Part Marking Detection Using Machine Learning

[1]Radha Tiwari, [2]Sangita Bharkad

Department of Electronics and Telecommunication Engineering,
Government Engineering College Aurangabad, INDIA

***Abstract*: For Paper presents Faster R-CNN based deep learning implementation of numbering of mechanical parts such as gears. Parts are photographed in real time. In this research works the parts are sorted into three categories- the correctly numbered parts, non-numbered parts and over-ride numbered parts- through image processing followed by deep learning algorithm. For this work Mobile-Net Model on TensorFlow Machine Learning platform to accomplish part identification. Visual inspection validates the technique to 95% accuracy in real time detection.**

Index Terms: Image Processing, Object Detection, Faster R- CNN, Machine Learning, Deep Learning.

I. INTRODUCTION

Inventory Management, Supply Chain Management and Product Recall and Replacement deal with the problem of identical part numbers for different parts or same parts with different part numbers. This problem requires visual inspection and correction. Manual inspection is prone to mistakes and is time consuming. A number of manufacturers for same part allot different part numbers and at times two different parts may have a numbering clash. We demonstrate a real time computer vision and deep learning algorithm that automates this process. Part numbers also follow different naming conventions to provide additional details like process of manufacturing, dimensions of parts, materials used, date of manufacturing etc. At times parts may be stamped second time with over-ride part numbers. A machine learning algorithm is used not only to validate part with its stamped part number but also to extract additional information about the part. In this paper we demonstrate computer vision algorithm to detect gears and validate stamped part numbers in real time.

Constant picture securing of apparatuses proceeding onward a transport line has issue of quick picture obtaining under factor light conditions, with surface sparkle and foundation reflection. These encompassing conditions change with lighting conditions, speed of the transport line, part arrangement and wrapping up. Profound learning calculations [1,2]used in our exhibition, oblige these mistakes without trading off the preparing time. In past, Industrial applications [3,4] have utilized comparative profound learning strategies yet at much more slow speeds [5]. In this paper, we exhibit the utilization of Faster R- CNN [6] to moderate location and article order speed without trading off exactness.

II. RESEARCH BACKGROUND

2.1 Conventional Methodology

First step in this problem is image acquisition, image registration, edge detection, image segmentation into objects and object classification using boxing. Object classification which is based on supervised learning using class-specific and class-agnostic boxes [7,8,9]. Traditionally used Over- Feat method [8] uses a Fully-Connected (FC) neural net layer to establish box coordinates for a single object for localization of processing field. FC layer is transformed to CONV Layer to aggregate multiple boxes around class-specific objects. Multiple box method [9,10] with couple of hundred boxes is used in our Faster R-CNN Method for class-specific object detection.

Over feet Method and Multiple Boxing can be applied to a solitary picture or littler picture crops as info. An order picture pyramid is set up with various degrees of exactness as a piece of preparing grouping. An adaptively-sized pool (SPP) of CONV highlights from this pyramid is utilized for semantic division [2] followed by district based item discovery [7]. Quick Region-based Convolutional Neural Networks (R-CNN) Method in [5] shows start to finish indicator preparing with significant exactness and speed utilizing CONV layer highlights. Rismayati and Rahari exhibited utilization of CNN to identify Salak Fruits [11]. They utilized six layered channel of size 3 x 5 x 5 as first layer and 18 layered channel of size of 6 x 2 x 3 and accomplished a precision of 81.5% in recognition. Faster R-CNN method proposed in this research work is more precise in comparison with R-CNN. Fast R-CNN method for Image detection provides higher accuracy at the level of 1 Frame per Second (FPS). Following Table shows the correlation between Region-based Convolutional Neural Networks (R-CNN), Fast Region-based Convolutional Neural Networks (R-CNN) and Faster Region-based Convolutional Neural Networks (R-CNN).

Table 1: Comparison of R-CNN techniques

	R-CNN	Fast R-CNN	Faster R-CNN
Test time (per image)	50 sec	2 sec	0.2 sec
Speed up	1x	25x	250x

In this proposed research work, Tensor Flow and Portable Net are utilized to actualize Faster R-CNN for recognizing and grouping gear images. Regional Proposal Network (RPN) and sigmoid capacities gives us better output performance from our framework.

2.2 Basic Concept

Convolutional Neural Network (CNN) is an integration of deep neural networks and feed-forward artificial neural networks which is used for providing accuracy in performance for computerized vision tasks like classification of images and also for detection of images [12]. CNNs are just like similar to a traditional neural network, but it has deeper layers which differs it from traditional neural network. It consists biased values, weight function and output layers from the non-linear activation function. The CNN neurons are organized in a height, depth and width volumetric way wise. Figure 1 shows the CNN architectural framework, it consists of convolutional layer, pooling layer and totally related layer. Generally Convolutional layer and pooling layer are alternated with each other. Here, from left to right the depth size of each filter is in increasing manner whereas the output size of height and width is in decreasing manner. The last stage in CNN architectural framework is of fully connected layer which is same as the last layer in traditional neural network. Here the image is provide as input which have some pixel values. The Image is of RGB mode and has three dimensions of height, width and depth [13]. The output from neurons which are attached to local regions of input are computed by Convolutional layer. The parameter of layers are made up of set of kernels or learnable filters and it is convolved over height and width of a volume object present at input and computes dot product of input and filter entries. From this Two-dimensional activation map of that filter is produced as a consequence the network commands filters that it has to be trigger when it detects some specific feature at some dimensional position at the input. The elementwise activation function is provided by the layer called Rectified Linear Unit (ReLU) . The function Rectified Linear Unit is characterized as,

$$f(x) = \max(0, x) \qquad (1)$$

The above function provides zero for negative values and for positive values it rises linealy. This will not make any affect on volume size. The maximum activation in a region is provided as a output by the pooling layer. This provides down-sampling of height and width dimensions. Here fully connected layer is present at output layer which is relevant to the Final Layer in neural network. This output layer uses softmax activation for providing output probability distributions over number of output classes.

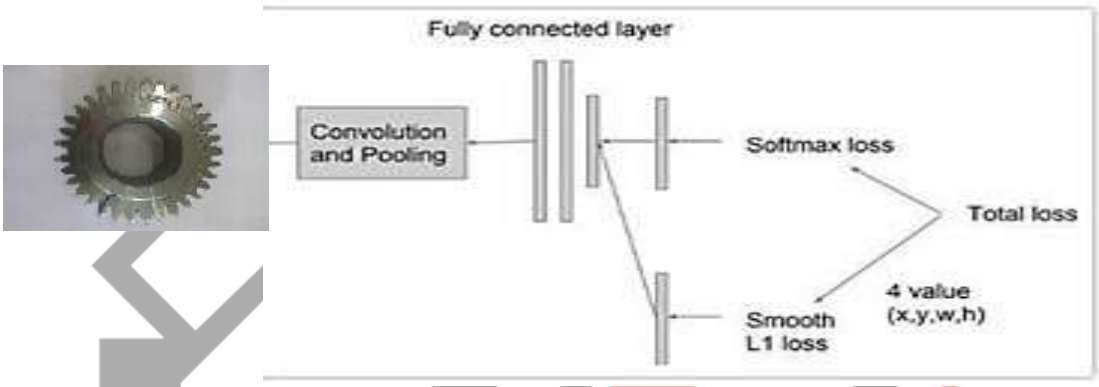

Fig.1. Architecture of CNN

III. PROPOSED METHODOLOGY

In this research work, detection of gear part numbering and object is achieved using Faster R-CNN on TensorFlow plat form. Here, a Soft-Max classifier is used. Training sequence is 104 images of gear from a manufacturer with correct gear numbers engraved on them, 100 images of gears with two part-numbers stamped and 118 images with no part numbering. Various steps of object and part number detection are detailed in this section.

3.1Object Detection using Faster R-CNN

Figure 2 shows processing phases of Faster R-CNN. Here, The input data layer is firstly passed to the CONV layer which computes the feature maps. Region Proposal Network (RPN) collects all images having different sizes and shapes to forecast the group consists of object and object-ness (object contain or not) provided by the feature map. Extraction of fixed-length of each subsequent is achieved using RoI for proposal from feature maps. Further, Feature vector is placed inside of fully connected layer sequence which has two output layer. The system separates gear part numbering. Then, output of regressor produces four real images for reconstructing the proposal location. The Convolutional process Regional Proposal Network (RPN) is used here for designing addition of n x n convolution layers and two relative 1 x 1 convolution layer. Here n=3. In the layer of Conv, The process of mapping for every sliding window is takes place for production of low-dimensional feature by putting input value as 3 x 3 from convolutional feature map. (512-for VGG16) [15]. In next layer, there are two layers Classification layer and regression layer which produces score of objectness and co-ordinate bounding box of every layer.

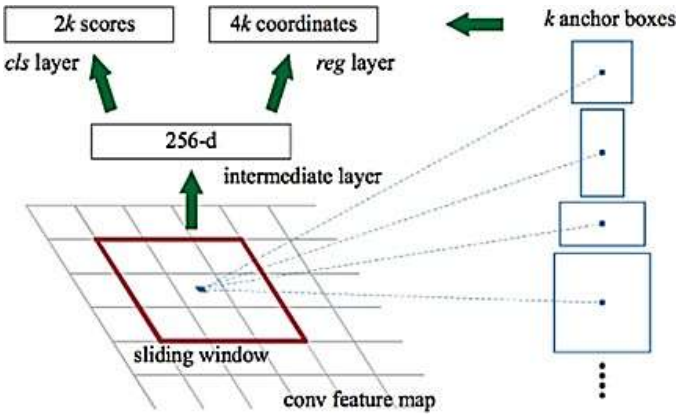

Fig. 3. Faster R-CNN

An anchor-based strategy is utilized for investigating of different scale sizes and object sizes. Each anchor has its own Aspect ratio and specific scale and is focused in its specific window. For each box proportion, we set 3 scale Proportions $(128^2, 256^2, 512^2)$ and three different aspect ratios (1:1, 1:2, 2:1). Here as per the instruction given, ablation experiment is provided on the its effect in further stage [14]. For Classification and regression branches here in this work Scholastic Gradient Distance (SGD) by using end-to-end RPN is implemented. Due to classifying and regressing processes RPN will creates a various sizes with proposal K in every window. So, 2k score of objectness probability is provided by every layer of classification and 4k box coordinates are provided by output regression layer.

3.2 *ROI Pooling Layer*

Here in this part, max pooling is used for conversion of features which are trained in the section where selection coverage of small feature is to be achieved. Here, Application tools of labeling and marks the region which is to be select for obtaining RoI and provides image labelling to each for giving images as per the information. When the image labelling is done, we sort the data and store in xml record. The xml record gives the information about fixed- length partition H*W (H=height; W=width) each H and W are layer hyper-limits that are self-sufficient. For simplicity of planning data, conversion of xml file to .csv plan is carried out.

3.3Classification Problem for Binary

Faster R-CNN consists of two modules, first module is of region based on Deep fully Convolutional Network and another module is detector module. The tool used for detection purpose in Fast-RCNN is Softmax. Softmax can compute the output probability for different objects by following expression:

$$\sigma_i = \frac{\exp(z_i)}{\sum_j^m \exp(z_i)}, i = 1, \dots, m,$$

Where, σ_i = output fully connected layer

z_i = corresponding i-th

i = i-th

m = Number of classes

For this section, part numbering is achieved with the help of two-class softmax by implementing problem of binary classification. Here, sigmoid function is implanted for achieving better performance. The sigmoid has ability to provide solution for problem regarding binary classification ans probability. Probability of binary values can be calculated by using following expression:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Here, for helping to training process multi-task loss is provided. Multi-task loss is defined as - $L = L_{cls} + L_{reg}$, where, L_{cls} is classification defined as log loss over two classes (object is contained or not) and L_{reg} is defined as regression loss over regressor target and it find out the true class for every pixel to provide its location. Rather than Softmax and Multinominal Cross entropy sigmoid cross entropy loss is used for L_{cls}. Additionally, for the detection of gear object sigmoid function is used. Whenever training data in not efficient, this process can be preferred for tasks regarding classification.

IV. IMPLEMENATION

In this research work, train and testing of model is achieved. We train for the mechanical assembly part numbering using Mobile-Net. Here randomly various images are taken then for training it is placed into batches. Then the images are resized according to height to weight ratio. Least dimension is of 600 pixels and maximum dimension is of 1024 pixels. Then, Faster R-CNN is run with 2 classes with the steps number = 10k, threshold limit = 0.2 and rate of learning = 0.00003. Another testing is run using step size = 12k, edge = 0.7 and learning rate = 0.000003.

We tests these model of apparatus part numbering by using 3 datasets. We split all the photos into 3 orders - right part numbering, override part numbering and non-part numbering. 1stcategory contains104 pictures, second grouping contains 100 pictures and second rate class contain 118 pictures.

4.1 Experimentation

We see the experiment setup for gear part numbering classification model using Mobile-Net on TensorFlow framework.

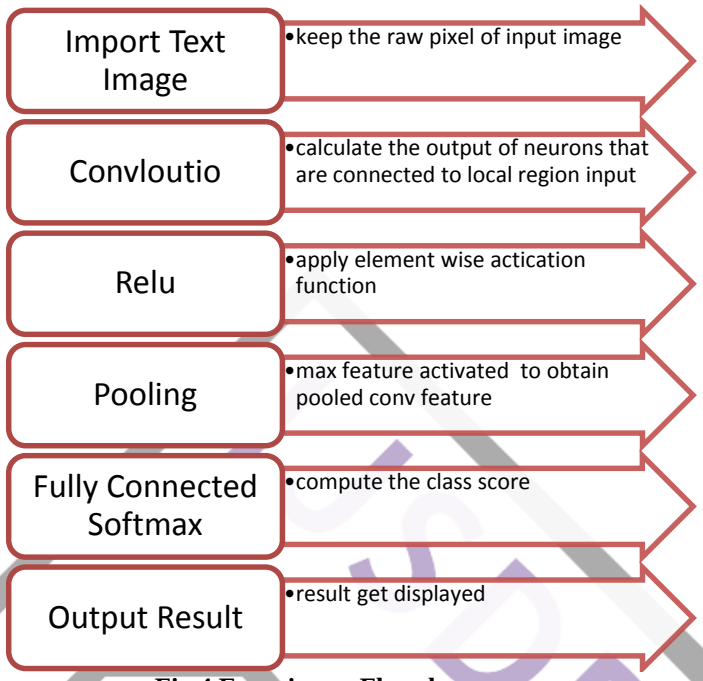

Fig.4.Experiment Flowchart

4.2 Image Setup

Here, Faster R-CNN Part Numbering Detection model is trained on the given dataset. There are 50 pictures for right part numbering, 50 pictures for supplant part numbering and 50 pictures for non-part numbering. We separate data into 3 classes, five star is correct part numbering beneath normal is repeal part numbering and below average class is non-part numbering. Some images of the mechanical assembly part numbering dataset are randomly testified. Here experiment is performed based on Python based MobileNet [16] model TensorFlow [17].

4.3Model Training

Here, collection of various part images is takes place for providing information to train the model. For training purpose every images is sequentially arranged. In the hidden layer, the process of computation and grouping of every image before output layer takes some more timg There are 50 pictures for right part numbering which results in process of learning is saved and can be used further again as there is not any change in lowest layer.

V. RESULTS

By performing this experiment, the result is obtained as 300 images and 95% accuracy. Process of right numbering the picture and uniqueness is significant factor for determining accuracy and precision.
Figure 5, 6 and 7 show the location result for three classifications.

Fig. 5. Detected Correct Part Numbering

Fig. 6. Detected Over-ride Part Numbering

Fig. 7. Detected Non- Part Numbering

Figure 8 show output of series of training which determines accuracy of validation, cross entropy and training precision. The outcomes are outlined as the level of picture named exactly. The haphazardly chosen picture with right marking is approved through visual assessment. Cross entropy examination is the proportion of the nature of the learning cycle pushing forward which are misfortune work S as the quantity of preparing tests are expanded as appeared in figure 8.

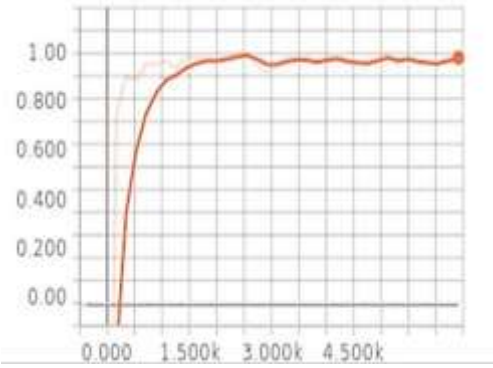

Fig. 8. Cross Entropy Vs. Training Samples

If the train model accuracy is low and the precision is high, it means that network over fitting has occurred. Hence, right feature selection process must be specified for ease of grouping of images. Training samples have to be increased for improving precision and accuracy.

VI. CONCLUSION

In this paper, Faster R-CNN method using Mobile-Net to recognize gear part numbering is presented. In examine, we found that Faster R-CNN can detect objects with high precision and speed with little models in an arrangement progression. Faster R-CNN can create a readied model and request right, override, non-part numbering model by using width multiplier with inspected the rightness level of size and shape. Hence, We conclude that if we need consistent application with high precision Faster R- CNN system is a good choice.

REFERENCES

[1] J. Yang and G. Yang, "Modified convolutional neural network based on dropout and the stochastic gradient descent optimizer," Algorithms, vol. 11, no. 3, pp. 28–43, 2018.

[2] J. Yang, W. Sheng, and G. Yang, "Dynamic gesture recognition algorithm based on ROI and CNN for social robots," in 2018 13th World Congress on Intelligent Control and Automation (WCICA), pp. 389–394, Changsha, China, China, July 2018.

[3] Y. Xu, C. Yang, J. Zhong, N. Wang, and L. Zhao, "Robot teaching by teleoperation based on visual interaction and extreme learning machine," Neurocomputing, vol. 275, pp. 2093– 2103, 2018.

[4] M. Trobe and M. D. Burke, "The molecular industrial revolution: automated synthesis of small molecules," Angewandte Chemie International Edition, vol. 57, no. 16, pp. 4192–4214, 2018.

[5] J. Yang, S. Li, Z. Gao, Z. Wang, and W. Liu, "Real-time recognition method for 0.8 cm darning needles and KR22 bearings based on convolution neural networks and data increase," Applied Sciences, vol. 8, no. 10, p. 1857, 2018.

[6] S. Jana and S. Basak, "Automatic Fruit Recognition from Natural Images using Color and Texture Features," Devices Integr. Circuit, pp. 620–624, 2017.

[7] C. Szegedy, A. Toshev, and D. Erhan. Deep neural networks for object detection. In NIPS, 2013.

[8] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. In ICLR, 2014.

[9] C. Szegedy, S. Reed, D. Erhan, and D. Anguelov. Scalable, high-quality object detection. arXiv:1412.1441v2, 2015

[10] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In CVPR, 2014

[11] Rismiyati and S. N. Azhari, "Convolutional Neural Network implementation for image-based Salak sortation," in Proceedings - 2016 2nd International Conference on Science and TechnologyComputer, ICST 2016, 2017, pp. 77–82.

[12] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in neural information processing systems, 2012.

[13] A. Karpathy, October 2017. [Online]. Available: http://cs231n.github.io/convolutional-networks/.

[14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 1096, pp. 1137–1149, 2015.

[15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Int. Conf. Learn. Represent., 2015.

[16] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," ArXiv, 2017.

[17] P. Goldsborough, "A Tour of TensorFlow," ArXiv, 2016.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Arxiv.Org, 2015.