https://doi.org/10.2352/ISSN.2470-1173.2017.10.IMAWM-164

Detection and characterization of Coordinate Measuring Machine (CMM) probes using deep networks for improved quality assurance of machine parts.

Binu M. Nair, Vidur Prasad, and Nilesh Powar; University of Dayton Research Institute, Dayton, OH

Abstract

We propose the use of a deep network to detect, segment and characterize a Coordinate Measuring Machine (CMM) probe used in measuring various machine parts. Our motivation is to accelerate the time taken for an operator to input various parameters of a CMM probe into the system, so that delay in quality assurance of machine parts can be negated. Using imagery from a high resolution EO sensor, we design a probe recognition and characterization framework which can segment probe regions, classify various probe-region proposals into generic or specific probe components, and estimate the various configuration parameters of the probe. In order to measure a specific machine part, an operator provides the CMM machine with an image of an assembled probe. This end-to-end deep network-based framework will then generate configuration parameters suitable for the measurement task. Since the number of machine parts are in the order of thousands, the probe can have multiple configurations. In this work, we do extensive analysis on a probe dataset captured in our lab and evaluate two main aspects of the framework: its ability to segment regions, and classify those regions as probe components.

Introduction

Vision-based automation has become the de-facto standard in various processes in manufacturing sectors such as aerospace and automotive. One of the main processes in the manufacturing pipeline is the quality assurance and final inspection to verify a part dimensions using Coordinate Measuring Machines (CMM). However, the inspection processes in place are not adaptable to different types of manufacturing parts. These parts are in the order of thousands and each type requires a different configuration of a CMM measurement probe. For every new part, the operator needs to assemble a probe and manually enter the configuration parameters onto the CMM system. Since this process is time consuming and tedious especially in older machines, we require a data-driven method which can model and characterize the probe and its configuration from high-resolution imagery.

A CMM probe consists of $3-4$ main components such as the probe head, probe body, probe extension (optional) and stylus head. Different variants of these parts are used to create a unique probe configuration. In order to automatically recognize a probe, its characteristics such as shape, size, the structure and other dimensions should be extracted as features by the vision system. These extracted features should then be compared to the corresponding known probe configuration within a pre-defined database. The proposed framework in this work addresses this issue of extracting relevant feature which are suitable to discriminate between very similar configurations. Sample images of the simple setup for probe image capture and the sample types of the components are shown in Figure 1 and Figure 2 respectively.

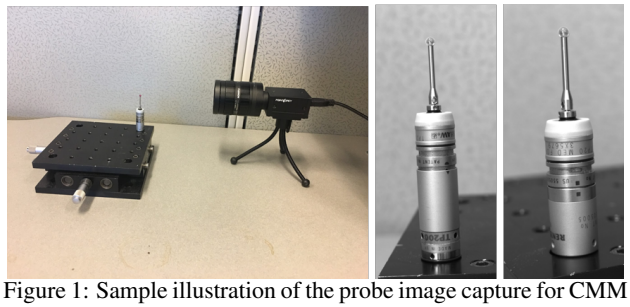

Figure 1: Sample illustration of the probe image capture for CMM machines.

Figure 2: Examples of some Renishaw probes used in CMM machines and the various stylus provided.

Objective

The aim of this project is to recognize and characterize a CMM probe configuration from a high resolution EO sensor so that accurate estimates of the system-specific probe parameters are generated. Some of the technical challenges that we will face are the following:

- Image data collection and ground truth annotation of large number of probe configurations for various manufacturing parts.
- Development of deep learning/data-driven models to capture fine features of the probe from high resolution imagery for accurate recognition among similar classes.

The main objective is to explore and develop a vision-based algorithm to recognize various configurations of a CMM (Coordinate Measuring Machine) probe used in measuring machine parts. In this work, we focus on 12 specific types of probe configurations which are similar to each other. Each configuration is obtained by selecting a probe stylus, probe body (extension) and a probe head from a set of 3 different types of probe stylus (*stylus*1, *stylus*2, *stylus*3), two different types of probe body extensions (*body*1, *body*2), and two different probe heads (*head*1, *head*2). One of the main challenges in probe recognition is that the various classes

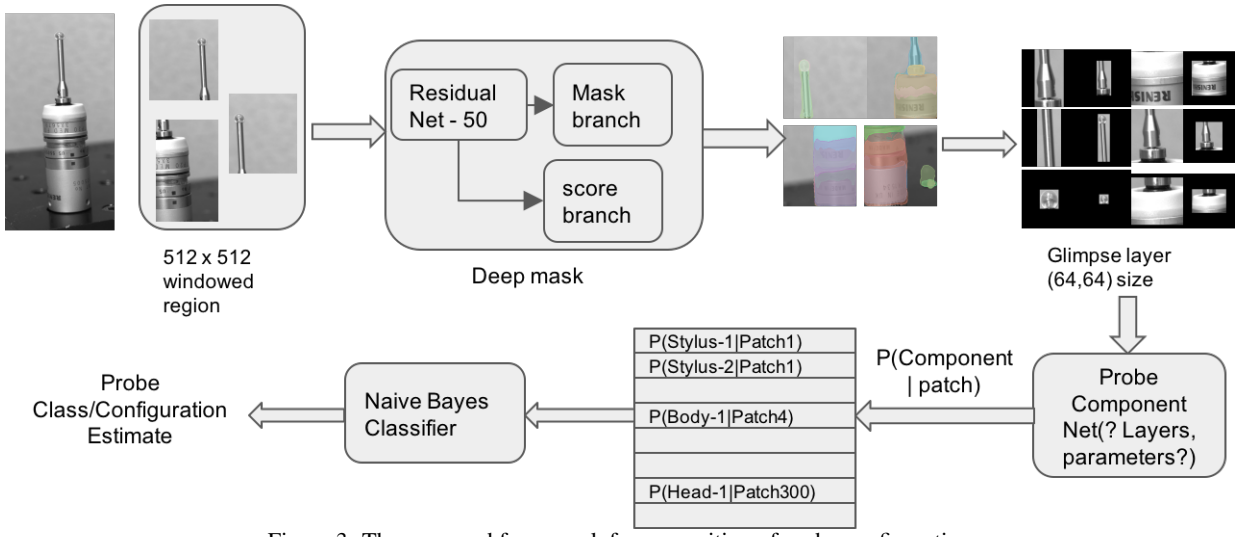

Figure 3: The proposed framework for recognition of probe configuration

are too similar to each other for a standard object classification problem. The attributes that need to be extracted are the subtle circular/straight line features in the probe body, the printed writing on the head, and the length/size of the stylus. To capture these subtle features effectively, we use a high resolution Point Grey 12MP monochrome EO sensor. However, due to high resolution of the image, current state of the vision algorithms cannot be applied directly as this would involve downsizing the image which would eliminate the subtle features necessary for the probe recognition. Our proposed novel algorithm leverages various deep learning models to extract probe-specific features from local regions, which can then be applied to a naive Bayes or Neyman-Pearson classifier to recognize the complete probe.

Our motivation for this algorithm is to combine a generic segmentation capability with a weak probe component detection criteria to retain regions containing specific parts of the probe within the image. Once we obtain the localized regions of the probe, we classify the various probe regions as a body, head, and extension (stylus). The novelty in our proposed approach is the use of a segmentation techniques to obtain certain regions of the probe and subsequently use deep residual neural networks to recognize generic or specific probe components from high resolution imagery. This is in contrast to state of the art CNN networks which are targeted for object recognition among a large number of classes with very distinct characteristics.

Related Work

Machine vision algorithms are widely used with CMM technology for several dimensioning tasks. Most of the time vision-based systems provided by commercial entities such as Cognex is used to improve the operation of a CMM. In traditional approaches, the vision camera is mainly used as a feedback to the CMMs motion controller to accurately guide the probe [6]. Another advantage of using the vision system is to register 2-D point from the camera and 3-D points from CMM and these points would be helpful for calibration. Combining a CMM with a vision system helps in reverse engineering of freeform surfaces accurately [1]. It also helps in reconstructing CAD model of objects of complex geometry with high accuracy. The fusion of optical and mechanical sensors makes it unique and is highly automated. A research group [12] has demonstrated usage of 3D active vision sensor in combination CMM equipped with motorized touch probe for part localization. In all the research above the common theme was to improve the CMM process and make it more accurate. Our proposed effort takes this concept to a whole new level where we would leverage state of the art models to enhance and increase the automation process without the need for extra calibration and environment setup. Our research focuses on using deep learning methods to effectively detect, segment and characterize a CMM probe before further evaluation and analysis.

Object segmentation is a critical portion of any classification or recognition problem. Traditionally this problem has been solved with various background modeling techniques including background subtraction, Gaussian Mixture Model (GMM), and statistical models. With the advent of deep learning, traditional techniques are being replaced and outperformed in many learning tasks. Selective segmentation[14], in conjunction with Convolutional neural networks have been combined to form region-centric CNNs (R-CNN) networks for object detection and classification [2, 10] . Here, region proposals are provided to the network to obtain efficient features for object classification. These features are then classified using SVMs or fully-connected layers to recognize the object present in the region. Convolutional neural networks have also been used in optimizing the masks surrounding the regions, and one such work was done by Pinheiro et al [7] where higher level layers of the VGG-Net [13] was replaced with two convolutional layer branches, one which outputs a mask, and the other a score reflecting the presence of an object. This approach is more generic as it could provide region proposals which has higher chance of containing an object.

Methodology

The proposed approach for probe configuration recognition from a high-resolution image is to segment various interesting regions corresponding to the probe, classify those individuals regions as probe components, and then use the various probe component detections to estimate the complete configuration. The framework is a combination of three algorithms: probe region

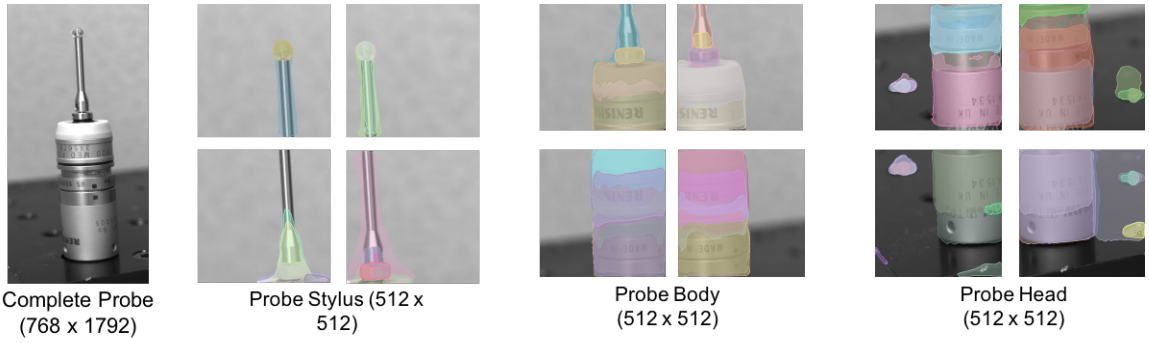

Figure 4: Illustration of deep mask segmentation on sections of the probe.

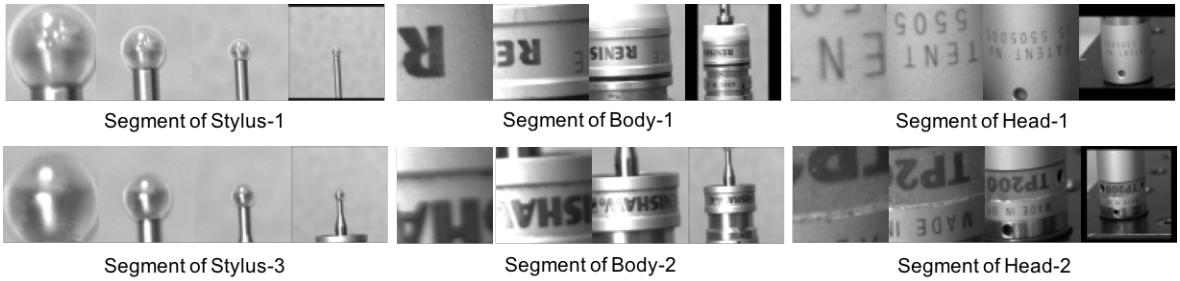

Figure 5: Illustration of glimpses extracted from various region segments. From each region, we obtain $4 \times 64 \times 64$ glimpses.

segmentation, probe region classification, and probe characterization. As shown in Figure 3, the probe region segmentation algorithm leverages a pre-trained deep neural network model known as deep mask [7, 8] where region proposals and their corresponding masks can be obtained from an image. Using the centroids of these region proposals, glimpses of such regions can be extracted at different resolutions. These glimpses of the probe regions are then classified using a trained probe component network using Residual Nets [3]. We will now explain each of these algorithms and its usage in the following sub-sections.

Probe Region Segmentation

One of the first tasks in this approach is to segment relevant regions of the probe present in the high resolution image. Selective segmentation [14] is an effective way to obtain regions proposals where these patches can be fed to a classifier for object recognition. However, we employ the use of a more accurate segmentation technique known as Deep Mask [7] which provides region proposals which has high probability of containing a possible interesting object. The Deep Mask is a deep convolutional neural network where it uses the first few trained layers of a VGG-Net [13] in the main branch of the network. The input to the main branch is an image of size $3 \times 224 \times 224$ and output is a 3D tensor of size $512 \times 14 \times 14$. From the main branch, this 3D tensor is fed to two parallel branches, named as mask branch, and score branch. The mask branch consists of a convolutional layer $(conv(1,1))$, fully connected layer, and upsampling to generate a binary mask of the possible object. The score branch consists of a convolutional layer $(conv(2,2))$, a fully connected layer and an output node which provides a confidence score which indicates how much of the object is fully contained in the estimated mask. The learning of the network uses a binary logistic regression loss which uses the pixel-by-pixel inference of the mask as well as the score value.

In our proposed model, we leverage the trained model of the deep mask (provided as open-source) and leverage the scores obtained when applied on the high resolution probe image. Due to memory constraints, we first divide the high resolution image of size 768×1792 into overlapping 512×512 blocks. For each block, we apply the deep mask to obtain k region proposals as shown in Figure 4. The first two rows of 512×512 blocks focuses on regions of the stylus, the next two focuses on the probe body, and the last two focuses on the probe head. The motivation behind this division of the large image is to obtain appropriate representations of the regions of the probe components at high resolution. We apply a filter to those regions by using the score obtained from those regions. Setting the score at 0.7, we obtain an average of 100 regions proposals from the whole image.

Now, to appropriately represent these regions at different resolutions, we make use of a neural network module known as glimpse. The glimpse layer has been very effective in attention networks [5] where the glimpse module provides multi-resolution multi-scale aspects of an object from a high resolution image at different instants of time and at different locations. In this work, we do not use recurrent models to capture the temporal transitions of the glimpse but is used only to get the multi-resolution multi-scale aspects of the probe component region. These glimpses of the probe region are shown in Figure 5

Probe Component Net : Classification of probe regions.

From the glimpses obtained, we can now train and model a neural network to classify each glimpse of a probe region into probe components. In this work, we focus on two types : one is a generic probe component classification which tells if the glimpse is of a stylus, body or head of the probe; the other is a more specific probe component classification which tells if the glimpse is a particular stylus, body or head from a set of styluses, bodies

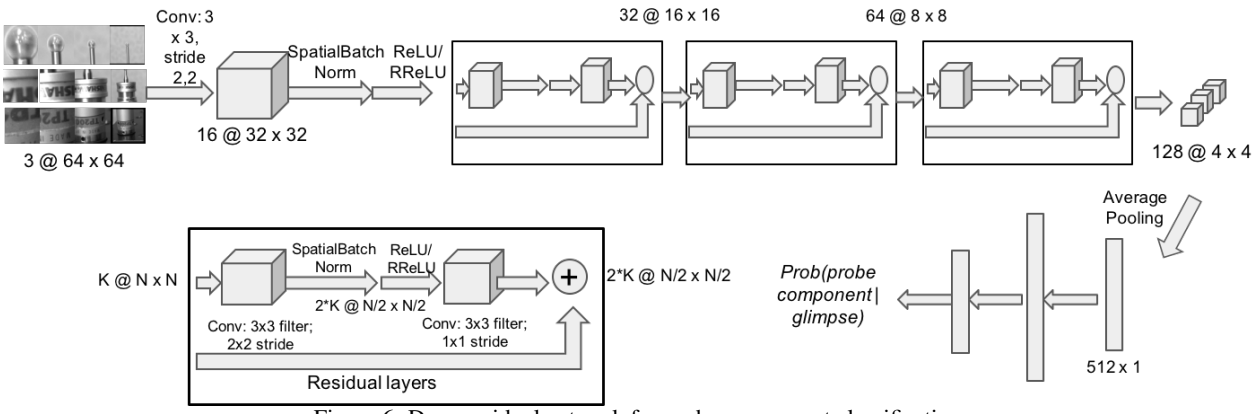

Figure 6: Deep residual network for probe component classification

and heads of the probe. For the generic probe component classification problem, we explore pre-trained Residual Net models as feature extractors for each glimpse image. The CNN features extracted from each glimpse is then fed to a fully-connected layer for a generic probe component classification. In this classification problem, we extract glimpses of size $3 \times 224 \times 224$ from each segmented region of the probe. We consider only two scales while computing the glimpse of a region. For the specific probe component classification problem, we explore a custom deep residual neural network which mirrors the Residual-Net-18 [3] architecture but with the input as a glimpse of much smaller size $3 \times 64 \times 64$. In this problem, we extract 4 scales of the glimpse. The only motivation for the small size and large number of scales of the glimpse is to have large number of samples and reduce memory consumption in training a deep network. Our objective to explore how well the network functions as a specific probe classification network. This network is illustrated in Figure 6.

The architecture of the network that we use in the specific probe component classification problem is as follows : 1 convolution layer, 1 spatial batch normalization layer, 3 residual layers as defined in [3], 1 average pooling layer, and a fully connected layer. We train this network using Stochastic gradient optimization for weights and biases with 7 different class labels representing the probe components (*stylus*1, *stylus*2, *stylus*3, *body*1, *body*2, *head*1, *head*2). To minimize the effect of over-fitting, we use a Randomized Rectified Linear units (RRelu) as the activation function. After training, what we expect is to obtain a class label and a confidence score. In our approach, this confidence score will reflect the probability of a specific probe component present in a glimpse of a segmented region. The complete probe characterization from the high resolution image will accumulate these probability scores computed from every glimpse of each segmented region in a Naive Bayes or Neyman Pearson classifier [9]. The probe characterization problem is currently work in progress.

Results and Analysis

We explore the capabilities of the proposed probe segmentation and classification framework on a data-set containing high resolution images of the probe. This data-set is captured from a 12 MP Point grey machine vision camera with custom region of interest set to focus only on the probe and the size set to 768×1792. Sample images of the probe and its various

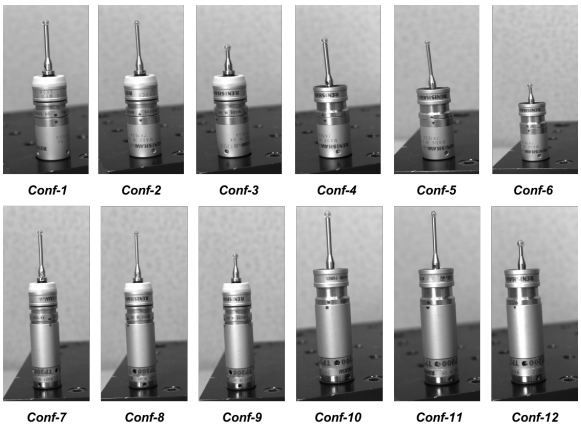

Figure 7: Illustration of various probe configurations

configurations are shown in Figure 7. There are 12 different configurations with each obtained by selecting a particular combination of the probe components such as the stylus, body (extension) and head. There are three different styluses (*stylus*1, *stylus*2, *stylus*3) which vary mostly in size and length, two different probe body (*body*1, *body*2) which vary in the concentric features, and the probe head (*head*1, *head*2) which vary in the marking, and its length. For this data-set, we captured 100 images of each configuration, leading to a total of 1200. The general variation in the image set is the small degree of rotation of the probe along its axis, its position in the image with respect to the x,y and z (depth) axis, and small changes in the background which are relatively uniform. This entire image capture setup used a micrometer stage where the probe could be placed. Some of the main challenges in this data-set are the following: high similarity (low variance) between the various probe configurations; high similarity between specific probe components; and high inter-class variance of the probe classes.

Analysis-1: Effectiveness of probe features and segmentation

We first qualitatively evaluate the effectiveness of capturing features which can potentially help in discriminating similar looking probe configurations. Traditional hand-engineered features such as SIFT, SURF or BRIEF have been used exten-

sively to solve image recognition tasks in robotic vision and automated guidance systems such as automated bin picking in manufacturing. Similar features can also be obtained by applying a pre-trained convolutional neural network model such as AlexNet [4]. An illustration of hand-engineered and CNN-based features on sample probe images are shown in Figure 8. We can see that the specific attributes of the probe such as elliptical features, line features and point features can be detected using both hand-engineered and CNN features, thereby providing us a cue for discriminating between two different probe configurations. Our motivation then for using a deep neural network is to obtain relevant and optimal features for probe recognition and characterization where the feature detectors are governed by data-driven techniques. We have also obtained some preliminary results showing

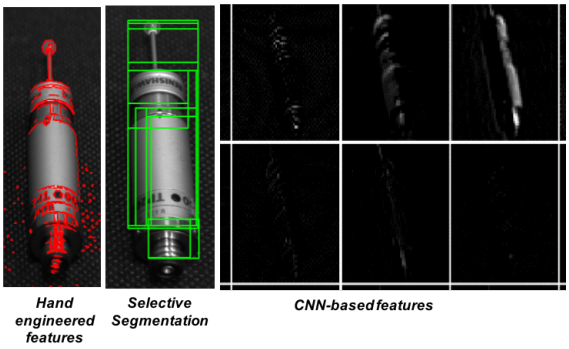

Figure 8: Features and region proposal illustration for the probe using hand-engineered techniques, selective segmentation [14] and deep mask segmentation [7].

the application of selective segmentation [14] on a sample image of the probe. We see that selective segmentation with probe-specific constraints are able to obtain region proposals which contain parts of the probe and more useful to characterize the probe. However, we use deep mask segmentation technique as it can provide more accurate region proposals than the selective segmentation technique.

Analysis-2 : Probe component type classification

In this analysis, we explore some common approaches towards learning and fine-tuning deep networks for probe component type classification where there are three classes *stylus*, *body*, and *head*. In our first approach, we use a pre-trained model such as $Residual\,Net-18$ which has 18 layers trained already on an ImageNet [11] dataset. This model takes an input image size of $3\times224\times224$ and outputs a 512 element features. So, in our framework, we capture glimpses of 224×224 on a segmented region of the probe with two scales and feed each scale to the pre-trained network to obtain CNN-based features. From the dataset, the total number of glimpses extracted and segmented for fine-tuning a model is around $217,000$ images. As shown in Table 1, there are three different fine-tuned models which varies in the architecture of the fully connected layers. The activation functions of these layers use a *ReLU* units with the $LogSoftMax$ layer at the output. We evaluate the various models by computing the precision ($\frac{TP}{TP+FP}$), recall ($\frac{TP}{TP+FN}$) and accuracy.

As shown, the precision, recall and accuracy of the model increases by about 5% rate with increase in complexity of the

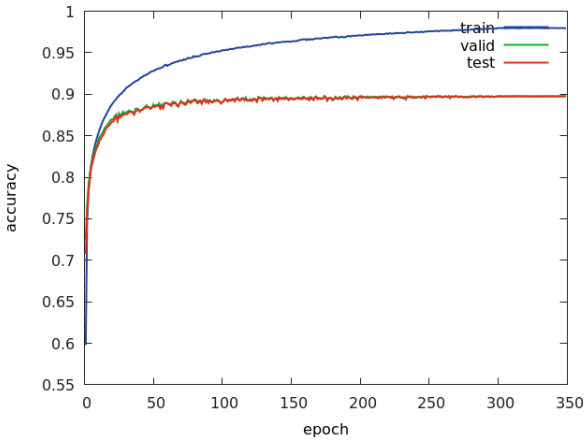

Figure 9: Accuracy curve for train, test and validation set obtained from the pre-trained model $512-1000-300-30-3$.

architecture of the fully connected layers. The precision of the stylus increases from 80.17% to 84.96% while its recall increases from 81.66% to 86.74%. Similarly, the precision of the body increases from 81.98% to 87.80% while the recall increases from 81.78% to 87.63%. For the probe head, we obtain a high precision of around 90.69% which increases to 95.06% while the recall increases from 89.86% to 94%. The overall accuracy increases from 84.65% to 89.72%. These results prove that using a pre-trained model (trained on a much larger and diverse dataset) is useful in recognizing and classifying regions on probe images captured using a different sensor in a different environment. The low-level feature maps which capture the relationships between edges, corners in the image are similar to those found on the probe. The higher level abstracted features which form building blocks of a generic ImageNet object class is applicable for representing similar building blocks of a image of a probe captured in a different environment. However, from the Figure 9, we see that the pre-trained models slightly overfit the data but not to a huge degree. This is due to the lack of refinement and fine-tuning of the higher abstracted feature layers (the convolution layers) towards the probe images. Therefore, we explore the option of training a custom model (ProbeComponentNet) from scratch using larger dataset but with smaller image/glimpse size of 64×64.

Analysis-3: Probe component classifications

In this analysis, we explore the state of the art deep residual neural network architecture for directly classifying specific probe components which make up the entire probe configuration. There are seven classes of components in this dataset, specifically *stylus*1, *stylus*2, *stylus*3, *body*1, *body*2, *head*1, and *head*2. In this approach, we train a custom deep residual network (known as Probe Component Net) which takes in $3\times64\times64$ image. Therefore, in our framework, each segment generates 4 glimpses of size $3\times64\times64$ resulting in a total of $490,000$ glimpses from this dataset. The motivation behind this approach is to explore the possibilities of directly associating a glimpse of a particular probe segment/patch to one of the specific probe components. The end goal is to obtain probability of each patch/segment of a probe belonging to one of the components. This will make the complete inference of a probe configuration easier as we can now

Conf	Input glimpse size	Stylus	Body	Head
Pre-trained ResNet-18 - 512-100-3	224×224	80.17%	81.98%	90.69%
Pre-trained ResNet-18 - 512-300-30-3	224×224	83.51%	86.90%	93.79%
Pre-trained ResNet-18 - 512-1000-300-10-3	224×224	84.96%	87.80%	95.06%
ProbeComponentNet	64×64			

(a) Precision Scores

Conf	Input glimpse size	Stylus	Body	Head
Pre-trained ResNet-18 - 512-100-3	224×224	81.66%	81.78%	89.86%
Pre-trained ResNet-18 - 512-300-30-3	224×224	85.80%	85.92%	93.34%
Pre-trained ResNet-18 - 512-1000-300-10-3	224×224	86.74%	87.63%	94%
ProbeComponentNet	64×64			

(b) Recall scores

Conf	Input glimpse size	Overall Accuracy
Pre-trained ResNet-18 - 512-100-3	224×224	84.65%
Pre-trained ResNet-18 - 512-300-30-3	224×224	88.56%
Pre-trained ResNet-18 - 512-1000-300-10-3	224×224	89.72%
ProbeComponentNet	64×64	

(c) Accuracy

Table 1: Precision, recall and accuracy per probe component class for various configurations using a common pre-trained Residual network, and changing the fully connected network for classification.

have each segmented region of a probe associated with a probe component. By the association of multiple probes regions with the corresponding probe component labels, the complete probe configuration from a high resolution image can be estimated. The performance of the Probe Component Net is summarized in Table 2.

In Table 2a, we see that the precision and recall of the task of identifying specific probe components is much less than what was observed for generic probe component classification. This can be attributed to the fact that the specific probe component classes are very similar to each other. One example is the tuple of categories (*stylus*1, *stylus*2, *stylus*3) where the only difference is in the diameter of the ball bearing at the end of stylus and the length of the stylus. Since we are using only a gray-scale camera, we lose the color information which can be a key component in distinguishing between various stylus components. The precision of a stylus-type category varies from 64% to 68% while the recall varies from 63% to 67%. However, the category tuple (*body*1, *body*2) have a higher precision and recall rate ranging from 70% to 79%. Moreover, the category tuple (*head*1, *head*2) have even higher precision and recall ranging from 75%to 84%. This is mainly because there are more variations in the *body* and *head* component type of the probe such as variation of lines, edges, elliptical features, small markings etc. Overall, the accuracy of the specific probe component detection peaks only at 73.33% and is mainly due to the low accuracies obtained by the stylus categories.

In Table 2b, we provide the true positive rate, the false positive rate and the false negative rate in the form of confusion matrix. Here, the rows correspond to true class while the columns refer to the estimated class by the network. The true positive rates for each category lie along the diagonal, the false negative rates fall only the rows, and the false positive rates fall along the columns. For the category tuple (*stylus*1, *stylus*2, *stylus*3) , the true positive rate lies between 48% to 49% with the false positive and negative rates spread across predominately along the other *stylus* and *body* categories. We see that around $4\% - 8\%$ of *stylus*1 segments falls in the (*stylus*2, *stylus*3, *body*1, *body*2) categories contributing to a high false negative rate of 21.86%. Also, regions belonging to other *stylus* and *body* component categories fall in the *stylus*1 category which contributes to a high false positive rate of 22.54%. Similar is the case with *stylus*2 and *stylus*3 categories where the true positive rate is $47 - 48\%$ while the false positive and false negative rate is around $25\% - 27\%$ and $16 - 18\%$ respectively. This large false positive and false negative rates can be attributed to two main factors: similarities within the *stylus* component categories; and the other is noisy labeling of the data where some samples of *body* component categories are present in the stylus and vice-versa. The former can be resolved by obtaining higher resolution glimpses and at large number of scales to capture the variations in the *stylus* component categories. However, the latter issue of noisy labeling can be attributed to the weak-labeling criteria in setting up the ground truth. This can be see in the performance of the network on the *body* categories where they achieve high true positive rates of $55\% - 59\%$ but the false positive and false negative rates are spread out across all other categories. When we consider the *head* component categories, the false positive rates and false negative rates are concentrated more along the other *head* component category and *body* component category. Here, a true positive rate of $62\% - 68\%$ is obtained with false positive rand false negative rates of 18% and 15% respectively. Another indication of weakly labeled and noisy ground truth is the indication of over-fitting of the model in spite of sufficient training data samples and necessary regularization techniques such as usage of *RRelu* units, spatial batch normalization techniques and drop out mechanisms. This fact is illustrated in the accuracy curves shown in Figure 10.

Conclusions/Future Work

In this work, we have explored the use of state of the art deep learning approaches to characterize a CMM probe configuration

Stats	Stylus-1	Stylus-2	Stylus-3	body-1	body-2	head-1	head-2	Overall
Precision	68.63%	64.81%	64.81%	76.37%	70.74%	78.66%	77.82%	
Recall	63.67%	63.78%	66.59%	67.44%	79.02%	75.23%	83.58%	
Accuracy								73.33%

(a) Precision, Recall and Accuracy scores for specific probe component classification.

True Class/Estimated Class	Stylus-1	Stylus-2	Stylus-3	Body-1	Body-2	Head-1	Head-2
Stylus-1	49.32%	7.46%	6.43%	3.76%	4.21%	0.22%	0.44%
Stylus-2	6.65%	47.37%	4.72%	2.51%	4.75%	0.21%	0.32%
Stylus-3	4.18%	3.67%	48.91%	1.78%	3.55%	0.13%	0.48%
Body-1	7.14%	8.39%	8.64%	55.80%	5.06%	4.32%	4.01%
Body-2	4.15%	5.76%	6.13%	0.022%	59.49%	4.22%	3.20%
Head-1	0.21%	0.32%	0.29%	4.49%	4.88%	62.48%	10.79%
Head-2	0.22%	0.13%	0.33%	2.51%	2.27%	7.85%	67.50%

(b) Confusion matrix reflecting percentage of true positives (along the diagonal), false negatives (along rows), and false positives (along columns)

Table 2: Performance of Probe Net component classification with RRelu (Randomized leaky rectified units) units and glimpse size of $3 \times 64 \times 64$

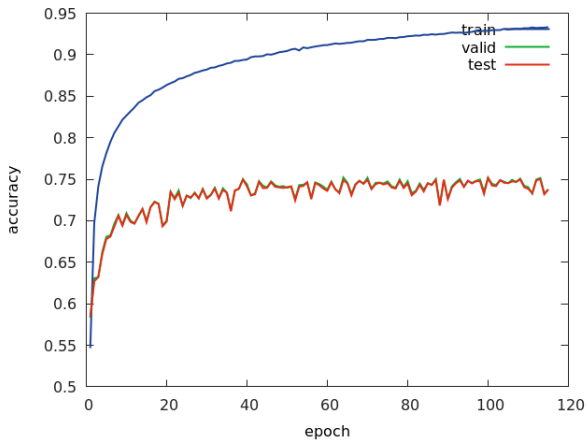

Figure 10: Accuracy curve from Probe Component Net

from high resolution imagery. This led to the development of a novel vision-based CMM probe recognition framework which can reduce cost and time during machine part inspection. In this approach, we leveraged the use of segmentation techniques to obtain certain regions of the probe, and develop a deep residual neural network model to associate the region with a generic probe component type and a specific probe component. In our analysis, we have obtained a fairly good performance with the generic probe component type classification and fairly moderate performance with the specific probe component classification. This paves way to another simple statistical approach which can leverage the confidence scores or probability scores of various regions in the image to estimate the complete probe configuration.

As part of future work, we will address the issues of overfitting, and miss-classification through a more accurate ground truth labeling, and introduction of more imagery under different background conditions. We will also explore the use of CAD models of the various probe components in our framework to obtain a match between the segmented regions and the corresponding 3D models for higher accuracy and robustness.

Acknowledgements

This work is supported by the grant awarded by Edison Advanced Manufacturing Program from the State of Ohio (Grant No: AMP 16-06). We would like to thank Michael'O Connor (Head of UDRI Sensor Systems Division), and Rob Gillen (program manager, UDRI FastLane Division) for their support in this research work. We also would like to appreciate Mary Miller (Technical proposal writer, UDRI FastLane Divison), Cory Bucksar (Software Engineer), and Philip Ratermann (Head of UDRI FastLane Division) for their support of promoting this initiative to small businesses in the Ohio region. And lastly, we would like to thank Advanced Industrial Measurement Systems (AIMS) for providing the CMM probes necessary for this research.

References

[1] V Carbone, M Carocci, E Savio, G Sansoni, and L De Chiffre. Combination of a vision system and a coordinate measuring machine for the reverse engineering of freeform surfaces. *The International Journal of Advanced Manufacturing Technology*, 17(4):263–271, 2001.

[2] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 38(1):142–158, 2016.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, 2016.

[4] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[5] Volodymyr Mnih, Nicolas Heess, Alex Graves, et al. Recurrent models of visual attention. In *Advances in neural information processing systems*, pages 2204–2212, 2014.

[6] Marilyn Nashman, Billibon Yoshimi, Tsai Hong Hong, William G Rippey, and Martin Herman. Unique sensor fusion system for coordinate-measuring machine tasks. In *Intelligent Systems & Advanced Manufacturing*, pages 145–156. International Society for Optics and Photonics, 1997.

[7] Pedro O. Pinheiro, Ronan Collobert, and Piotr Dollár. Learning to

segment object candidates. In *Proceedings of the 28th International Conference on Neural Information Processing Systems*, NIPS'15, pages 1990–1998, Cambridge, MA, USA, 2015. MIT Press.

[8] Pedro O. Pinheiro, Tsung-Yi Lin, Ronan Collobert, and Piotr Dollár. *Learning to Refine Object Segments*, pages 75–91. Springer International Publishing, Cham, 2016.

[9] H. Vincent Poor. *An introduction to signal detection and estimation*. Springer texts in electrical engineering. Springer-Verlag, New York, 1988. Adapted from a one semester graduate course taught at the University of Illinois.

[10] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.

[11] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.

[12] Tzung-Sz Shen, Jianbing Huang, and Chia-Hsiang Menq. Multiple-sensor integration for rapid and high-precision coordinate metrology. *IEEE/ASME Transactions on mechatronics*, 5(2):110–121, 2000.

[13] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[14] Koen EA Van de Sande, Jasper RR Uijlings, Theo Gevers, and Arnold WM Smeulders. Segmentation as selective search for object recognition. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1879–1886. IEEE, 2011.

Author Biography

Binu Nair *is a Research Engineer with University of Dayton Research Institute and specializing in computer vision and machine learning for video analysis, text analysis and machine vision. He graduated with a PhD in Electrical Engineering from the University of Dayton and did his postdoctoral research for University of Nevada Las Vegas. His research interests are in activity modeling, video semantic description, and object detection for surveillance and machine vision applications. His current work involves object detection using convolutional neural networks, topic mining for discovering device vulnerabilities, and activity modeling using temporal networks.*

Vidur Prasad *is a student at the University of Michigan pursuing a dual major in Computer Science and Economics. He has worked at the University of Dayton Research Institute and at Aptima, Inc. on a variety of Computer Vision and Machine Learning applications. He has presented his work at the Intel International Science and Engineering Fair. Moving forward, he would like to continue his work in data analytics through a career in Management Consulting.*

Nilesh U. Powar *received his Bachelor of Engineering in Electronics from University of Bombay, India in 1999, and a M.S. in Computer Engineering from Wright State University, Dayton, OH in 2002 and a PhD degree in engineering from University of Dayton, Dayton, OH in 2013. He is currently a senior researcher at University of Dayton Research Institute (UDRI) and had joined UDRI in 2003. His research interests include image processing, statistical pattern recognition and system integration.*