Open Access
Full Length Article

ORIGINAL ARTICLE

AI-Enabled Elderly Care Robot

Iram Noreen[1], Adeel Akbar[2], Uzair Siddiqui[3]

[1]Department of Computer Science, Bahria University, Islamabad, Lahore 54000, Pakistan.
[2]Department of Computer Science, Bahria University, Islamabad, Lahore 54000, Pakistan.
[3]Department of Computer Science, Bahria University, Islamabad, Lahore 54000, Pakistan.

ABSTRACT

Increasing average human age and growing number of old age population over the globe has emerged the need of automation in health care domain. Eventually, health care robots will contribute to manage the workload of care provider in future health care domain. Elderly people face major challenge to identify and pronounce names of daily use objects and different family and friends due to dementia related issues. One of the many tasks by AI-Enabled elderly care humanoid robots will be the communication with such patients. They will aid them to remember objects and persons they find difficult to identify and recognize due to dementia. This study has investigated the potential of JD humanoid robot for object identification in such scenario. The study has presented prototype AI-Enabled elderly care robot which can aid elder persons to identify different objects. The robot is trained on ImageNet data set for object identification. Further, a personalized data set comprising of faces of different persons labelled by names of family and friends is also used to train robot for family recognition aid. The JD Humanoid robot captures image of object/person and also tracks it as the object/face moves. It identifies input image and outputs in voice the name of the object/person to guide the elderly person. The third feature of robot is the notifications of daily routine tasks for elder person like reminder for prayer, exercise walk and medicine intake. Robot training has shown object and person identification with 95% accuracy.

Keywords: Elderly Care, Convolutional Neural Network, humanoid Robot, Object Detection

Author`s Contribution
[1,2,3] Manuscript writing, Data Collection Data analysis, interpretation, Conception, synthesis, planning of research, and discussion

Address of Correspondence
Irum Noreen
iram.bulc@bahria.edu.pk

Article info.
Received: March 17, 2020
Accepted: October 28, 2020
Published: December 30, 2020

Cite this article: *Minallah N, Ashfaq M. Iterative non-rigid image registration in Brain MRI Images. J. Inf. commun. technol. robot. appl.2020; 11(2):22-29*

Funding Source: *Nil*
Conflict of Interest: *Nil*

INTRODUCTION

Robots makes use of existing ICT infrastructure and also imply a continued technological evolution of computer networks. Humanoid is a category of robots that is designed to look like humans, and their interactions/behavior may demonstrate intelligence. Nowadays, every physical work is shifting on the machine, therefore, robots are being used as workforce in automation. Humanoid robots are being used in the scenarios where human interaction is of central focus such as health care domain, office assistance and client interaction in restaurants. Human society is on verge of an old dream come true, i.e., a friendly device that resembles to human and can interact with humans. Robotics is on rise and in future it will extend with automation of child care, elderly care and medical assistance facility as well [1]. The elderly care robots will not only provide support to old people and patients to perform daily tasks but they will also assist to monitor their behavior and will serve as companion to the elderly persons [2]. However, the use of robots for tasks in close vicinity of humans is very complicated due to inbuilt intelligence needed for the automation of numerous human activities [3].

Recently, Japan has shown tremendous progress in medical and health care robots due to their growing and un-intended old age population. Future prediction of increase in average human age has emerged the need of health care robots to manage the workload of care providers [4]. The need of this era is to find an ease for the elderly people. Elderly persons especially with dementia ailment, suffer from memory loss and cognitive impairment. They face difficulties to identify and narrate different objects of routine life. They also find it difficult to recognize the faces of family and friends [5]. This is the dilemma of old people in modern age, not only in Pakistan but in the whole world.

This study is conducted to develop the module for “JD Humanoid” robot to aid elderly person in identification and recognition tasks. “JD Humanoid Robot” is trained as an initial prototype for elderly care support and to help them in their dementia condition to identify different objects/persons. The project has trained JD robot using convolutional neural network algorithm to work as a daily routine assistant. It will enable robot to identify different objects/persons by using object detection and tracking features. The rest of the paper is structured as follows. Section 2 describes materials and methodology. Section 3 describes experimental setup. Section 4 presents results discussion followed by conclusion in Section 5.

LITERATURE REVIEW

Robotic applications are in rise in recent year specifically for health-care therapy in elderly patients [6]. Object detection and tracking is done in literature using a number of techniques. Objects are detected and tracked by different aspects, different angles and dimensions. There are basic three models to track the objects. One of the well-known method is spatial model, which is related to the geometric shape and verification of an object. Second is temporal model, which is related to the multi tasked structure and uses adjustment of the frames and the third is discriminative feature, which is related to the intra-class compactness and inter class reparability. Many dimensions of the object can be analysed and investigated using feature representation. During feature representation, it separates the object from the background to its foreground. Further, object model learning performs classification of object [7].

Sensors used for object detection are required to provide rapid and quick response otherwise sensors may provide noisy or delayed data stream resulting in varied results. Therefore, fast and efficient algorithms are reported in literature to be used in robot’s Navigation Systems and Motion modelling such as HMM (Hidden Markov Model) [8]. The vision of robot is also affected by distance of object from robot and its motion angles as well. Similarly, there are different object tracking techniques such as drift alleviation based tracking, correlation filter based tracking, deep learning based tracking [7]. Similarly, use of reinforcement learning has also been in practice in research community or object detection and tracking [9]. The object may be tracked in the 2D or 3D, the sum of squared differences (SSD) calculates the displacement vector during each time interval [10]. Recently, Convolutional Neural Networks (CNN) have gained immense popularity for object classification due to their ability to deal with the complex video and/or image data and affine transformation invariant features [11]. Chiatti et al. [12] proposed a method based on Convolutional Siamese Network for object recognition through shot-image matching for mobile robots.

METHODOLOGY

The project methodology comprises of two distinct phases. First, phase performs training for object classification; second, phase is related to object detection and tracking. A top level architecture and data flow of the proposed approach is shown in Figure 1 and

Figure 2 respectively.

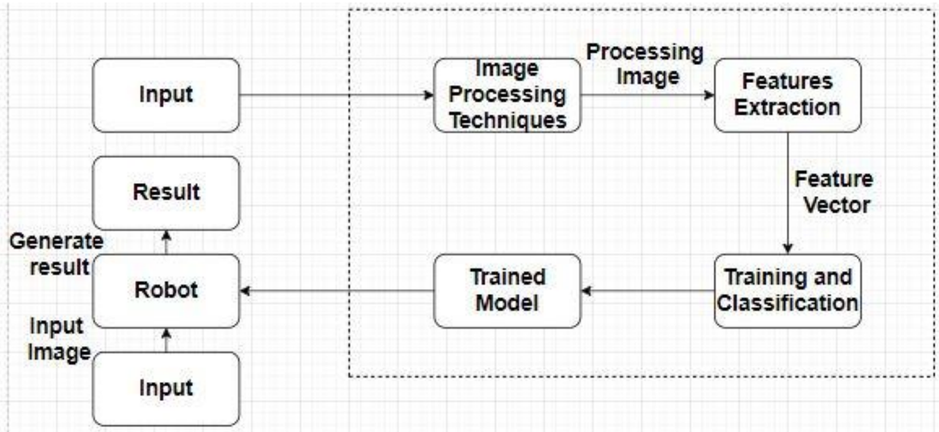

Figure 1. Top Level Architecture

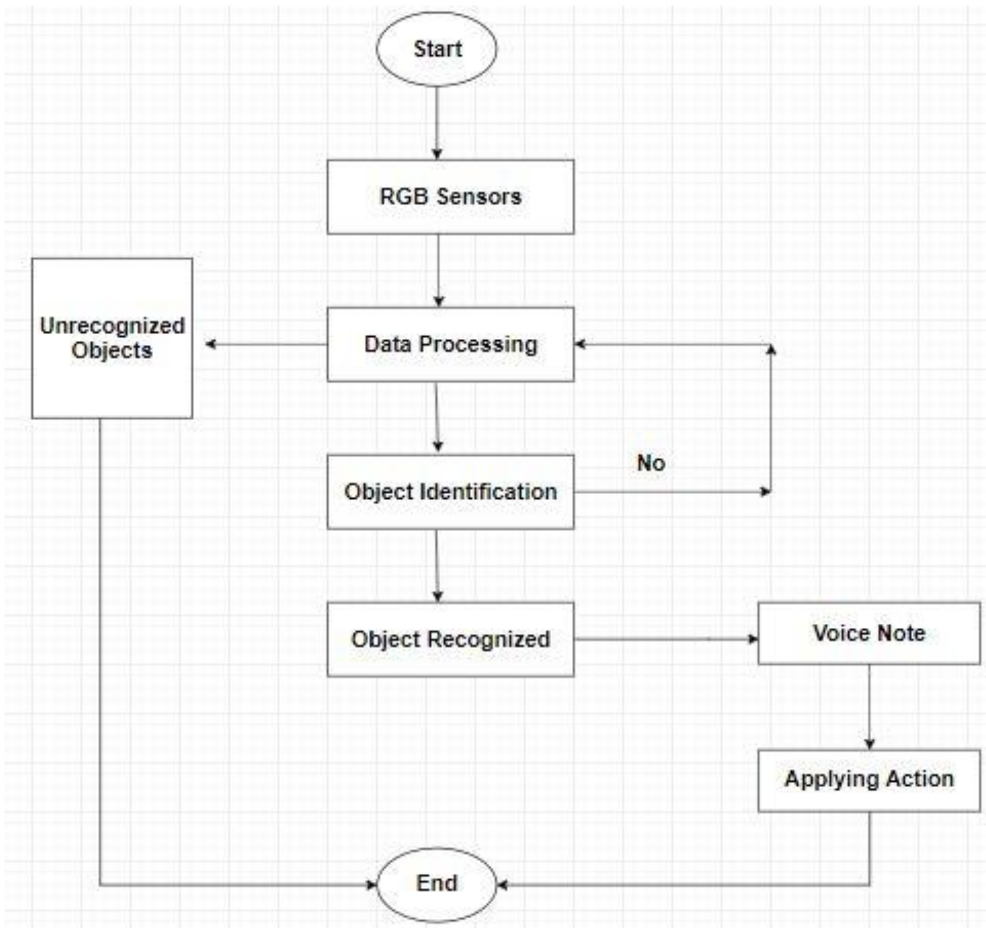

Figure 2. Date Flow Diagram

A. Object Training and Classification

Convolution Neural Network (CNN) is used for training of different objects and persons face classification. Feature Extraction is also automatically managed by CNN. CNN comprises of three types of layers; input, hidden and output. Hidden layers are further categorized as convolutional layers and max pooling layers. The architecture of a standard CNN [13] is presented in Figure 3.The detail of different layers of CNN is provided as follows:-

Input Layers: This layer is responsible to received input for model. The number of neurons in this layer is equal to total number of features in data set samples and in this scenario the number of pixels of our image. Image frame is extracted from live video stream received by robot's sensors. The robot receives image of object as an input and applies different pre-processing techniques such as image rescaling for 32 x 32 x 3 dimension. So this layer holds the raw input of image with width 32, height 32 and depth 3.

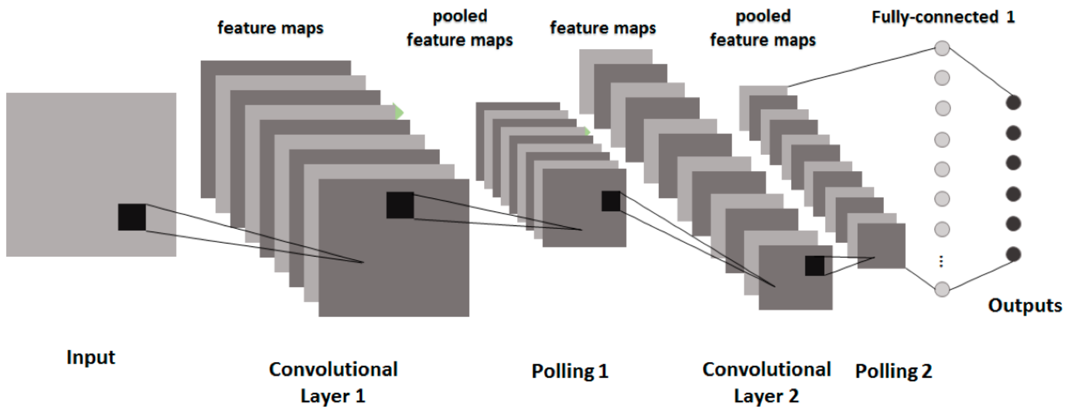

Figure 3. A standard CNN Model [13].

Hidden Layers: The input from input layer is then feed into the hidden layers. This layer is responsible for training and feature learning. There can be many hidden layers depending upon model and data size. Each hidden layers can have different numbers of neurons which are generally greater than the number of features. The output from each layer is computed by matrix multiplication of output of the previous layer with learnable weights of that layer and then by addition of learnable biases followed by activation function which makes the network nonlinear. In CNN a number of convolution and pooling layers are used alternatively to extract feature. Lower layers extract high level features whereas higher layers extract the fine grain detail from the image. Thus every layer transforms one volume to another through differentiable function.

Convolution Layers compute the output volume by computing dot product between all filters and image patch. We have used total 12 filters for this layer and we'll get output volume of dimension 32 x 32 x 12. Leaky ReLU activation function is applied after convolution process. It applies element wise activation to the output of convolution layer. The volume remains unchanged hence output volume will have dimension 32 x 32 x 12.

Max pooling layers are periodically inserted and its main function is to reduce the size of volume which makes the computation fast reduces memory and also prevents from overfitting. Two common types of pooling layers are max pooling and average pooling. We have used Max pooling layer as it also removes noise from image and provides better results. We have used a max pool with 2 x 2 filter and a stride size of 2 and producing a resultant volume with dimension 16x16x12. Fully-Connected (FC) layers add a regular fully connected neural network after convolution and pooling layers. A dense layer is formed using flattening to provide input to FC. It takes input from the previous layers and computes the class scores and outputs the 1-D array of size equal to the number of classes.

Output Layer: Fully connected layers are connected with the output from the hidden layer is then fed into logistic function softmax which converts the output of each class into probability score of each class.

B. Object detection and Tracking

The robot receives image of object as an input and applies different image processing techniques such as pre-processing using thresholding, image segmentation and transformation. Object tracking is done using RGB images from the frames of live video stream. Object tracking process first extracts red channel, then blue and in the end extracts the green channel of the image frame. After color channel extraction, it gets final colored images with enhanced pixels. Then this module applies color channel information to separate foreground object from its back ground using segmentation technique. After object segmentation, it is labelled and then object tracking process continues till the objects remains in front of the robot.

EXPERIMENTAL SETUP

EZ builder Library is used to code using python and EZ script (based on C++) for JD Humanoid Robot. EZ builder IDE is used to deploy code over robot using Windows operating system. The computer/laptop containing the EZ builder IDE and the JD Humanoid robot will be on the same network and communicate through the network. Both the laptop and the JD Humanoid robot must be on the same network at all times to work properly. A deployment model is shown in Figure 4.

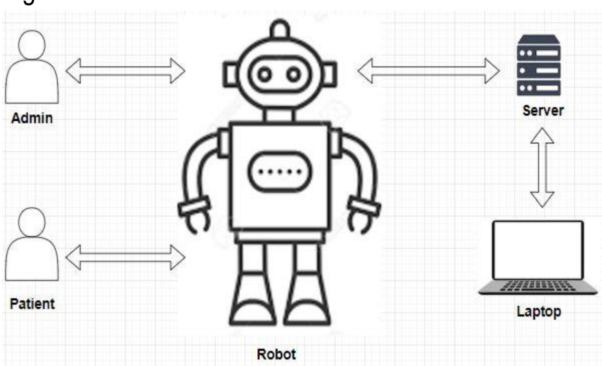

Figure 4. Deployment Model

Data Set

An image dataset of daily routine objects ImageNet [14] is used. It provides more than 14 million images to train. The average image resolution in ImageNet is 469x387 pixels. They are rescaled for 32 x 32 dimension. Moreover, different persons (to be identified by names as family members by elderly person) images are captured by Robot's eyes and then these images are also added in the database. In the end, a unified dataset is created by merging different sys-net images from ImageNet and different person images collected using robot's visual sensors. All images are pre-processed on same scale for training using CNN. Figure 5 shows process of image capturing using robot's camera. A sample of data used during training for detection is shown in Table 1.

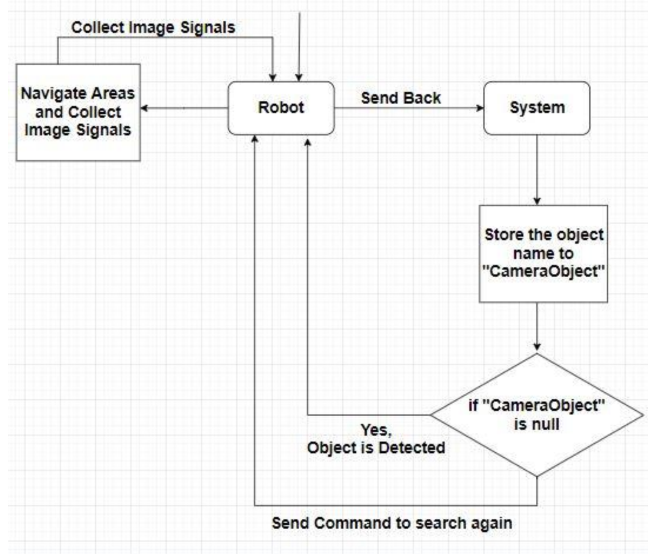

Figure 5. Data capturing flow using robot's eyes camera

Table 1. Processing Time for Different objects

Board	Robot				
	1	2	3	4	5
CameraObjectName	Nurse	Umar	Adeel	Iram	Uzair
CameraObjectTracking	1	2	1	3	1
Frame Volume	100	90	110	98	95

Training Session

Data is categorized as 90:10 ratio for training and testing. Further a 10-fold cross validation is used. To include family and relatives faces in dataset, faces are captured live from robots camera and are saved. Different stages during training session are shown in

Figure 6. Experimental setup during training of JD humanoid is conditional with the following assumptions:-

- The JD Humanoid robot must be kept in a Dry Environment as water can damage the circuits of robot and affect its performance.
- Don't sudden switch off the robot while performing any functionality on the robot.

Further, notifications of different timeslots for medicine intake, walk reminder and prayer reminder are fed into the robot using python programming language. Robot alerts person using audio message about the scheduled activity.

Figure 6. Training Session

RESULTS AND DISCUSSION

Robot was trained for object recognition in two different scenarios. Detail of each scenario is described in the following sub-sections.

Scenario 1: Identify Object from front angle

From the front angle the detection take place from multiple ways its angle of freedom of 16 degree. The detection takes place when the object is placed in front of the robot visual sensor. The robot have 18 sensors, during the detection all sensors becomes active, during training it picks up the different points of an image. It takes up the different frames of the image and uses those frames during the detection. When it detects an object it gives response using audio channel and tells the specific objects name which we have trained. All the working has been done on the bases of CNN algorithm. Robot can detect an object from front angle by moving its Servo at 180 shift. The detection can take place at up, down, right, left, forward, roll, and pitch and yaw position getting view from its front side.

Scenario 2: Robot Get Confused

Sometimes, it gets confused due to illusion factor, motion blur then its sensor lights fade or blinks. The error occurs when the buffer of the previous object while cleaning and the other object loaded on the buffer so as the result the robot become confuse. This is because conditional statement executes so fast that the previous buffer was removing while the next buffer start loaded so the robot become confuse. Summary of different evaluation measures are shown in Table 2 and results of different evaluation measures are shown in Table 3.

CONCLUSION

Elderly care robots are the future of health care domain and robotic sector. Object detection and face recognition is one of the many challenging tasks required by elderly care humanoid robots. A prototype AI-Enabled elderly care robot is presented for object and person identification in order to assist dementia effected elder persons. In this study, convolutional neural network for object identification is trained on JD Humanoid robot as a care taker prototype for elderly persons with dementia. The robot demonstrated the testing accuracy of 95%. The future plans are to extend this prototype to next intelligence level of care taker, which includes implementation of different computer vision based algorithms in robot to make different types of interactions with elderly person. Other recommendations to enhance the prototype are gestures and voice control function, and speech recognition function. Moreover, an android smart application will also be developed to facilitate elderly person and/or care taker of elderly person to perform different tasks and configure robot's friends and family recognition training.

Table 1. Summary of Evaluation Metric		
	Description	**Formula**
True Positive (TP)	Correctly predicting a label	we predicted "yes", and it's "yes"
True Negative (TN)	Correctly predicting the other label	we predicted "no", and it's "no"
False Positive (FP)	Falsely Predicting a label	we predicted "yes", but it's "no"
False Negative (FN)	Missing and incoming label	(we predicted "no", but it's "yes"
Sensitivity	Sensitivity or True Positive Rate (TPR)	TPR = TP / (TP + FN)
Specificity	Specificity (SPC) or True Negative Rate (TNR)	SPC = TN/ (FP + TN)
Precision	Precision or Positive Predictive Value (PPV)	PPV = TP / (TP + FP)
Negative Predictive Value	Negative Predictive Value (NPV)	NPV = TN / (TN + FN)
False Positive Rate	Fall-out or False Positive Rate (FPR)	FPT = FP / (FP + TN)
False Discovery Rate	False Discovery Rate (FDR)	FDR = FP / (FP + TP)
False Negative Rate	Miss Rate or False Negative Rate (FNR)	FNR = FN / (FN + TP)
Accuracy	Accuracy (ACC)	ACC = (TP + TN) / (P + N)
F1 Score	F1 Score (F1)	F1 = 2TP / (2TP + FP + FN)
Matthews Correlation Coefficient	Matthews Correlation Coefficient (MCC)	TP*TN – FP*FN / sqrt ((TP + FP) *(TP + FN)*(TN + FP)*(TN + FN))

Table 2. Testing Results of different Metric	
Measure	Value
Sensitivity	0.6667
Specificity	1.0000
Precision	1.0000
Negative predictive value	0.9444
False-positive rate	0.0000
False discovery rate	0.0000
False-negative rate	0.3333
Accuracy	0.9500
F1 score	0.8000
Matthews correlation coefficient	0.7935

REFERENCES

1. e. a. VERCELLI, “Robots in Elderly Care,” DigitCult - Scientific Journal on Digital Cultures, vol. 2, no. 2, pp. 37-50, 2018.
2. M. &. M. H. Niemelä, “Robots as social and physical assistants in elderly care. In,” in Human-centered digitalization and services. Translational Systems Sciences, Singapore, 2019.
3. O. K. D. Poulsen A., “The Ethics of Inherent Trust in Care Robots for the Elderly,” in IFIP International Conference on Human Choice and Computers, Cham, 2018.
4. K. K. M. &. R. T. Coco, “Care personnel's attitudes and fears toward care robots in elderly care: a comparison of data from the care personnel in Finland and Japan,” Journal of Nursing Scholarship, vol. 50, no. 6, pp. 634-644, 2018.
5. G. J. H. A. S. D. A. C. B. S. B. C. B. e. a. Livingston, “Dementia prevention, intervention, and care: 2020 report of the Lancet Commission,” The Lancet , vol. 396, no. 10248, pp. 413-446, 2020
6. A. e. a. Cifuentes, “Social Robots in Therapy and Care,” Current Robotics Reports , pp. 1-16, 2020.
7. H. W. a. F. X. Y. Zhang, “Object detection and recognition of intelligent service robot based on deep learning,” in International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), 2017.
8. B. M. G. D. H. M. H. a. N. B. M. Haid, “Inertial-Based Gesture Recognition for Artificial Intelligent Cockpit Control using Hidden Markov Models,” in IEEE International Conference on Consumer Electronics (ICCE), 2019.
9. P. S. F. Z. W. L. T. Z. a. Y. W. W. Luo, “End-to-end active object tracking and its real-world deployment via reinforcement learning,” IEEE Transactions of Pattern Analytics and Machine Intelligence , vol. 42, no. 6, pp. 1317-1332, 2019.
10. Z.-Q. e. a. Zhao, “Object detection with deep learning: A review,” IEEE transactions on neural networks and learning systems, vol. 30, no. 11, pp. 3212-3232, 2019
11. M. C. Y. W. W. e. a. Hsu, “A Convolutional Fuzzy Neural Network Architecture for Object Classification with Small Training Database,” International Journal of Fuzzy Systems , vol. 22, pp. 1-10, 2020.
12. A. e. a. Chiatti, “Task-agnostic object recognition for mobile robots through few-shot image matching,” Electronics, vol. 9, no. 3, p. 380, 2020.
13. J.-R. J.-E. L. a. Y.-M. Z. Jiang, “Time Series Multiple Channel Convolutional Neural Network with Attention-Based Long Short-Term Memory for Predicting Bearing Remaining Useful Life,” Sensors, vol. 20, no. 1, p. 166, 2020.
14. J. e. a. Deng, “Imagenet: A large-scale hierarchical image database,” in IEEE conference on computer vision and pattern recognition, 2009.