Automatic Recognition System For Mechanical Hand Tools Using Convolutional Neural Networks

Ghazi Alnowaini
Department of
Mechatronic Engineering
Taiz University
Taiz, Yemen
Dr.ghazi.aln@gmail.com

Hesham Al-khateeb
Department of
Mechatronic Engineering
Taiz University
Taiz, Yemen
HeshamAlkhateeb88@gmail.com

Amer Al-basser
Department of
Mechatronic Engineering
Taiz University
Taiz, Yemen
AmerAlbasser@gmail.com

***Abstract*— the identification of high-precision mechanical tools is an important problem faces mechanical engineers. Indeed, hundreds of different instruments are typically used for one task. In several cases and multiple workshop environments each tool will be used. Mechanical engineering is a practical field that need different mechanical tools during workshop work. These tools such as Wrench, hammer, toolbox, Gasoline Can, and pebble have different size and style. During work time of the mechanical engineers, they need these tools frequently and the identification process of these tools is a difficult task for automated system. In this paper, an automated recognition system for mechanical tools using convolution neural networks. A CNN-based model is discussed using four versions of residual network classifiers; ResNet-18, ResNet-34, ResNet-50 and ResNet-152. This model can be integrated with a robot to give it the ability to recognize the specific mechanical tool and deliver it to the mechanical engineer. The feasibility of this method is illustrated in the achieved results. The obtained results are very promising to be used in practical use. In term of testing accuracy, the results achieved 84%, 85%, 86% and 87% for ResNet18, ResNet-34, ResNet-50 and ResNet-152 respectively.**

Keywords—Mechanical Tools, Hand Tools, Deep Learning, CNN, ResNet.

I. INTRODUCTION

Mechanical tools are a set of tools that a mechanic needs while performing his work. Mechanics use many mechanicals tools of multiple size and functions. There are many innovative ideas through the use of artificial intelligence to facilitate the selection process and obtain the appropriate tool to perform the process that requires the suitable tool.

The human tendency to see regularities in observations is the pattern recognition. Early on, scientists and engineers have attempted, either partly or in its whole, to mimic the potential by mechanical means.

Pattern recognition (PR) can be identified as a classification mechanism that aims at extracting patterns from a data set and categorizing them into different classes [1]. It is also possible to classify the system by which machines can analyze the context, discern patterns of interest from their history, so as to make accurate and feasible decisions concerning pattern categories [2]. Over the years, there have been many PR concepts offered. Pattern Identification is relevant because in many functional issues it is required. The human beings do this very well, but it is impossible to get a machine to do the same.

Any provided data can also be evaluated and a scene in fact that is useful for executing those tasks is the primary objective of a pattern recognition system. The identification of patterns requires several methods that guide the creation of multiple applications in various files.

It is common practice to mark the wrenches by stamping, to differentiate between the varying shapes, which desirably creates an impact on the instrument. When size marks are clearly visible, this technique is satisfactory. However, in certain cases, when a mechanic works on an engine, for instance, he may try to hit the tool without focusing on scale marking. It is also a matter for the invention to make it possible to distinguish work parts and instruments of equal size by size. A similar purpose is to make individual socket, box end and open end wrenches of about the same size openings easier to spot.

Another innovation purpose is to allow the consumer to rapidly pick a tool of the desired dimension without seeing either the tool's opening or the scale.

Pattern recognition field have become an important approach to robotics, industrial engineering, mechanical engineering, or medical fields. For many computer vision applications, object recognition and classification are an essential approach. Function descriptors and detectors have been used prominently to resolve object variations in location, rotation, size, color, brightness, etc. Different types of issues with automatic recognition usually need different features and learning programs. In modern industrial and mechanical applications, such as mechanical engineering or automotive manufacturing. It is always important to be able to control various instruments on the plant equipment with field devices, such as sensors or actuators, which are needed for the functionality of the tool.

Recently, in 2021, P. Jain et al [3] provide an artificial neural network application to identify mechanical instruments independent of their location, scale, and orientation. They try to deal with the recognition of objects in the image given, a supervised learning technique in the form of neural network feed-forward with back-propagation is used.

This paper presents a CNN-based model to recognize mechanical tools. Four models have been applied to identify mechanical tools. Six mechanical tools have been included in this study. It is also a matter for the invention to make it possible to distinguish work parts and instruments of equal size by size. A similar purpose is to make individual socket, box end and open end wrenches of about the same size openings easier to spot. Another innovation purpose is to allow the consumer to rapidly pick a tool of the desired dimension without seeing either the tool's opening or the scale.

II. METHODOLGY

A. Exprimintal Dataset

The dataset that was used in this paper is available on Kaggle website [4], it has multiple and different object classes from mechanical tools scenes. It contains about 7527 images for 8 classes but we trained and test on 2961 images for 6 classes such as *gasoline can*, *hammer*, *pebbles*, *Rope*, *toolbox* and *wrench* which have different size and style. We divided 70% for training (2073 images), 20% for validation (592 images) and 10% for testing (296 images).

B. Proposed Model

The proposed system for recognizing mechanical tools is shown in the following figure.

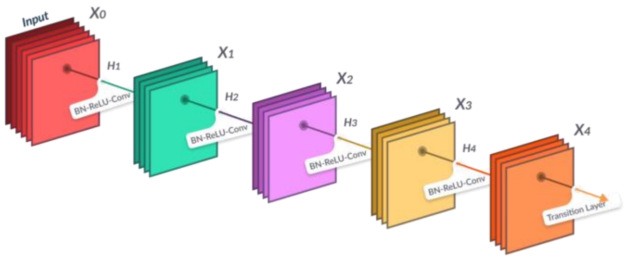

Fig. 1. The proposed CNN-based model to recognize mechanical tools.

In our approach, we trained the six classes of the dataset on the deep convolutional ResNet network that was trained on more than a million images from the ImageNet model, this training gave it feature-rich representations of a wide range of images. This network aims to identify the objects of the input pictures; the size of the image input for the network is 224 x 224. We trained the dataset on several layers of the ResNet network, the architecture of ResNet layers is shown in (Fig1).

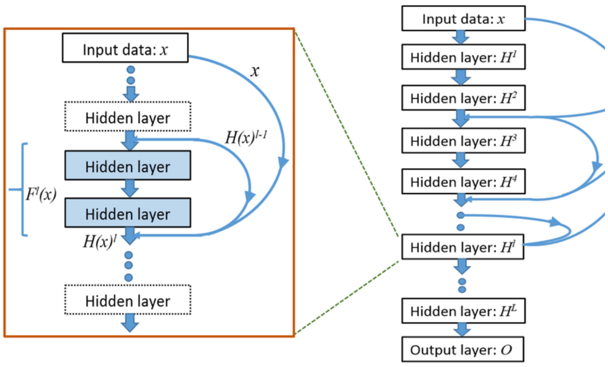

Fig. 2. Resnet structure.

As shown in the above figure (Fig. 2), the images are inputted to the CNN. Then, the hidden layers apply filters and many operations to the model as extracting the features of the input image. Finally, the classifier which is located at the end of the neural network will shows the recognized class.

C. ResNet Archeticture

The residual neural network (ResNet) is an artificial neural network which applies mapping that relies on well-known structures that pass the entry directly from one layer to another[7].

ResNet was appeared after the success of AlexNet in the LSVRC 2012 classification contest [8], it was able to be the leader in the fields of computer vision and the deep learning community within the recent years. Its powerful representational ability has enhanced the performance of many computer vision applications, such as object detection and face recognition, additionally to image classification.

Before the appearance of ResNet, there were numerous attempts across various methods to solve the vanishing gradient problem which a problem associated with deep learning, but none of them really addressed the problem permanently and forever[9, 10, 11].

After the success of residual network architecture that proposed by Microsoft Research researchers in 2015[12], which worked to solve the vanishing gradient problem by using skip connection technique which means skipping some layers in the neural network and feeding the output of one layer to another layer by skipping layers in between. The main approach of using this architecture is to allow network to fit the residual mapping rather than layers learn the basic mapping. We can represent this proportion by the following equation [8]:

F (x): = H (x) – x, which gives H(x): = F(x) + x.

The basic idea of how the ResNet works is the "Identity Shortcut Connection" that skips from one or more layers and connects directly to the output. As shown in the following figure:

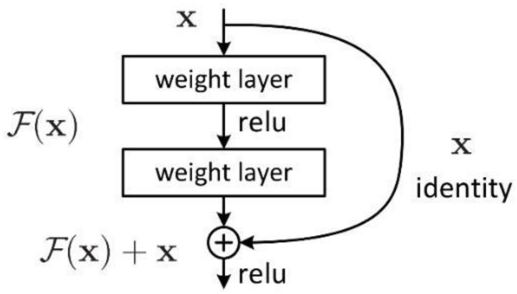

Fig. 3. dentity connections.

We can design ResNet architecture (including full blocks) using Tensorflow and Keras API, from scratch.

.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-lay
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	1×1, 64 3×3, 64 1×1, 256
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	1×1, 128 3×3, 128 1×1, 512
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	1×1, 256 3×3, 256 1×1, 1024
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	1×1, 512 3×3, 512 1×1, 204
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×1

Fig. 4. The architecture of Residual Network.

As shown in the above table [12], the residual networks are constructed of convolution layers with skip connections. These connections are the most are the main difference between ResNet networks and other deep learning pre-trained models.

D. Transfer Learning

Transfer learning [5] endeavor to store learned weights knowledge from one domain and enforcement it to another different domain but related to the first domain. Because when training from scratch, it sometimes takes a lot of time. So, instead of training DNN (dense neural network) from scratch, the model will convey the features or weights that were learned from different dataset that have executed the same tasks. A pre-trained CNN models [6] are effective feature extractor. We used the transfer learning technique on the pre-trained ResNet-18, ResNet-34, ResNet-50, ResNet-152 and ResNet-101 models which were trained on ImageNet to train them on Mechanical Tools detection. To employ the perfect performance of ResNet, a new structure dependent upon ResNet is built with parameters transferred from the pre-trained model.

*E. **Training and Testing Phases***

The training phase is done using transfer learning. Four pre-trained models are used including ResNet-18, ResNet-34, ResNet-50, and ResNet-152.

III. RESULTS AND DISCUSION

First we trained the dataset images on the high performance ResNet18 and ResNet34 architectures, the structure of both networks includes five convolutional stages, one convolution and pooling step followed by 4 layers of similar behavior, achieving an accuracy of 84 and 85% respectively. We also trained on the ResNet50 architecture with a 50-layer depth, this model consists of 5 stages, each stage includes a wrapping block containing 3 wrapping layers and an identity block including 3 wrapping layers, after training with this model we got an accuracy of 86%. ResNet152 network, with a depth of 152 layers, also used in training and achieved an accuracy of 87%. The training accuracy and loss shown in fig (5) and fig (6) respectively. The proposed model is implemented using Tensorflow [16] and Keras [17]. The optimizer that we used is Adam optimizer [18] with learning rate of 1e-4. The model trained for 24 epochs with a batch size of 32. It takes 7 hours on computer with I5-2450M CPU at 2.5 GHz and 6 GB memory.

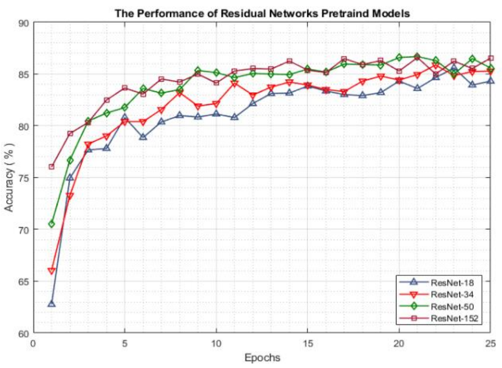

Fig. 5. Training accuracy.

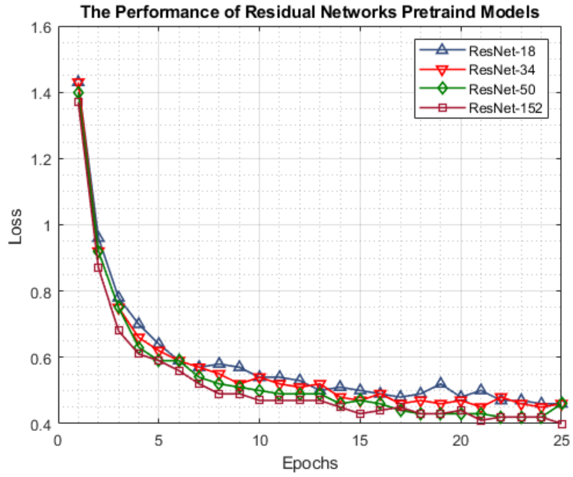

Fig. 6. Training loss.

Training Model	Accuracy	Loss
ResNet-18	84%	0.4640
ResNet-34	85%	0.4684
ResNet-50	86%	0.4608
ResNet-152	87%	0.4002

Table.1. Results of Training Accuracy and Loss.

IV. CONCLUSION AND FUTURE WORKS

We trained a set of classes for mechanical tools images dataset on a CNN model like ResNet that was pre-trained on a millions of images[12]. The CNN-based model was trained using four networks of residual net ResNet-18, ResNet-34, ResNet-50 and ResNet-152. ResNet-152 achieved the highest accuracy of 87%. One of the difficulties that we faced in this paper is that the dataset is not available for a lot of classes, which makes the classification process limited to certain mechanical hand tools. In the future, we aspire to create our own dataset that includes the largest possible number of classes of mechanical hand tools, and we are planning to build our own model with a high-precision and performance to train all the classes of the dataset on it.

REFERENCES

[1] Editorial, "Advances in Pattern Recognition,Pattern Recognition Letters," 2005, pp. 395-398.

[2] M. Paolanti, "Multidisciplinary Pattern Recognition applications," June 2020.

[3] P. Jain, "Recognition of Mechanical Tools Using Artificial Neural Network," in *Recent Advances in Mechanical Engineering*: Springer, 2021, pp. 637-644.

[4] Mechanical Tools Dataset From Kaggle. Last Accessed Jan 2021.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 7695-6321.

[6] S. Marcel, R. Erik, and D. Joachim," ImageNet pre-trained models with batch normalization," in arXiv:1612.01452v2 [cs.CV], 6 Dec 2016.

[7] Rezende, E., Ruppert, G., Carvalho, T., Ramos, F., & de Geus, P, " Malicious Software Classification Using Transfer Learning of ResNet-50 Deep Neural Network," in Proceedings of the IEEE International Conference on Machine Learning and Applications (ICMLA), 2017, pp. 770-778.

[8] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., … Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), 211–252.doi:10.1007/s11263-015-0816-y .

[9] B. D. Ripley. Pattern recognition and neural networks. Cambridge university press, 1996.

[10] C. M. Bishop. Neural networks for pattern recognition. Oxford university press, 1995.

[11] W. Venables and B. Ripley. Modern applied statistics with s-plus. 1999.

[12] H, Kaiming., Z, Xiangyu., R, Shaoqing., S, Jian, "Deep Residual Learning for Image Recognition," Microsoft Research, arXiv:1512.03385v1 [cs.CV], 10 Dec 2015.

[13] R. K. Srivastava, K. Greff, and J. J. A. i. n. i. p. s. Schmidhuber, "Training very deep networks," vol. 28, pp. 2377-2385, 2015.

[14] Y. LeCun, Y. Bengio, and G. J. n. Hinton, "Deep learning," vol. 521, no. 7553, pp. 436-444, 2015.

[15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700-4708.

[16] M, Abadi., P, Barham., J, Chen., Z, Chen., A, Davis., J, Dean., M, Devin., … Z, Xiaoqiang., "TensorFlow: A System for Larg-Scal Machine Learning", Open access to the proceeding of the 12th {USENIX} symposium on operating systems design and implementation (OSD'16). November 2-4,2016. ISBN 978-1-931971-33-1.

[17] Chollet, F. & others, 2015. Keras. Available at https://github.com/fchollet/keras.

[18] Kingma, Diederik P. and Jimmy Ba. Adam: A Method for Stochastic Optimization. CoRR abs/1412.6980 (2015): n. pag.