During training, sum of squared error loss is used. And objectness score is predicted using logistic regression. It is 1 if the bounding box prior overlaps a ground truth object by more than any other bounding box prior. Only one bounding box prior is assigned for each ground truth object.
- Class Prediction: Softmax is not used as used in previous YOLO [8]. Instead, independent logistic classifiers are used and binary cross-entropy loss is used. Because there may be overlapping labels for multilabel classification such as if the YOLOv3 is moved to other more complex domain such as Open Images Dataset.
- Prediction Across Scales: 3 different scales are used. Features are extracted from these scales like FPN. Several convolutional layers are added to the base feature extractor Darknet-53. The last of these layers predicts the bounding box, object**ness** and class predictions. On COCO dataset, 3 boxes at each **scale**. Therefore, the output tensor is $N \times N \times [3 \times (4+1+80)]$, i.e. 4 bounding box offsets, 1 object prediction, and 80 class predictions. Next, the feature map is taken from 2 layers previous and is **unsampled** by 2×. A feature map is also taken from earlier in the network and merge it with our **unsampled** features using concatenation. This is actually the typical encoder-decoder architecture, just like SSD is evolved to DSSD. This method allows us to get more meaningful semantic information from the **unsampled** features and finer-grained information from the earlier feature map. Then, a few more convolutional layers are added to process this combined feature map, and eventually predict a similar tensor, although now twice the size. k-means clustering is used here as well to find better bounding box prior. Finally, on COCO dataset, (10×13), (16×30), (33×23), (30×61), (62×45), (59×119), (116×90), (156×198), and (373×326) are used.

ISSN: 2233-7857 IJFGCN

- Feature Extractor: Darknet-53: Darknet-19 classification network is used in YOLOv2 for feature extraction. Now, in YOLOv3, a much deeper network Darknet-53 is used, i.e. 53 convolutional layers. Both YOLOv2 and YOLOv3 also use Batch Normalization.

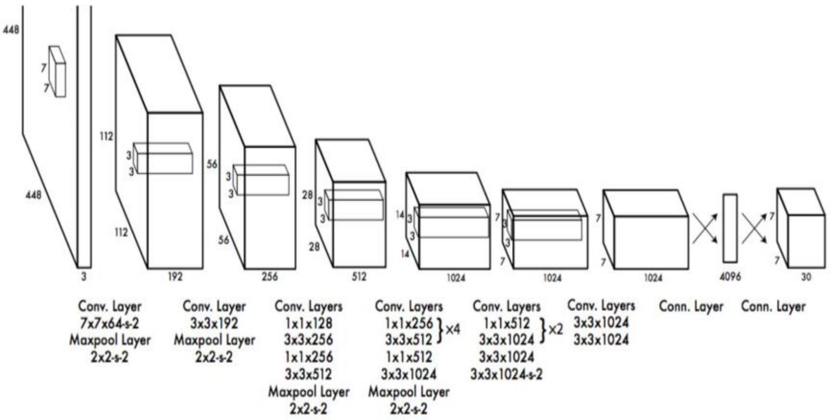

Figure 1.2 YOLO Framework

YOLO network has 24 convolutional layers followed by 2 fully connected layers, 1×1 reduction layers followed by 3×3 convolutional layers. The final output of our network is the $7 \times 7 \times 30$ tensor of predictions.

V. IMPLEMENTATION

In the following subsections, the setup and implementation details are specified and corresponding results are presented.

A. *Inspection data generation*

A hardware setup is assembled, a VGA camera is mounted on a tripod and this camera is connected to the computer. Images of defective parts and non-defective parts are captured. The images are captured with different angles. The light conditions are kept constant. Since the key task of the network is to determine the class of the image, namely if a defect exists or not, it is especially important that the two-image data sets only differ by the existence of a defect and not differ from other factors.

B. Data augmentation and splitting

The risk of overfitting is a major challenge during the training process of a neural network. In order to generate different sets of images for training, validation and test purposes, the following measures are taken to minimize overfitting. For the same, previously described reasons, the training data set is augmented by rotations and translations of its contents. The data set is randomly split in three different data sets: the training, validation and test data set, in the ratio of 60:20:20.

C. Training data

The training data set is used to train the weights of the CNN. The data used for training is labelled with its respective classes. The training progress is continuously checked by calculating the output for the validation data set. The training stops if loss function between predicted and labelled output of each of the last five validation epochs is greater than the previous (sixth last) epoch. The training of the CNN model is done using YOLO network.

ISSN: 2233-7857 IJFGCN

D. Testing data samples

The test data is then tested over the trained YOLO network. The components on the parts are detected and labels are also detected. On the basis of components present or missing, the part is said to be good or bad part. Thus inspection of parts takes place.

E. Results

The result is in the form of image which specifies the bounding boxes of object detected along with the class labels. The accuracy percentage of each component detected is displayed.

The results also mention the performance measures of the model like accuracy, mAP of the model. mAP of YOLO model is 61.39%

F. Storing the reuslts in database

The reuislts of the inspected image are later stored for record. The image is stored with labelled image and specifications.

V. CONCLUSION

Images taken via a real-time camera were analyzed with the image processing algorithm designed within the scope of the study and feature extraction was performed. After performing the same operations on the test images, feature matching method was used and defects on the products were detected. A system to detect and classify defects on metal sheet is developed based on the object detection algorithms. In this paper, a system is developed based on YOLOv3 to detect the components embedded on a metal sheet and label them. The components are further classified on the basis of different classes. These classes are: correctly classified parts, missing parts, misplaced parts, defective parts, etc. Machine vision provides the automation in inspection work field providing better results and minimizing manual errors.

ACKNOWLEDGEMENT

We express our gratitude to our guide Prof. Mrs. B. D. Shendkar for competent guidance and timely inspiration. It is our good fortune to complete our project under her able competent guidance. This valuable guidance, suggestions, helpful constructive criticism, keeps interest in the problem during the course of presenting this ``Inspection of parts using image analytics and classification of good parts and bad parts'' project successfully. We are very much thankful to Dr. G. S. Navale, Head, Department of Computer Engineering and also Dr. R. S. Prasad, Principal, Prof. S. A. Kulkarni, Vice principal, Sinhgad Institute of Technology and Science, Narhe for their unflinching help, support and cooperation during this project work. We would also like to thank the Sinhgad Technical Educational Society for providing access to the institutional facilities for our project work.

ISSN: 2233-7857 IJFGCN

REFERENCES

[1] Thomas Konrad, Lutz Lohmann, and Dirk Abel1, “Surface Defect Detection for Automated Inspection Systems using Convolutional Neural Networks,” in 2019 27th Mediterranean Conference on Control and Automation (MED) (IEEE).

[2] Ananthu Raj , Sajeena A, “Defects Detection in PCB Using Image Processing for Industrial Applications,” in 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018) (IEEE).

[3] Zhipeng Li ,Jun Zhang ,Tao Zhuang ,Qiuyue Wang , “Metal surface defect detection based on MATLAB,”in 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Con- trol Conference(IAEAC 2018).

[4] Oumayma Essid1, Hamid Laga, Chafik SamirI,“Automatic Detection And Classification Of Manufacturing Defects In Metal Boxes Using Deep Neural Networks” in PLOS ONE.

[5] S. Ren, K. He, R. Girshick, X. Zhang, and J. Sun, “Object detection networks on convolutional feature maps,”arXiv:1504.06066, 2015.

[6] Ross Girshick Jeff Donahue Trevor Darrell Jitendra Malik. “Rich feature hierarchies for accurate object detection and semantic segmentation”. arXiv:1311.2524v5, 22 Oct 2014.

[7] S. Ren, K. He, R. Girshick, and J. Sun. “Faster r-cnn: Towards real-time object detection with region proposal net-works.arXiv” preprint arXiv:1506.01497, 2015

[8] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. “You Only Look Once:Unified, Real-Time Object Detection”. arXivpreprint arXiv:1506.02640, 2015.

[9] J. Redmon and A. Farhadi. Yolo9000: Better, faster, stronger. In Computer Vision and Pattern Recognition (CVPR), 2017IEEE Conference on, pages 6517–6525. IEEE, 2017.

[10] J. Redmon and A. Farhadi. Yolov3: An incremental improvement. arXiv, 2018.

ISSN: 2233-7857 IJFGCN