



## Building a High Accuracy Transfer Learning-Based Quality Inspection System at Low Costs

**Ahmed Najah\***

**Faiz F. Mustafa\*\***

**Wisam S. Hacham\*\*\*\***

*\*, \*\*Department of Automated Manufacturing Engineering / Al-Khwarizmi College of Engineering/  
University of Baghdad*

*\*\*\*Department of Mechatronics Engineering/ Al-Khwarizmi College of Engineering/ University of Baghdad*

*\*Email: [ahmednajah5049@gmail.com](mailto:ahmednajah5049@gmail.com)*

*\*\*Email: [faizalrawy@yahoo.com](mailto:faizalrawy@yahoo.com)*

*\*\*\*Email: [wisam@kecbu.uobaghdad.edu.iq](mailto:wisam@kecbu.uobaghdad.edu.iq)*

(Received 5 October 2020; accepted 6 December 2020)

<https://doi.org/10.22153/kej.2021.12.001>

### Abstract

Products' quality inspection is an important stage in every production route, in which the quality of the produced goods is estimated and compared with the desired specifications. With traditional inspection, the process rely on manual methods that generates various costs and large time consumption. On the contrary, today's inspection systems that use modern techniques like computer vision, are more accurate and efficient. However, the amount of work needed to build a computer vision system based on classic techniques is relatively large, due to the issue of manually selecting and extracting features from digital images, which also produces labor costs for the system engineers.

In this research, we present an adopted approach based on convolutional neural networks to design a system for quality inspection with high level of accuracy and low cost. The system is designed using transfer learning to transfer layers from a previously trained model and a fully connected neural network to classify the product's condition into healthy or damaged. Helical gears were used as the inspected object and three cameras with differing resolutions were used to evaluate the system with colored and grayscale images. Experimental results showed high accuracy levels with colored images and even higher accuracies with grayscale images at every resolution, emphasizing the ability to build an inspection system at low costs, ease of construction and automatic extraction of image features.

**Keywords:** *Backpropagation Algorithm, convolutional neural networks, computer vision, deep learning, image classification, quality inspection, transfer learning.*

### 1. Introduction

Traditional layouts, has consistently been a drawback with regards to an excellent product. Manual inspection is one of the problems that adds to inferior production. Thus, quality inspection (QI) is considered as one of the most important stages in the flow of production, which decides if the item has conformed to the ideal specifications, and fit to be packed or it should be turned into scrap. However, sometimes an item can be reprocessed, in which the damage can be relatively small such as bad surface finish.

There are numerous devices that are utilized for QI, contingent upon the necessary inspection task. Conventionally, experts took care of the vast majority of those assignments manually, yet those techniques were not proficient. They did not just increase lead times, yet additionally expanded labor and product costs, lower rate of production, and eventually prompted a decrease in income. These days, QI experts and designers, lean towards discovering advanced solutions and new methodologies, to achieve the objective of ideal quality at low expenses, and less consumption of time. Inspection is considered as a tool within quality control, which distinguishes the defects

and assures the level of quality of the product, new methodologies that achieve this task comprise of: inspection systems that are fully automated rather than manual techniques, sensors systems to perform on-line assessment rather than off-line review [1].

Since QI is an important stage in the production cycle, attention must be focused on its methodological improvement. To satisfy this demand; advanced technologies such as computer vision (CV), which is one of the artificial intelligence (AI) applications should be used. This technology gives frameworks the ability to interpret the visual world by using Deep learning (DL) models and digital images, in which the machine can distinguish and classify the objects. Generally, gears are recognized as essential components in any machine, they can be found separated or working together in a gearbox to manipulate the movement of wheels. Since gears play an essential role in machinery, they must be manufactured carefully. Therefore, optimum quality of gears must be obtained and that can be assured by performing inspection on the produced gears.

Recent progressions that are concerned with the inspection of gears have broadly utilized mathematical analysis strategies to achieve inspection tasks, for example, detection of plastic gear defects with image processing [2], using wavelet transform for fault detection of planetary gears system [3], detection of gear faults using: morlet-wavelet filter [4], adaptive wavelet threshold de-noising [5] and cosine similarity, wavelet transform and Hilbert transform [6]. Moreover, gear faults diagnosis using: adaptive impulsive wavelet transform [7], utilizing extreme learning machines and numerical simulation [8], discrete wavelet packet for feature selection of gear faults [9] and inspection of polymer spur gears [10]. Advanced technologies like AI and CV are also employed for inspection, such as: using machine vision for spur gears parameters measurement [11], using CV to detect gear tooth number [12], using artificial vision for quality control of spur gears [13], inspection of gear faults using support vector machines (SVMs) and artificial neural networks (ANNs) [14], determining fine-pitch gears centers using machine vision [15], gear faults with convolutional neural networks (CNNs) [16], gears diagnosis using CNNs [17] and inspection of plastic gears using ANN and SVM based method [18]. AI and CV are also used for other AI inspection related application like: dimensions inspection with machine vision [19], detection of

defects in products [20], sugarcane varieties inspection [21], welding inspection [22] inspection of optical laser welding [23] and inspection of aerospace components [24]. Vibration signals were the source information in most of the gears related literature mentioned above.

Therefore, to stay updated with the state of the art, in this research, we present a transfer learning approach that is based on CNNs to classify helical gears into healthy and defected classes in attempt to build an accurate, low cost inspection system with automatic features extraction. Since this methodology depends on neural networks, manual features extraction was not involved. Also, the issue of lacking data is bypassed by using a deep neural network (DNN) that was previously trained on (1.2 million) pictures from ImageNet [25]. The parameters from the trained network are imported to the new architecture as the initial segment and an undeveloped neural network fills in as the subsequent part, which accommodates the task of gears faults identification. The subsequent part is trained on test data that comprised of 4000 pictures. As will be demonstrated later, a high accuracy system can be established at low expenses and with no preprocessing methods for the extraction of features.

Despite the fact that features extraction with the empirical and manual techniques indicated success at various levels, evidently their applicability depends on the features that are extracted from the analysis and may not function on other systems. This method is referred to as a descriptive analysis, in which the analyzer needs to gather process information, assemble a hypotheses on information patterns, and compare the results of the descriptive model with the genuine results to verify the hypotheses [26]. However, it is unsafe to form this type of models, as there is a risk of not modeling some of the variables that scientists and engineers do not include due to lack of information or not understanding the problem [27].

On the other hand, predictive analysis, finds the rules that underlie a phenomenon and establishes a predictive model that limits errors between the desired and the predicted results, with all the involved factors taking into consideration [26]. In contrast with conventional CV techniques, predictive analysis is used by DL to solve issues, which grant DL the advantage of reaching high accuracies with CV applications such as image classification, semantic segmentation and object detection. Since DL relies upon DNNs that are trained rather than

programmed, applications that depend on this strategy involves basic analysis, calibrating and exploit the huge amount of information that is accessible today within every system. Moreover, DL is viewed as a truly adaptable technique as CNNs structures can be re-employed for custom data with various applications by training them again, unlike conventional algorithms which are generally intended for a particular domain [28]. The remainder of this paper is organized in the following matter, in section two, CNNs are briefly introduced, transfer learning and the proposed architecture, in section three experimental work with training and evaluation are explored, results and discussion are elaborated in section four, and finishing with sections five and six by including the conclusions and the used references respectively.

## 2. Convolutional Neural Networks

A convolutional neural network (or CNN) is one of the DL algorithms that is similar to the neurons connections in the visual cortex of the

animals [29]. The CNNs are considered as a form of DNNs, because they are consist of many layers, as appeared in Fig. 1. CNNs utilize a linear operation called convolution Rather than general multiplication of matrices, which is used by standard ANNs [30]. They are referred to as the most accurate object detection/recognition algorithm, and like different DNNs, they rely upon substantial amounts of data to be trained and give accurate results. This algorithm learns every filter and extract features automatically in contrast to other manual algorithms.

As appeared in Fig. 1, a typical architecture for a CNN comprises of convolutional layers, an activation function (for example ReLU), a pooling layer (for example max pooling), and a flattened layer. A single vector is produced by flattening the pooled images, the vector is then used as an input for a fully connected ANN for the processing of the features. After training through forward propagation and backpropagation with a number of epochs, the last network will be prepared to give the decision on the image class that is intended to search for.

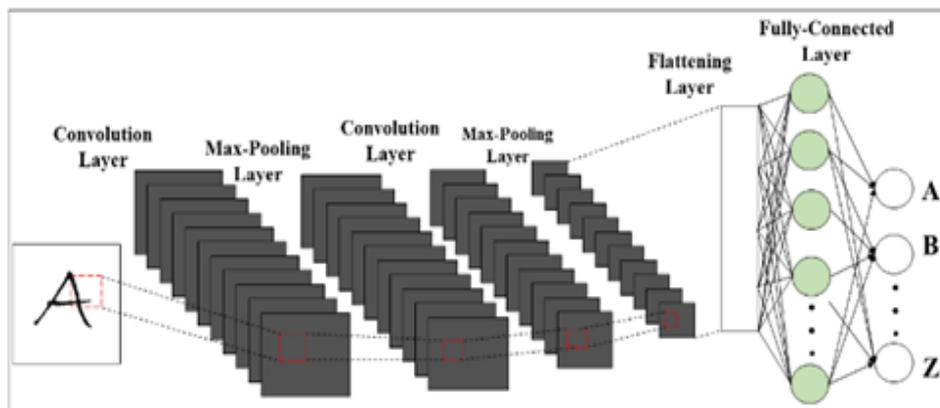


Fig. 1. Typical CNN architecture.

**Convolutional Layer:** In mathematics, convolution is an operation that describes the mixing of two information sources. When images are being convolved, the two information sources are the convolution filter (kernel) that is a matrix of (e.g.  $5 \times 5$  or  $3 \times 3$ ) used for edge detection, sharpening or any other image processing

algorithm, and the image pixels matrices, in which there are three that represent each color channel (RGB channels) or one if it is a grayscale image. The two sources convolve into a map of features after applying a dot product between them, as illustrated in Fig. 2.

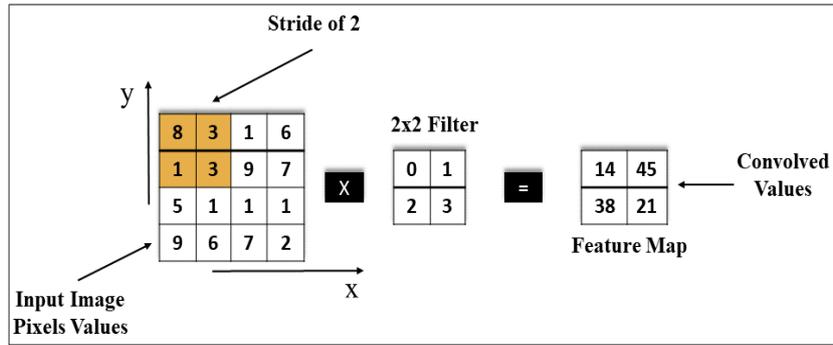


Fig. 2. Convolution operation.

A feature map is the result of the values that are convolved, which highlights an image feature, additionally multiple feature maps can be possessed by a convolutional layer that highlight more than one feature. A neuron's filter window (receptive field) moves over the image, contingent upon the size of the stride.

**ReLU:** Images possesses high non-linearity, and the used network needs to be capable of training on that nature. However, when convolution is applied, it causes an expansion in the input's linearity. Thus, to recover the non-linearity, an activation function must be used to convert the values of the input, as appeared in Equation 1:

$$y = f(\sum_{i=1}^m (a_i b_i) + c), m= 2, 3, 4... \quad \dots(1)$$

where (f) is an activation function and y is the output of the neuron. The value of the input in this equation comprise of a single layer perceptron, with (a<sub>i</sub>) being an input value, (b<sub>i</sub>) as the connector weight and c representing the value of the bias. This function plays a major role as it

increases the non-linearity in the network by multiplying it with the feature maps. This function was utilized in this research as it is capable of training the network at a rate that is higher than any activation function.

With ReLU function, the output value is (z) if the input (z) is positive and 0 if it is not, as it can be seen in Equation 2:

$$f(z) = \max(z, 0) \quad \dots(2)$$

where f(z) refers to the output of the activation function.

**Pooling:** The output of the convolutional layer (feature maps) is affected by an operation called pooling in which it reduces its dimensionality by decreasing the pixels count (as illustrated in Fig. 3). This process results in spatial variance reduction as it removes unimportant details, and this is accomplished by taking the receptive field's maximum value in case of max pooling, or the average value in case of average pooling, in which both of those operations are used in this research.

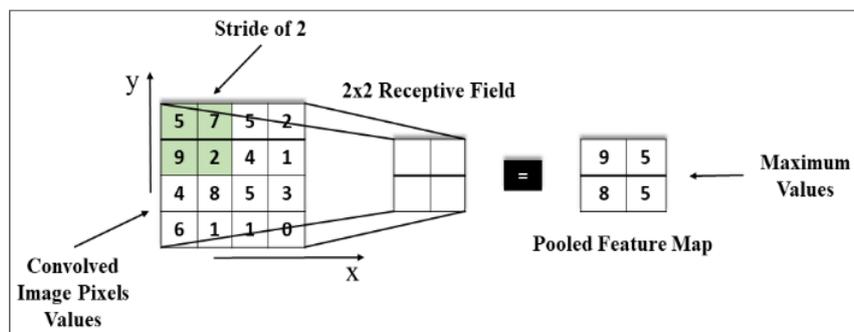


Fig. 3. Pooling Operation.

Pooling (equation 3) makes objects recognition much easier in spite of their location in the image, not to mention that diminishing the count of pixels implies there are less parameters to tune, which eliminates overfitting. Pooling is fairly like

convolution, in the sense that the size of filters, the type of padding and stride must be selected.

$$q(z) = \max z_{ab} , a, b \in R_{ij} \quad \dots(3)$$

where q(z) represents the pooling function and z<sub>ab</sub> is the value of the pixel on the a-th row and b-th column of the filter window (R<sub>ij</sub>). Max

pooling focuses on the pixels that are important; in which the pixels with high values are thought of as the ones that are highly activated. According to [31], max pooling showed superior performance when compared with average pooling. It has been utilized in many advanced models [32, 33, 34, 35].

**Global Average Pooling:** a global average pooling (GAP) is a type of pooling that is applied onto the dimensions of the tensor until every dimension is converted to one, in which; this process decreases the trainable parameters and reduces overfitting. A layer of GAP has been used in this research instead of a flattening layer, as it showed better performance.

## 2.1 Transfer Learning

To upscale the performance of a CNN, a number of modifications can be executed like increasing the number of hidden layers or neurons. But that action can cause an increase in the number of trainable parameters, which requires more data. The effect of data amount on the performance of the network is depicted in Fig. 4.

Although large-scale networks performance is superior to different techniques, it still contingent upon the size of training data. But, transfer learning on the other hand can achieve a

remarkable performance which can be compared custom CNNs using small datasets [36, 37]. By utilizing parameters (knowledge) gained from prior tasks that have a sufficient dataset, transfer learning can be used with insufficient data. Instead of using a custom CNN, which would result in issues with its performance. With this methodology, the primary  $m$  layers of the network that are already trained, are transferred to another network with untrained layers that would use the new target's data to train with. Generally, the transferred  $m$  layers are trained to be features extractors when applied on input data, in which they do not depend on the domain of the application.

Transfer learning is considered as a solution to the problem of small datasets, as recent studies proved that the main layers (pooling, ReLU and convolution) for a CNN, serve as a features extraction tool regardless of the target task, as for the rest of the layers (classification and sigmoid), they are associated with task [38, 39]. However, there were insufficient data to build a network from scratch for this research. Therefore, transfer learning approach was used to construct a classifier for the purpose of QI, to classify helical gears into damaged and healthy gears. The used trained model was the DenseNet121, which will be discussed in the following section.

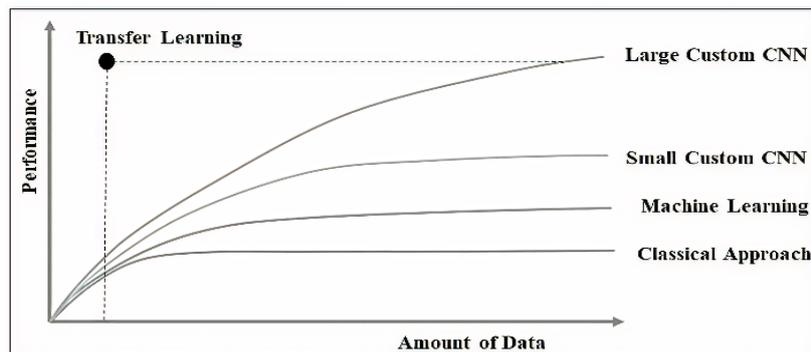


Fig. 4. Performance of learning approaches vs. data size.

## 2.3 Implemented Architecture

DenseNet (or dense convolutional neural network) is a huge CNN that comprises 121 layers, created by a group of analysts [40]. The layers in this network are connected in a feed forward arrangement, in order that each layer receives feature maps as inputs that came from previous layers. The feature maps from this layer are also employed as inputs to the following

layers. Regular CNNs with  $m$  layers have just  $m$  connections, in which there is one connection between every two layers, however DenseNet has  $m(m+1)/2$  straight connections. The design of the network include 121 layers shaped as 4 dense blocks, comprising convolutional and other layers between for information flow improvement among layers, as illustrated in Table 1 [40]. The trainable classification layer is also shown in the table.

**Table 1,  
Transferred and trainable layers.**

Stage	Layers
Transferred 1 (Conv.)	Convolution
Transferred 2 (Pooling)	Max Pooling
Transferred 3 (Dense Block 1)	Convolution × 6
Transferred 4 (Transition 1)	Convolution Average Pooling
Transferred 5 (Dense Block 2)	Convolution × 12
Transferred 6 (Transition 2)	Convolution Average Pooling
Transferred 7 (Dense Block 3)	Convolution × 24
Transferred 8 (Transition 3)	Convolution Average Pooling
Transferred 9 (Dense Block 4)	Convolution × 16
Transferred 10 (Classification)	Global Average Pooling Fully Connected
To be trained (Classification)	Global Average Pooling 256D Fully Connected 50 % Dropout 1D Sigmoid

### 3. Experimental Work

To collect the data, a phone camera was used to capture 4000 images with different backgrounds for training and validation (Fig. 5). The object (helical gear) was taken from the gearbox of an automobile's transmission unit. The gear serves as an important part for the gearbox dynamics because it transmits power between axles.

An open-source web based development environment known as Jupyter Notebook was used to develop the model and execute different data manipulation operations using Python.

After feeding the images into the notebook, they were converted to pixels grids that describe the red green blue (RGB) image, then decoded into three dimensional matrices (floating point tensors), because CNNs works with three dimensional matrices only. Since the DenseNet121 was trained with images of size (224 x 224) pixels, then it was only reasonable to convert the training images into that size. The final step in preprocessing was to rescale the values of the pixels to values between 0 and 1 by dividing by 255.

**Fig. 5. Training images.**

### 3.1 Model Training

Training refers to the process of optimizing the weights and biases of a network. Optimizers are employed to update the values of those biases and weights continuously until the global minimum point is reached. Adam optimizer was the employed optimizer for this research. To estimate the error within the network, binary cross-entropy function was used as the cost function:

$$\text{Error} = -\frac{1}{M} \sum_{x=1}^M (R \cdot \log(Y)) \quad \dots(4)$$

where M refers to the number of the model's training data, R is the goal value and Y is the output value (prediction). A (log) term used in the equation to provoke accentuation onto the right predictions.

Since insufficient data was used for this research, hyperparameters tuning was mandatory to avoid overfitting. 18 trials were performed using an application programming interface called Keras, moreover, dummy data was added to the training dataset by using data augmentation method, to upscale the data.

DL models consist of millions of parameters that requires a massive amount of calculations. So, to perform those calculations, a large computing power was required especially GPU power, as DL trains faster on GPUs than on CPUs, a CPU was also needed to perform data augmentation processes. The used GPU was NVIDIA GeForce GTX 1660Ti (6GB) and the CPU was Intel Core i7-9750H.

### 3.2 Model Evaluation

Following the training and validation stage, a final test stage is mandatory to confirm the success of the training. The system had to be tested to evaluate its performance with new data, in the sense that the model should be able to classify unseen gear images, as it was trained to identify the damaged gears from the ones in a healthy condition.

Therefore, to perform the evaluation, 30 colored images (RGB) were captured by 3 different cameras with differing resolutions of 24, 8 and 5 Mega-pixels, to test the model classification accuracy. Moreover, 12 grayscale images were also tested to explore the model versatility with images. The whole process was carried out in the same “Jupyter Notebook” environment, using Python to upload and initiate the model, in addition to building a function to convert the inserted images’ size to the standard DenseNet121’s (224 × 224) pixels.

## 4. Experimental Results

Experiments have been performed to select the proper model for transfer learning, and to show how it is more suitable for small datasets. Moreover, system assessment has also been done to explore its limitations and are all shown in the following sections.

### 4.1 Training Results

The trials results showed how using transfer learning can achieve a better performance than building a model from scratch, in terms of easier implementation. Moreover, several encountered issues were observed from the trials that could be discussed. Firstly, the issue of training time, which was resolved using more powerful hardware. And that is one of the drawbacks of DL, in which larger systems would require stronger GPUs to train at a reasonable rate. But this problem can still be avoided by using online cloud services like (Alibaba, AWS, Google Colab...etc.) that provide development environments.

The second issue is the high divergence between the losses of training and validation, which convey the significance of the occurring overfitting (as shown in Fig. 6), and also that the model is becoming stuck at local minimum points. This is a common issue when dealing with neural networks, especially with insufficient data, the network fails to generalize to the validation set which means the model will not be able to identify unseen information.

A bundle of solutions distributed separately and combined within trials were used to prevent overtraining. The effect of the used methods like data augmentation and dropout was observed throughout the trials, as the reduction of training parameters and the addition of dummy data contributed in minimizing the cost function.

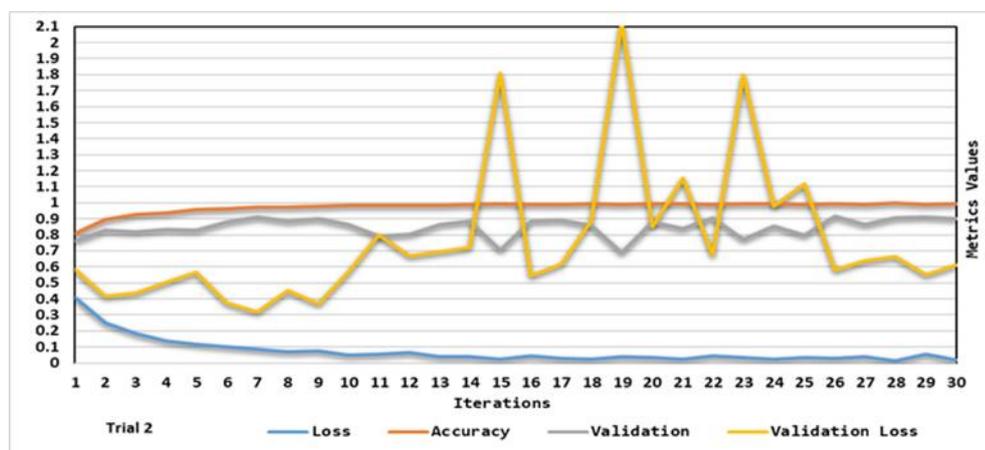


Fig. 6. Using VGG16 model in trial 2.

Last issue encountered was selecting the proper architecture. Pre-trained models that were used like Xception, VGG16 and ResNet50 failed to provide the desired performance; as not every pre-trained model can be used in transfer learning

for a custom target dataset. The proper model can either be selected by tediously testing several architectures or by using a selection technique like the one proposed in [41].

It was observed from trial 14 and up that tweaking the hyperparameters was improving the model, especially with the DenseNet121 model, as the model showed a noticeable enhancement, conveying the effect of the used architecture and the importance of selecting the proper one. Trial

18, reflected an outstanding performance with 98.43% validation accuracy and high convergence between the training and the validations losses (as shown in Fig. 7), that means that the network was able to converge towards the global minimum point.

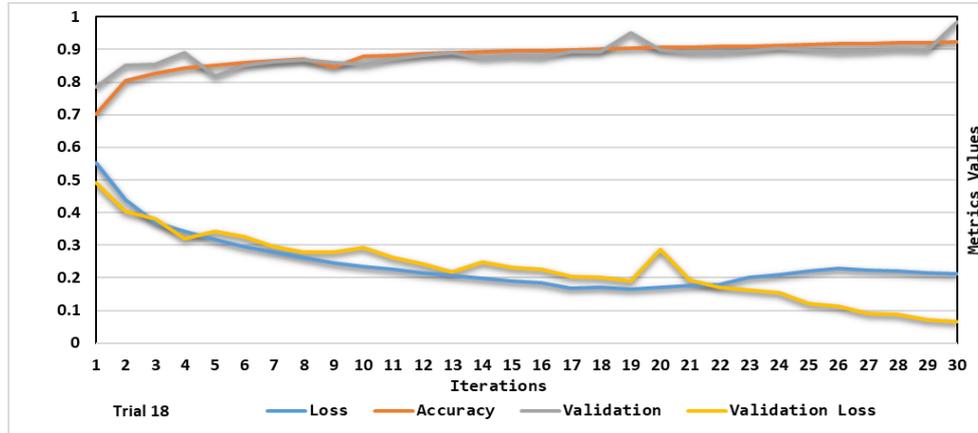


Fig. 7. Using DenseNet121 model in trial 18.

## 4.2 Evaluation Results

### a. Colored Images (RGB)

To evaluate the model, and test its ability to distinguish the damaged gears from the ones in a healthy condition, it was mandatory to be tested on unseen images; and observe how accurately it will perform.

Table 2, Illustrates the model evaluation results.

No.	24 Mega-pixels	8 Mega-pixels	5 Mega-pixels
1	0.9988	0.9961	0.9770
2	0.9786	0.9763	0.9957
3	0.9974	0.9697	0.9422
4	0.9951	0.9907	0.9602
5	0.9854	0.9936	0.9343
6	1.000	0.9995	0.9883
7	0.9931	1.000	0.9930
8	0.9959	0.9966	0.9914
9	0.9961	0.9911	0.9991
10	0.9801	0.9914	0.9985
Sum	9.9205	9.9050	9.7797
%	99.20	99.05	97.79

The accuracy results shown in Table 2, implies that the camera resolution had a very small effect on the system performance. Moreover, the model memorized and identified the object attributes with high certainty, which means low-resolution

cameras can be used for QI systems. Those cameras can be purchased at low costs.

### b. Grayscale Images

More images were tested for evaluation (table 3), but those images were in a black and white format, to see if the model is susceptible to color information.

Table 3, Shows a comparison between grayscale and RGB images.

Resolution	Metrics	Grayscale (%)	RGB (%)
24 mega-pixels	Accuracy	99.57	91.13
	Accuracy	99.78	99.71
	Accuracy	83.77	73.14
	Accuracy	93.04	90.15
8 mega-pixels	Accuracy	99.05	98.46
	Accuracy	95.28	94.69
	Accuracy	99.90	99.88
	Accuracy	98.12	97.63
5 mega-pixels	Accuracy	98.84	92.89
	Accuracy	95.42	84.82
	Accuracy	70.74	54.21
	Accuracy	78.46	61.08

Evaluation of the grayscale images showed outstanding results in the accuracy measures, as it can be seen from Table 3, when compared with the exact image but in RGB format. The accuracy

of the model increased with the black and white format with an average of (5.5% for the 24 mega-pixels, 0.42% for the 8 mega-pixels and 12.61% for the 5 mega-pixels). Moreover, with the last image in the 5 mega-pixel images' set, the model made a wrong prediction with the colored image, but when the same image was tested in the grayscale format, the model was able to make the right prediction with a rather higher accuracy.

According to [42] those results make good sense, as the classification accuracy of a CNN model depends largely on the image lighting, and colored images are susceptible to lighting. The effect of the lighting intensity can be seen in [43] where it was studied thoroughly along with other factors.

## 5. Conclusions

In this work, a quality inspection system based on deep convolutional neural networks was proposed to study the effect of utilizing DL with CV as an easier way to build an inspection system and more accurate way to classify objects at a low cost. The used object to be classified was a helical gear. It was planned to train the network to classify it according to its shape state, whether being damaged or not. The system was built using Keras framework. A dataset consisted of images of the gears in a healthy and defected state, were used as input for the system. Eighteen experiments were conducted using transfer learning and custom CNNs to build a high-performance system. Employing DenseNet121 with transfer learning yielded a high accuracy model.

The model was trained on the GPU of the computer and took an average of 4 hours for each experiment. The system performance was evaluated on 30 images dataset. They were captured at 3 different resolutions (24, 8 and 5 Mega-pixels). The system showed high accuracies of 99.20%, 99.05% and 97.79% respectively. In addition, it was evaluated with 12 grayscale images which showed results that were superior to their exact colored copies.

Experimental results concluded that using CNNs in quality inspection tasks produced some advantages over traditional computer vision techniques regarding accuracy, cost, feature extraction and implementation difficulty, in which high accuracy was achieved at low costs with automatic feature extraction ability and ease of implementation. At the same time, a disadvantage was also noticed about the CNNs, as they require

large datasets, but that issue was avoided by using transfer learning.

## Nomenclature

<i>AI</i>	Artificial Intelligence
$a_i$	Neuron Input Value
<i>ANN</i>	Artificial Neural Network
$b_j$	Connector Weight
$c$	Bias Value
<i>CNN</i>	Convolutional Neural Network
<i>CV</i>	Computer Vision
<i>CPU</i>	Central Processing Unit
<i>DL</i>	Deep Learning
<i>DNN</i>	Deep Neural Network
$f$	Activation Function
$f(z)$	Activation Function Output
$y$	Neuron Output
<i>RGB</i>	Red Green Blue
<i>GAP</i>	Global Average Pooling
<i>GPU</i>	Graphic Processing Unit
$M$	Number of Training Data
<i>QI</i>	Quality Inspection
$q(z)$	Pooling Function
$R$	Goal Value
<i>ReLU</i>	Rectified Linear Unit
$R_{ij}$	Filter Window
<i>SVM</i>	Support Vector Machine
$Y$	Model Prediction
$z$	Activation Function Input
$z_{ab}$	Pixel Value

## 6. References

- [1] M. P. Groover, "Inspection principles and practices", in Automation, Production Systems and Computer Integrated Manufacturing, 2nd ed. M. P. Groover, Eds. Harlow: Prentice Hall, 2001, pp. 681-682.
- [2] U. Mane, A. Mahajan, E. Kargutkar and K. Dhuri, "Detection of defects in plastic gears using image processing", IJISRT, vol. 2, issue 7, pp. 2456-2165, 2017.
- [3] Y. Qin, Y. Mao, B. Tang, Y. Wang and H. Chen, "M-band flexible wavelet transform and its application to the fault diagnosis of planetary gear transmission systems", Mechanical Systems and Signal Processing, vol. 134, issue 0888-3270, 2019.
- [4] M. Gao, G. Yu and T. Wang, "Impulsive gear fault diagnosis using adaptive morlet wavelet filter based on alpha-stable distribution and

- kurtogram," in IEEE Access, vol. 7, pp. 72283-72296, 2019.
- [5] J. Cai, "Gear fault diagnosis based on a new wavelet adaptive threshold de-noising method", Industrial Lubrication and Tribology, Vol. 71 No. 1, pp. 40-47, 2019.
- [6] T. Bettahar, C. Rahmoune, D. Benazzouz, B. Merainani, "New method for gear fault diagnosis using empirical wavelet transform, Hilbert transform, and cosine similarity metric", Advances in Mechanical Engineering, vol. 12, issue 6, 2020.
- [7] G. Yu, M. Gao and C. Jia, "A fast filtering method based on adaptive impulsive wavelet for the gear fault diagnosis", Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. February 2020.
- [8] X. Liu, H. Huang and J. Xiang, "A personalized diagnosis method to detect faults in gears using numerical simulation and extreme learning machine", Knowledge-Based Systems, vol. 195, issue 0950-7051, 2020.
- [9] P. Ong, T.H.C. Tieh, K. H. Lai, et al. "Efficient gear fault feature selection based on moth-flame optimisation in discrete wavelet packet analysis domain", J Braz. Soc. Mech. Sci. Eng., vol. 41, issue 266, 2019.
- [10] U. Urbas, D. Zorko, B. Černe, J. Tavčar and N. Vukašinović, "A method for enhanced polymer spur gear inspection based on 3D optical metrology", Measurement, issue 0263-2241, 2020.
- [11] K. Joshi, B. Patil, "Measurement of Spur Gear Parameters Using Machine Vision," Proceedings of International Conference on Intelligent Manufacturing and Automation. Lecture Notes in Mechanical Engineering. Springer, Singapore, 2020.
- [12] Y. Wu et al., "Detection of Gear Tooth Number and Common Normal Length Change Based on Computer Vision", 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Dublin, Ireland, 2019, pp. 618-621.
- [13] F. I. J. Ramírez and J. M. J. Barrionuevo, "Cyber-physical system for quality control of spur gears through artificial vision techniques", 2019 IEEE Fourth Ecuador Technical Chapters Meeting (ETCM), Guayaquil, Ecuador, 2019, pp. 1-6.
- [14] P. Kane and A. Andhare, "End of the assembly line gearbox fault inspection using artificial neural network and support vector machines", International Journal of Acoustics and Vibrations, vol. 24, issue 1, pp. 68-84, Mar 2019.
- [15] X. Zuo, X. Lei and X. Wang, "Research on machine vision measuring method for fine-pitch gears", Proc. SPIE 11343, Ninth International Symposium on Precision Mechanical Measurements, 2019.
- [16] P. Cao, S. Zhang and J. Tang, "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning", IEEE Access, vol. 6, pp. 26241-26253, May 2018.
- [17] L. Yu, X. Yao, J. Yang and C. Li, "gear fault diagnosis through vibration and acoustic signal combination based on convolutional neural network," Information, vol. 11, issue 6, pp. 266, 2020.
- [18] K. D. Joshi, V. Chauhan and B. Surgenor, "A flexible machine vision system for small part inspection based on a hybrid SVM/ANN approach", J Intell Manuf, vol. 31, pp. 103–125, 2020.
- [19] M. Y. Sallom, "Machine vision application in manufacturing: inspection of dimensions," The Iraqi Journal for Mechanical and Material Engineering, vol. 16, issue 3, 2016.
- [20] J. Wang, P. Fu and R. X. Gao, "Machine vision intelligence for product defect inspection based on deep learning and Hough transform", Journal of Manufacturing Systems, vol. 51, pp. 52-60, 2019.
- [21] M. Alencastre-Miranda, R. R. Johnson and H. I. Krebs, "Convolutional Neural Networks and Transfer Learning for Quality Inspection of Different Sugarcane Varieties", IEEE Transactions on Industrial Informatics, issue 1941-0050, pp. 1-1, 2020.
- [22] Y.J. Cruz, M. Rivas, R. Quiza, G. Beruvides and R. E. Haber, "Computer vision system for welding inspection of liquefied petroleum gas pressure vessels based on combined digital image processing and deep learning techniques", Sensors, vol. 20, issue 16, pp. 4505, 2020.
- [23] Y. Yang, L. Pan, J. Ma, R. Yang, Y. Zhu, Y. Yang and L. Zhang, "A high-performance deep learning algorithm for the automated optical inspection of laser welding," Appl. Sci., vol. 10, issue 3, pp. 933, 2020.
- [24] C. Beltrán-González, M. Bustero and A. Del Bue, "External and internal quality inspection of aerospace components", IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace), Pisa, Italy, 2020, pp. 351-355.

- [25] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, and L. FeiFei, "Imagenet: A large-scale hierarchical image database", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Florida, 2009, pp. 248-255.
- [26] G. Bonaccorso, Machine Learning Algorithms: Popular algorithms for data science and machine learning, 2nd ed. Birmingham: Packt Publishing, 2018.
- [27] N. Mahony, T. Murphy, K. Panduru, D. Riordan and J. Walsh. "Improving controller performance in a powder blending process using predictive control", Irish Signals and Systems Conference (ISSC), Killarney, 2017, pp. 1-6.
- [28] J. Walsh, N. Mahony, S. Campbell, A. Carvalho, L. Krpalkova, G. Velasco-Hernandez, et al. "Deep learning vs. traditional computer vision", Computer Vision Conference (CVC), Nevada, 2019, pp. 128-144.
- [29] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", Biol. Cybernetics, vol. 36, pp. 193-202, Apr 1980.
- [30] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MA: MIT press, 2016.
- [31] D. Scherer, A. Muller and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition", ICANN'10 Proceedings of the 20th International Conference on Artificial Neural Networks: Part III, Greece, 2010, pp. 92-101.
- [32] M.D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," European Conference on Computer Vision, Switzerland, 2014, pp. 818-833.
- [33] A. Krizhevsky, I. Sutskever and G.E. Hinton, "ImageNet classification with deep convolutional neural networks", Advances in neural information processing systems, vol. 25, issue 2, pp.1097-1105, Jan 2012.
- [34] A. Krizhevsky, "Convolutional deep belief networks on cifar-10," May 2010.
- [35] Y. LeCun, L.D. Jackel, L. Bottou, C. Cortes, J. Denker, H. Drucker et al. "Learning algorithms for classification: A comparison on handwritten digit recognition", Proc. 12th Int. Conf. Pattern Recognition and Neural Networks, Singapore, 1995, page 261-276.
- [36] C.K. Shie, C.H. Chuang, C.N. Chou, M.H. Wu and E.Y. Chang, 2015, August. "Transfer representation learning for medical image analysis", 37th Annual International Conference of the IEEE, Milan, 2015, pp. 711-714.
- [37] R. Zhang, H. Tao, L. Wu and Y. Guan, "Transfer learning with neural networks for bearing fault diagnosis in changing working conditions", IEEE Access, vol. 5, pp.14347-14357, June 2017.
- [38] M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks", 1st International Conference on Learning Representations, Scottsdale, 2013.
- [39] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," 2nd International Conference on Learning Representations, Banff, 2014.
- [40] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely connected convolutional networks", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Hawaii, 2017, pp. 2261-2269.
- [41] A. Meiseles and L. Rokach, "Source model selection for deep learning in the time series domain", IEEE Access, vol. 8, pp. 6190-6200, Jan 2020.
- [42] H. M. Bui, M. Lech, E. Cheng, K. Neville and I. S. Burnett, "Using grayscale images for object recognition with convolutional-recursive neural network," 2016 IEEE Sixth International Conference on Communications and Electronics (ICCE), Ha Long, 2016, pp. 321-325.
- [43] A. Najah, F. F. Mustafa and W. S. Hacham, "Effect of environmental factors on the accuracy of a quality inspection system based on transfer learning", Submitted to Al-Khwarizmi Engineering Journal.

## بناء نظام فحص نوعية عالي الدقة مبني على التعلم بالنقل بكاف قليلة

احمد نجاح\*  
فائز فوزي مصطفى\*\*  
وسام سعدي خضير حجام\*\*\*

\* قسم هندسة التصنيع المؤتمت/ كلية الهندسة الخوارزمي/ جامعة بغداد  
\*\* قسم هندسة الميكاترونكس/ كلية الهندسة الخوارزمي/ جامعة بغداد  
\*\*\* قسم هندسة الميكاترونكس/ كلية الهندسة الخوارزمي/ جامعة بغداد

البريد الإلكتروني: [faizalrawy@yahoo.com](mailto:faizalrawy@yahoo.com)  
البريد الإلكتروني: [ahmednajah5049@gmail.com](mailto:ahmednajah5049@gmail.com)  
البريد الإلكتروني: [wisam@kecbu.uobaghdad.edu.iq](mailto:wisam@kecbu.uobaghdad.edu.iq)

### الخلاصة

تعتبر مرحلة فحص جودة المنتجات من اهم المراحل في كل مسار إنتاجي، حيث يتم فيها تقدير جودة البضائع المنتجة ومقارنتها بالموصفات القياسية المطلوبة. تحدد الأدوات المستخدمة في الفحص وكيفية إجراء الفحص مدى تعقيد النظام. تعتمد وسائل الفحص التقليدية على الطرق اليدوية التي تتطلب تكاليف مختلفة واستهلاكاً كبيراً للوقت، على العكس من ذلك، فإن أنظمة الفحص الحالية التي تستخدم التقنيات الذكية الحديثة مثل رؤية الكمبيوتر، تكون أكثر دقة وكفاءة، ومع ذلك، فإن حجم العمل المطلوب لبناء نظام فحص تقليدي قائم على رؤية الكمبيوتر كبير جداً، نظراً لمشكلة اختيار واستخراج الخصائص يدوياً من الصور الرقمية، مما ينتج عنه أيضاً تكاليف العمالة لمهندسي النظام.

في هذا البحث، تم تقديم نهجاً متبنياً يعتمد على الشبكات العصبية التلافيفية لتصميم نظام لفحص الجودة بمستوى عالٍ من الدقة وتكلفة منخفضة. تم تصميم النظام باستخدام التعلم بالنقل لنقل الطبقات من نموذج تم تدريبه مسبقاً وشبكة عصبية متصلة بالكامل لتصنيف حالة المنتج إلى حالة صالحة أو تالفة. تم استخدام التروس الحلزونية كحالة دراسية وتم استخدام ثلاث كاميرات بدقة مختلفة لتقييم النظام. أظهرت النتائج التجريبية مستويات دقة عالية بتكاليف منخفضة واستخراج تلقائي للميزات.