

Article

Using Convolutional Neural Networks Based on a Taguchi Method for Face Gender Recognition

Cheng-Jian Lin ^{1,2,*} , Yu-Chi Li ¹ and Hsueh-Yi Lin ¹ 

¹ Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung 411, Taiwan; nancy19960222@gmail.com (Y.-C.L.); hyl@ncut.edu.tw (H.-Y.L.)

² College of Intelligence, National Taichung University of Science and Technology, Taichung 404, Taiwan

* Correspondence: cjlin@ncut.edu.tw

Received: 10 June 2020; Accepted: 27 July 2020; Published: 30 July 2020



Abstract: In general, a convolutional neural network (CNN) consists of one or more convolutional layers, pooling layers, and fully connected layers. Most designers adopt a trial-and-error method to select CNN parameters. In this study, an AlexNet network with optimized parameters is proposed for face image recognition. A Taguchi method is used for selecting preliminary factors and experiments are performed through orthogonal table design. The proposed method filters out factors that are significantly affected. Finally, experimental results show that the proposed Taguchi-based AlexNet network obtains 87.056% and 98.72% average accuracy of image gender recognition in the CIA and MORPH databases, respectively. In addition, the average accuracy of the proposed Taguchi-based AlexNet network is 1.576% and 3.47% higher than that of the original AlexNet network in CIA and MORPH databases, respectively.

Keywords: convolutional neural network; Taguchi method; parameter optimization; image recognition

1. Introduction

In recent years, face recognition has been widely used in various fields. Because the face is a very complex and important biological feature, it contains a lot of information, such as gender, age, and expression. Refs [1–4] Automatic gender and age analysis has produced many business model applications. For example, demographic data collection and product sales [5]. The intelligent facial recognition system can be based on the big data of the customer's gender and age, so that the seller can accurately and clearly understand the customer group that purchased this product. Therefore, this study focuses on gender classification analysis.

Machine learning has been successfully applied in gender recognition. Saatci et al. [4] proposed a method to identify the gender and expression of facial images through the active appearance model (AAM). Features extracted using trained AAM can be used to construct support vector machine (SVM) classifiers. These classifiers are arranged in a cascade structure to optimize the overall recognition performance. Mäkinen and Raisamo [5] conducted experiments by arithmetically combining the output of the gender classifier. This leads to improved classification accuracy. Singh et al. [6] used local binary pattern (LBP) and histogram of oriented gradient (HOG) feature extraction algorithms and SVM classifiers to obtain the more efficient gender classification from face images. Eidingner et al. [7] proposed the dropout-support vector machine approach for face attribute estimation in order to avoid over-fitting. This method was inspired by the popular dropout learning technology in deep belief networks. In addition, a robust face alignment technique is also proposed that explicitly considers the uncertainty of the face feature detector. Although the aforementioned methods can successfully achieve

gender recognition, the user needs to determine the features of gender images in advance. Therefore, this study will adopt a convolutional neural network (CNN) to overcome the above problems.

Recently, a CNN is a type of deep neural network that is commonly used to identify and detect visual images [8–11]. CNN has been widely used in various fields, such as object detection [12], face recognition [13,14], and speech recognition [15]. For the gender classification, Duan [8] proposed a hybrid of CNN and extreme learning machine (ELM) to incorporate the synergy of two classifiers for classifying gender images. Levi and Hassner [14] used deep-convolutional neural networks to learn representations. A significant increase in performance can be obtained on the gender classification. Ozbulak et al. [16] explored the portability of the existing CNN models for age and gender classification. A common AlexNet-like architecture and domain-specific VGG-Face CNN model were used and fine-tuned with the Adience data set to classify age and gender in an uncontrolled environment. In order to obtain better gender classification performance, users usually use the trial-and-error method to adjust the structural parameters of CNN during training process.

Many researchers [17–19] used various methods to optimize the structural parameters. Esteban et al. [17] proposed a genetic algorithm (GA) and deoxyribonucleic acid (DNA) based coding scheme to determine the CNN architecture parameters. Benteng et al. [18] proposed an autonomous and continuous learning (ACL) algorithm to generate automatically a deep convolutional neural network (DCNN) architecture for each given vision task. Zoph and Le [19] proposed neural architecture search (NAS) with reinforcement learning. The recurrent network architecture is used to generate neural networks. CNN is trained through reinforcement learning to maximize the expected accuracy of the architecture generated on the validation set. NAS is based on reinforcement learning and uses a total of 800 graphics processing unit (GPUs). This method has a very large search space and is also very time-consuming.

Competition is becoming increasingly fierce in various manufacturing sectors. To optimize product quality, Taguchi methods are usually used. Taguchi methods [20–22] were proposed by Dr. Genichi Taguchi from the 1950s to the early 1960s. Well known robust design methods now face competition from Taguchi quality engineering and Taguchi experimental design methods. A Taguchi method usually selects an appropriate orthogonal table based on the number of design parameters and the number of levels; an appropriately selected table can minimize the number of tests. The goal is to discover the best optimal product quality characteristics and maintain the stability of product quality. With Taguchi methods, the average loss of the product during the manufacturing process can be minimized.

To reduce the number of experiments and to optimize the parameters of a CNN architecture for gender image recognition, this study proposes a Taguchi-based convolutional neural network. Four well-known CNN architectures, LeNet [23], AlexNet [24], GoogLeNet [25], VggNet [26] and ResNet [27] are commonly used by researchers. In this study, we focus on the popular AlexNet network architecture, which is bigger than LeNet and smaller than GoogLeNet and ResNet. In addition, transfer learning [28] is suitable for large neural networks and small data sets. For the case where the data set itself is small, it is unrealistic to start training a large neural network with thousands of parameters because the generated model requires basic data. The problem of overfitting is unavoidable. Since this study does not use large neural networks, it does not use transfer learning in this study. Recently, researchers proposed to use transfer learning before training GoogLeNet, SqueezeNet, and ResNet50 network architectures [29]. These network architectures are relatively complex, so transfer learning is needed to avoid overfitting. In this study, using AlexNet and omitting transfer learning can also achieve high accuracy. To avoid using the trial-and-error method to determine the architectural parameters of the AlexNet network, we use a Taguchi method with an orthogonal table based on the numbers of design parameters and levels. In our experiments, the Taguchi-based AlexNet network recognized the genders of faces in the Computational Intelligence Application (CIA) and MORPH Facial Recognition database. The major contributions of this study are presented as follows:

- (1) To avoid using the trial-and-error method to determine the parameters of the network architecture, the Taguchi-based AlexNet network is proposed.
- (2) To avoid time-consuming for optimal parameters of the network, the orthogonal array is used. Without the orthogonal array, the 486 experimental number would be needed; however, there is 18 needed in this study.
- (3) Experimental results proved that the average accuracy rates of the proposed Taguchi-based AlexNet network is higher 1.576% and 3.47% than the original AlexNet network in CIA and MORPH databases, respectively.
- (4) Although this study only uses the Taguchi method to optimize the structural parameters of the AlexNet network, other networks can also adopt the proposed method to determine their network structure and obtain a better average accuracy.

The remainder of the paper is organized as follows. Section 2 introduces the Taguchi-based AlexNet network. The experimental results of the proposed method in CIA and MORPH databases are described in Section 3. Section 4 offers conclusions and suggestions for future work.

2. The Taguchi-Based AlexNet Network

This section presents a Taguchi-based CNN for detecting gender from an image data set. Figure 1 shows the flowchart of the proposed Taguchi-based AlexNet [24] architecture

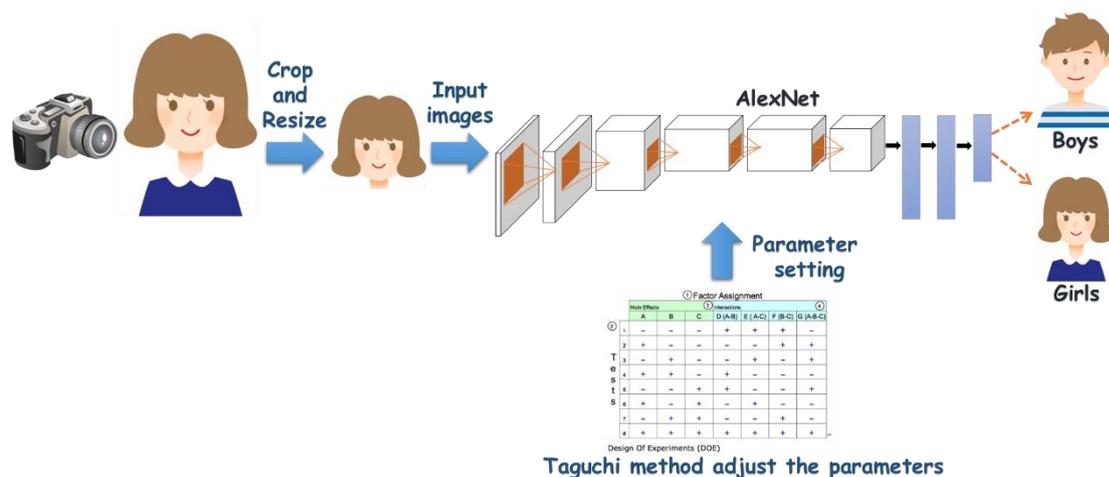


Figure 1. Flowchart of the proposed Taguchi-based AlexNet architecture.

2.1. Convolutional Neural Networks

Any major breakthrough in deep learning is relevant to CNNs. CNNs constitute the main force in the current development of deep neural networks; they recognize images even more accurately than humans. A CNN is a feed-forward neural network structure that responds to the pixel surrounding relevant image sections; CNNs excel at large image processing. LeCun et al. [23] have demonstrated that CNNs can have powerful image classification capabilities. CNNs mainly consists of three types of layers: convolutional layers, pooling layers, and fully connected layers. Convolutional layers and pooling layers are the most vital layers. A convolution layer is used to extract features by convolving an image area with multiple filters. Because its number of layers is increased, a CNN can more accurately parse the features of its input image. The pooling layer reduces the size of the convolution output mapping. If these two layers are managed properly, the number of parameters and the calculation complexity in a CNN can be minimized.

AlexNet is a classic model for image classification. As illustrated in Figure 2, the AlexNet architecture has five convolution layers (Conv), three pooling layers (Pool), and three fully connected

layers (each marked in the figure with FC and a number in Figure 2). Table 1 lists the parameters of each layer.

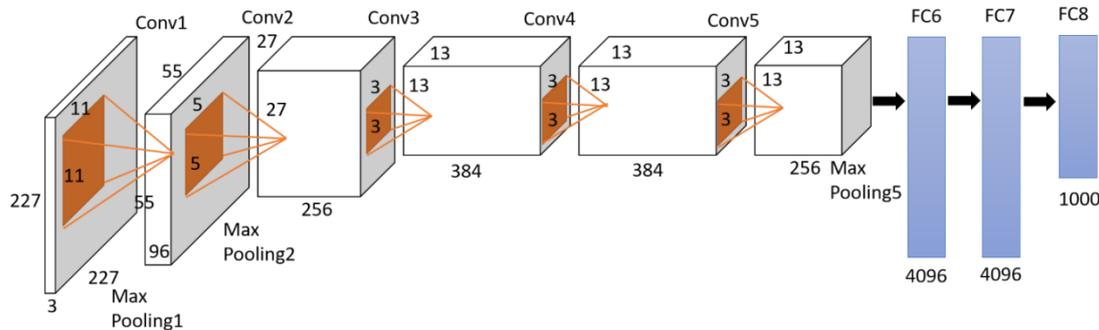


Figure 2. Architecture of AlexNet.

Table 1. Parameters for each layer of AlexNet.

	Conv-1	Pool-1	Conv-2	Pool-2	Conv-3	Conv-4	Conv-5	Pool-5
Kernel size (KS)	11 × 11	3 × 3	5 × 5	3 × 3	3 × 3	3 × 3	3 × 3	3 × 3
Stride (S)	4	2	1	2	1	1	1	2
Padding (P)	0	0	2	0	1	1	1	0

- Convolution layer

The convolution layer uses a specific feature detector called the convolution kernel to perform the convolution operation on the original image. The picture size is $227 \times 227 \times 3$ during training. In the first convolution layer, 96 number of $11 \times 11 \times 3$ convolution kernels are used for convolution calculation to generate new pixels. The input data of the second layer is the $27 \times 27 \times 96$ pixel layer output by the first layer. In order to facilitate subsequent processing, the upper, lower, left, and right edges of each pixel are filled with 2 pixels. The size of the convolution kernel size of the second layer is 5×5 , and the stride is 1 pixel, which is the same as the calculation formula of the first layer. The input data of the third layer is a set of $13 \times 13 \times 128$ pixel layers output by the second layer. For the convenience of subsequent processing, the upper, lower, left, and right edges of each pixel layer are filled with 1 pixel. The input data of the fourth layer is sets of $13 \times 13 \times 192$ pixel layers output by the third layer, similar to the third layer. The input data of the fifth layer is a set of $13 \times 13 \times 192$ pixel layers output by the fourth layer. For the convenience of subsequent processing, the upper, lower, left, and right edges of each pixel layer are filled with one pixel.

The KS, stride (S), and padding (P) in the convolution layer are used to determine the sizes of input and output images. Specifically, the calculation formula of the convolution layer.

$$O = \frac{(I - KS + 2P)}{S} + 1 \quad (1)$$

- Pooling layer

This study mainly uses max pooling, which picks the maximum value in the array. When the image is translated by several pixels, the max pooling does not affect the judgment at all. The major advantage of max pooling is to remove noise efficaciously. Max pooling layers are only after the first, second and fifth convolution layers.

- Fully connected layer

In AlexNet, there are three fully connected layers. 4096 convolution kernels are in the first and second fully connected layer, and the size of each convolution kernel is $6 \times 6 \times 256$. It is because the

size of the convolution kernel is exactly the same as the size of the feature map to be processed. Each coefficient in the convolution kernel is only multiplied by a pixel value of the size of the feature map, and one-to-one correspondence. Therefore, the size of the pixel layer after convolution is $4096 \times 1 \times 1$, that is, there are 4096 neurons. Because the number of categories is 2 in this study, 2 neurons in the third fully connected layer are set.

2.2. Proposed Taguchi Method

The proposed Taguchi method is a low-cost, high-efficiency quality engineering method. It emphasizes that product quality is improved by design rather than by inspection. In this study, the optimization parameters of Taguchi-based AlexNet are taken as the research object to explore the stability of AlexNet design. First, numerous factors are screened for the most pivotal factors to determine the optimal performance conditions. We use the screened factors as the design parameters of the Taguchi method. The quality of our objective function must be defined, and the factors and levels that affect the objective function must be identified. We use orthogonal arrays (abbreviated as OA in Figure 3) to determine the number of experiments and the factor allocations of experiments. It is expected that reference data equivalent to the full factor can be obtained with a few experiments. Finally, the key factor levels must be determined to verify that the Taguchi-based AlexNet parameters are optimal. This experiment considers cost-effectiveness to discover the best combination of factor levels. A flowchart of the Taguchi method is depicted in Figure 3. The detailed process of the Taguchi method is explained as follows:

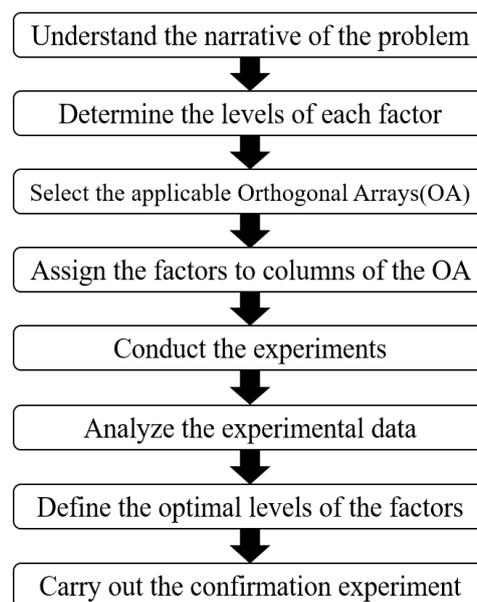


Figure 3. Flow chart of the Taguchi method.

- Step 1: Understand the narrative of the problem

To understand the tasks need to be completed. To achieve simplified experiments and a good accuracy, this study adopts the Taguchi method to adjust the parameters of the convolutional neural network.

- Step 2: Determine the levels of each factor

This study uses the Taguchi method to determine the structural parameters of the AlexNet network. AlexNet has five convolutional layers. Many researchers suggest that the convolutional layers of LeNet or AlexNet have the highest impact [30,31]. Therefore, we also use the first and

fifth convolution layers of the AlexNet network as the structural parameters. In the first and fifth convolution layers, the structural parameters include the kernel size, stride, and padding. In this study, there are six affecting factors in total. Except for the two levels used for the padding of the first convolutional layer, three levels are selected among these affecting factors.

- Step 3: Select the applicable Orthogonal Arrays (OA)

Before selecting the OA, the number of affecting factors and levels must be considered. According to the selected levels and factors, a suitable OA is selected. In this study there is one factor of two levels and five factors of three levels. So that L18 ($2^1 \times 3^5$) OA is selected. Up to 6 factors can be configured, with 3 and 2 representing the number of levels.

- Step 4: Assign the factors to columns of the OA

The purpose of using the OA is to obtain the desired result with a very small number of experiments. For example, when parameter design only considers the main effect and does not consider the interaction between various factors. After the configuration of the factors and levels of the OA is completed, the experiment can be conducted according to the experimental combination configured in the OA to obtain the optimal level combination. This experiment selects 3 levels and 2 levels for each of the selected 6 factors, so an L18 orthogonal array is used for the experiment. The factors KS of conv-1, S of conv-1, P of conv-1, KS of conv-5, S of conv-5, and P of conv-5 are arranged in rows 1, 2, 3, 4, 5 and 6, and shown in Table 2.

Table 2. L18 ($2^1 \times 3^5$) orthogonal arrays.

	KS of Conv-1	S of Conv-1	P of Conv-1	KS of Conv-5	S of Conv-5	P of Conv-5
1	13	4	2	7	3	2
2	13	3	2	5	2	1
3	13	2	2	3	1	0
4	11	4	2	7	2	1
5	11	3	2	5	1	0
6	11	2	2	3	3	2
7	9	4	2	5	3	0
8	9	3	2	3	2	2
9	9	2	2	7	1	1
10	13	4	1	3	1	1
11	13	3	1	7	3	0
12	13	2	1	5	2	2
13	11	4	1	5	1	2
14	11	3	1	3	3	1
15	11	2	1	7	2	0
16	9	4	1	3	2	0
17	9	3	1	7	1	2
18	9	2	1	5	3	1

- Step 5: Conduct the experiments

According to the L18 ($2^1 \times 3^5$) OA and conduct the experiments.

- Step 6: Analyze the experimental data

Through the experimental factor level planning of the Taguchi experiment design, the observation values obtained from the experiment are carried out. It is decided that the number of repeats for each experiment is 5 times. Calculate the Signal-to-noise ratio (S/N ratio), and then analyze the experimental data.

- Step 7: Define the optimal the levels of the factors

According to the Step 6 results define the optimal the levels of the factors.

- Step 8: Carry out the confirmation experiment

Finally, carry out the confirmation experiment. Use the optimal parameter combination to train the AlexNet.

3. Experimental Results

In this experiment, there are two data sets. First is the CIA data set and the second is the MORPH database. Our trials randomly selected 70% of the data set for training and 30% of the data for testing. All experiments are performed on Ubuntu 16.04 with a 3.30 GHz Intel® Xeon® E3-1225 v5(4C4T) processor and Nvidia GTX 1080 8GB (MSI, New Taipei City, Taiwan).

3.1. AlexNet

The architecture of AlexNet consists of eight layers. The first to fifth layers are convolutional layers for convolution operation and pooling; the sixth to eighth layers are fully connected layers. In this experiment, the first and fifth convolutional layers are used for structural parameter optimization. Figure 4 shows the structural parameter optimization diagram of the AlexNet. Because the main purpose of the convolutional layer is to capture the characteristics of the image, it affects the quality of the feature map and the accuracy of recognition. Therefore, this paper uses the Taguchi method to optimize the structural parameters of the convolution layer has obtained better characteristics. As for the fully connected neural network, we adopt the same as the original AlexNet.

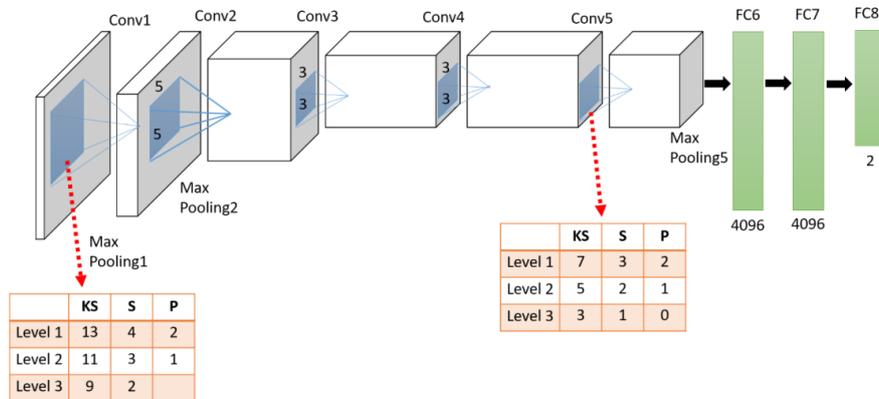


Figure 4. Diagram of structural parameter optimization in the AlexNet.

3.2. Computational Intelligence Application (CIA) Data Set

To evaluate the proposed Taguchi-based AlexNet network, CIA data set is used. The CIA data set is a small face database collected by our laboratory. The database images depict male and female persons in the age range of 6 to 80 years, as illustrated in Figure 5. This data set includes 1880 face pictures and comprises male and female subjects. Most subjects are East Asian in this data set.

3.2.1. Determine the Levels of Each Factor

In this experiment, only the convolution kernel parameters are considered in the first and fifth convolutional layers. The affecting factors selected in AlexNet include KS, stride (S), and padding (P) of the convolution kernel in the first and fifth convolutional layers. The system had six affecting factors in total. The levels of key factors are set by referring to the results of experts and actual experiments in Table 3.

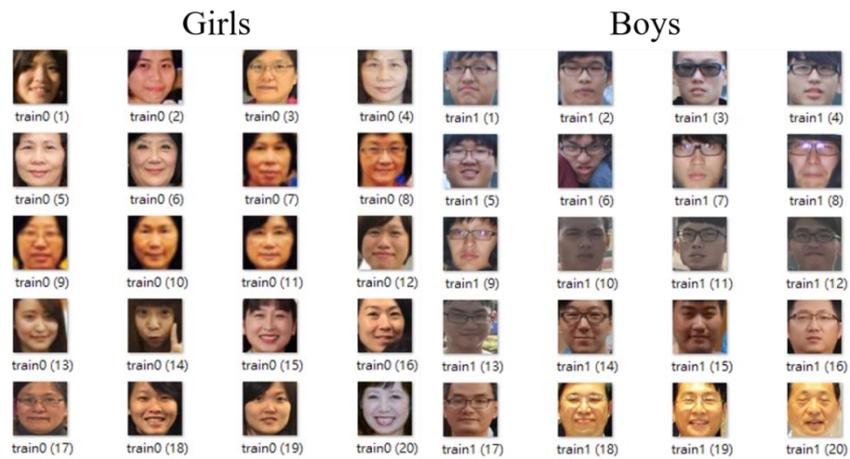


Figure 5. Computational Intelligence Application (CIA) data set.

Table 3. Control factors and levels.

	KS of Conv-1	S of Conv-1	P of Conv-1	KS of Conv-5	S of Conv-5	P of Conv-5
Level 1	13	4	2	7	3	2
Level 2	11	3	1	5	2	1
Level 3	9	2		3	1	0

3.2.2. Factor Allocation of Taguchi Experiments

This experiment is based on five previously selected factors with three experimental levels and one factor with two experimental levels. The total number of degrees of freedom is 12. Therefore, the L18 ($2^1 \times 3^5$) orthogonal array is used. The factors, KS of conv-1, S of conv-1, P of conv-1, KS of conv-5, S of conv-5, and P of conv-5, are configured in columns 2, 3, 4, 5, 6, and 7, as shown in Table 4. The selected factors and levels originally require $2^1 \times 3^5 = 486$ experiments, and only 18 experiments after using the Taguchi experiment.

Table 4. L18 ($2^1 \times 3^5$) orthogonal array.

	KS of Conv-1	S of Conv-1	P of Conv-1	KS of Conv-5	S of Conv-5	P of Conv-5
1	13	4	2	7	3	2
2	13	3	2	5	2	1
3	13	2	2	3	1	0
4	11	4	2	7	2	1
5	11	3	2	5	1	0
6	11	2	2	3	3	2
7	9	4	2	5	3	0
8	9	3	2	3	2	2
9	9	2	2	7	1	1
10	13	4	1	3	1	1
11	13	3	1	7	3	0
12	13	2	1	5	2	2
13	11	4	1	5	1	2
14	11	3	1	3	3	1
15	11	2	1	7	2	0
16	9	4	1	3	2	0
17	9	3	1	7	1	2
18	9	2	1	5	3	1

3.2.3. Accuracy and S/N Ratio Results

Because of the experimental factor level planning of the Taguchi experiment design, the quality characteristics of this study are larger the better (LTB). A cross-validation experiment is performed and tested. The S/N ratio of the experimental response value is calculated. The experimental accuracy values were calculated with five rounds of cross-validation, and the S/N ratios are listed in Table 5. This study does not consider the interaction between these factors. The LTB is calculated as follows.

$$S/N_{LTB} = -10 \log_{10} \frac{1}{t} \sum_{i=1}^t \frac{1}{a_i^2} \tag{2}$$

where t is the number of repeated experiments and a_i is the i th accuracy of repeated experiments. The larger S/N ratio presents a smaller standard deviation and illustrates higher stability for the quality.

Table 5. Accuracy and S/N ratio of AlexNet in CIA data set.

NO.	Cross-Validation 1	Cross-Validation 2	Cross-Validation 3	Cross-Validation 4	Cross-Validation 5	S/N Ratio
1	83.69	89.18	87.41	89.01	87.23	-1.18618
2	84.93	83.51	84.57	85.11	87.23	-1.40709
3	87.23	86.88	86.88	84.93	85.28	-1.28739
4	85.46	84.22	84.93	88.48	85.99	-1.33231
5	85.99	87.23	85.11	83.33	85.99	-1.36063
6	81.91	82.09	85.11	83.16	83.51	-1.6046
7	83.16	85.64	84.4	86.17	87.06	-1.38582
8	87.23	86.17	82.45	85.82	85.28	-1.37656
9	88.83	84.4	86.17	84.93	84.93	-1.32939
10	86.17	86.88	84.22	85.64	87.23	-1.3092
11	87.59	85.11	85.46	84.22	82.8	-1.41238
12	86.17	85.11	84.75	85.28	84.22	-1.40154
13	88.3	85.64	84.93	85.28	87.59	-1.27808
14	84.04	83.87	84.57	81.38	87.06	-1.50144
15	87.23	85.64	85.99	84.4	82.09	-1.41003
16	84.75	83.51	84.4	85.11	84.4	-1.47017
17	85.64	89.01	85.99	87.77	85.28	-1.23927
18	82.62	84.75	83.33	84.04	85.28	-1.51569

3.2.4. Quality Characteristics and Optimal Parameters

The average S/N ratio of each factor and level is listed in Table 5. The higher S/N ratio represents the higher stability of the quality. The optimal parameter combination is displayed in the last row in Table 6. That is, the best parameter combination has levels 1, 1, 1, 3, and 1 for KS of conv-1, S of conv-1, P of conv-1, KS of conv-5, S of conv-5, and P of conv-5, respectively. In addition, the S/N ratio effect of each factor is depicted in Figure 6. The first three significant factors are S of conv-5, KS of conv-5, and S of conv-1.

Table 6. Average S/N ratio of each factor and level in CIA data set.

Factors Levels	KS of Conv-1	S of Conv-1	P of Conv-1	KS of Conv-5	S of Conv-5	P of Conv-5
1	-1.33396	-1.32696	-1.36333	-1.31826	-1.43435	-1.3477
2	-1.41451	-1.38289	-1.39309	-1.39147	-1.39962	-1.39919
3	-1.38615	-1.42477	-	-1.42489	-1.30066	-1.38774
Delta	0.08055	0.097813	0.029757	0.106632	0.13369	0.051481
Rank	4	3	6	2	1	5
Best levels	1	1	1	1	3	1
Best factors	13	4	2	7	1	2

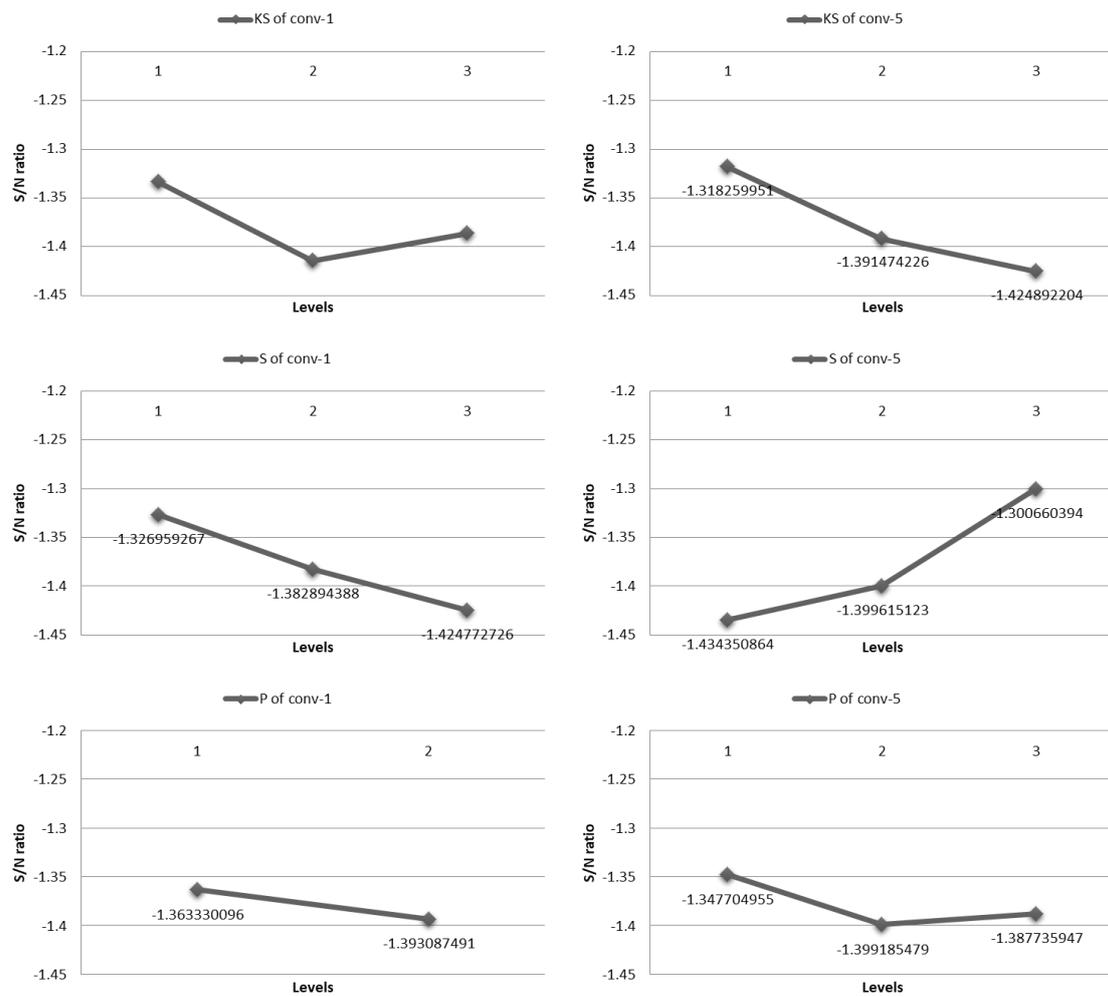


Figure 6. Effects of all factors in the CIA data set.

3.2.5. Optimal Combination of Experimental Parameters

In light of the experimental quality results, the final experiment is performed based on the optimal parameter combination. Table 7 lists five cross-validation experiments and the average accuracy of the optimal combination for experiments in the proposed Taguchi-based AlexNet network. In addition, the original AlexNet network is tested with the CIA data set and performance levels are compared. And Table 7 lists the accuracy of the original AlexNet network. The average accuracy values of the proposed Taguchi-based AlexNet network and the original AlexNet network are 87.056% and 85.48%, respectively. Experimental results prove that the proposed Taguchi-based AlexNet network has a better average accuracy than the original AlexNet network.

Table 7. Accuracy of the optimal parameter combination in CIA data set.

Optimal Parameter Combination	Cross-Validation 1	Cross-Validation 2	Cross-Validation 3	Cross-Validation 4	Cross-Validation 5	Average Accuracy
Original AlexNet network	85.99	84.22	86.11	84.93	86.17	85.48
Taguchi-based AlexNet network	86.88	86.93	86.17	87.89	87.41	87.056

3.3. MORPH Facial Recognition Database

MORPH database is a large database and contains the classes of age, gender and ethnicity. In this study, we focus on gender classification. MORPH database contains about 55,000 facial images, including about 46,660 boys and 8493 girls. In this experiment, the same reference to the CIA experiment only considers the convolution kernel parameters in the first and fifth convolutional layers. The selected affecting factors include the KS, convolution (S) and padding (P) of the convolution kernels in the first and fifth convolutional layers. There are six key factors. Therefore, the level and distribution of each factor in the Taguchi experiment are the same as those in the CIA data experiment.

3.3.1. Results of Accuracy and S/N Ratio

The experimental accuracy values are calculated with five rounds of cross-validation, and the S/N ratios are listed in Table 8. This study also does not consider the interaction between these factors.

Table 8. Accuracy and S/N ratio of AlexNet in MORPH database.

NO.	Cross-Validation 1	Cross-Validation 2	Cross-Validation 3	Cross-Validation 4	Cross-Validation 5	S/N Ratio
1	97.96	98.1	97.96	97.99	98.02	-0.17495
2	98.11	97.79	97.07	97.89	97.74	-0.2005
3	97.79	97.88	97.82	97.55	97.68	-0.19822
4	97.09	97.99	98.21	98.01	98.13	-0.18581
5	97.14	97.68	97.76	97.49	97.08	-0.22625
6	97.7	97.87	97.39	97.33	97.15	-0.22107
7	97.32	98.16	97.62	97.93	98.26	-0.18824
8	97.69	97.91	97.75	97.75	98.24	-0.18724
9	97.87	97.91	98.28	97.94	98.31	-0.17003
10	97.7	98.01	97.76	97.19	97.66	-0.20541
11	97.92	97.86	97.76	97.67	97.49	-0.19858
12	97.91	96.96	97.44	97.56	96.89	-0.2333
13	97.42	97.83	98.16	97.61	97.87	-0.19526
14	97.59	97.46	97.1	97.76	97.38	-0.22372
15	97.9	97.67	97.74	97.77	97.09	-0.20809
16	97.62	97.49	97.24	97.66	97.85	-0.21355
17	97.18	98.22	98.15	98.39	98.1	-0.17502
18	97.65	97.59	97.6	96.69	97.85	-0.22227

3.3.2. Quality Characteristics and Optimal Parameters

The average S/N ratio of each factor and level are listed in Table 8. The optimal parameter combination is displayed in the last row in Table 9. That is, the best parameter combination has levels 3, 1, 1, 1, 3, and 1 for KS of conv-1, S of conv-1, P of conv-1, KS of conv-5, S of conv-5, and P of conv-5, respectively. In addition, the S/N ratio effect of each factor is depicted in Figure 7. The first three significant factors are KS of conv-5, KS of conv-1, and S of conv-1.

Table 9. Average S/N ratio of each factor and level in MORPH database.

Factors Levels	KS of Conv-1	S of Conv-1	P of Conv-1	KS of Conv-5	S of Conv-5	P of Conv-5
1	-0.20183	-0.19387	-0.1947	-0.18541	-0.20481	-0.19781
2	-0.21003	-0.20188	-0.20836	-0.21097	-0.20475	-0.20129
3	-0.19273	-0.20883	-	-0.2082	-0.19503	-0.20549
Delta	0.017308	0.01496	0.013654	0.025556	0.009774	0.007681
Rank	2	3	4	1	5	6
Best levels	3	1	1	1	3	1
Best factors	9	4	2	7	1	2

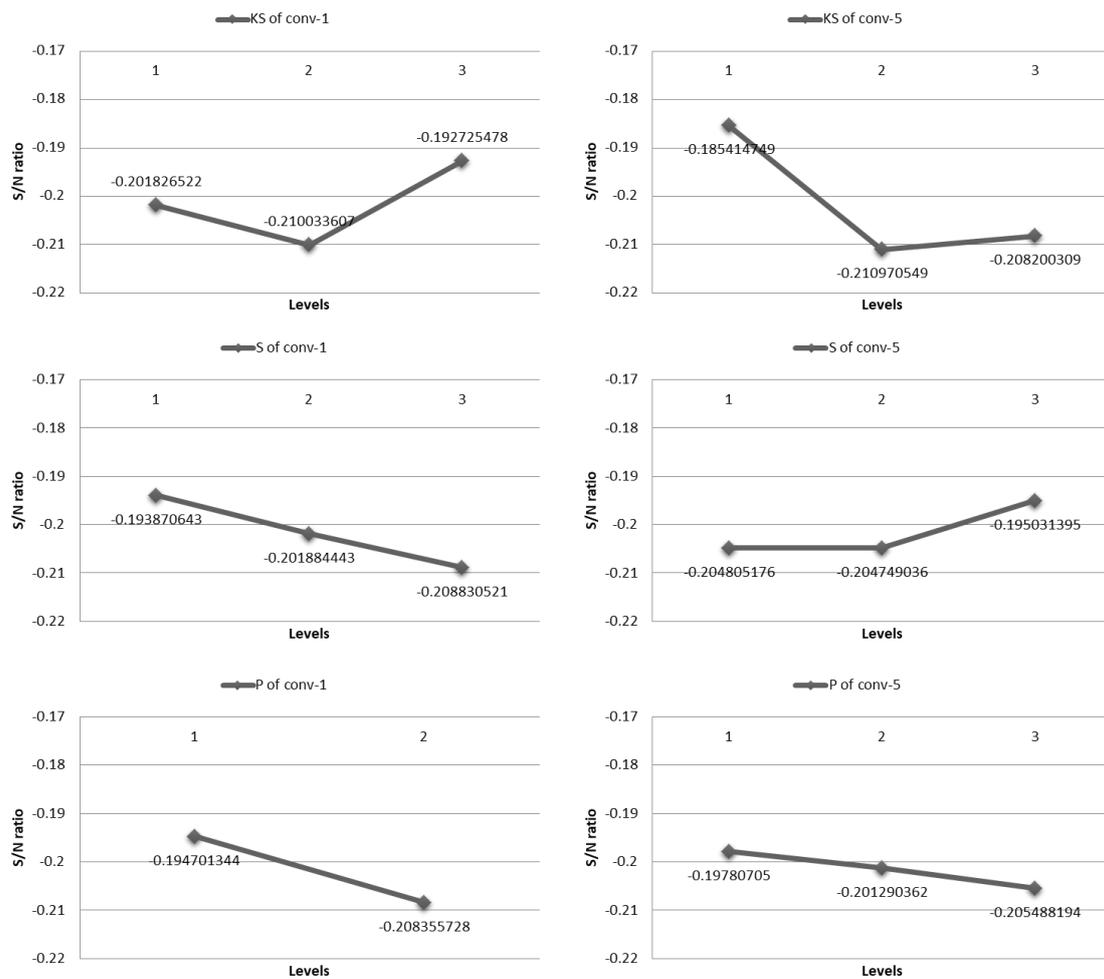


Figure 7. Effects of all factors in the MORPH database.

3.3.3. Optimal Combination of Experimental Parameters

This experiment is performed by using the optimal parameter combination. Table 10 lists the accuracy of the original AlexNet network. The average accuracy values of the proposed Taguchi-based AlexNet network and the original AlexNet network are 95.25% and 98.72%, respectively. Experimental results again prove that the proposed Taguchi-based AlexNet network has a better average accuracy than the original AlexNet network. Besides, Table 11 also shows the comparison results of accuracy and training time using various methods in the MORPH database. The results show that the proposed method not only has higher accuracy than other methods, but also has a shorter training time.

Table 10. Accuracy of the optimal parameter combination in the MORPH database.

Optimal Parameter Combination	Cross-Validation 1	Cross-Validation 2	Cross-Validation 3	Cross-Validation 4	Cross-Validation 5	Average Accuracy
Original AlexNet network	94.53	94.58	95.98	95.75	95.45	95.25
Taguchi-based AlexNet network	98.99	98.65	98.78	99.08	98.1	98.72

Table 11. Comparison results of accuracy and training time using various methods in the MORPH database.

Optimal Parameter Combination	Taguchi-Based AlexNet Network	LeNet [14]	ResNet [18]	GoogLeNet [16]	CNN-ELM + Dropout 0.5 [6]	CNN-ELM + Dropout 0.7 [6]
Accuracy	98.72%	96.30%	98.15%	97.48	87.30%	88.20%
Training time	40 min 23s	41 min 9s	1 h 2 min 52s	1 h 33 min 8s	-	-

4. Conclusions

Generally, the adjustment of CNN parameters is time-consuming. In this study, a Taguchi-based AlexNet network is proposed for optimizing the parameters of an AlexNet architecture. To avoid using the trial-and-error method to determine the network parameters of the AlexNet, a Taguchi method is used. The Taguchi method has an orthogonal array based on the number of design parameters and levels. Without the orthogonal array, 486 experimental number would be needed; however, there are 18 needed in this study. According to the average S/N ratio of each factor and level obtained, S of conv-5 and KS of conv-5 are significant factors. Experimental results proved that the average accuracy rates of the proposed Taguchi-based AlexNet network is higher 1.576% and 3.47% than the original AlexNet network in CIA and MORPH databases.

In future work, we will use the proposed Taguchi-based AlexNet network to realize practical hardware applications in embedded systems. An extensible open computing platform NVIDIA DRIVE™ AGX is used. It is like an artificial intelligence brain that can provide industry-leading performance and energy-efficient computing. In order to develop and produce functional safety and adopt artificial intelligence technology, we join this hardware and are able to practically apply various systems including this gender classification system. In addition, the experimental results prove that, in CIA and MORPH databases, the average accuracy of the Taguchi-based AlexNet network is higher than the original AlexNet network. Therefore, this method will be extended to LeNet, GoogLeNet and ResNet in the future.

Author Contributions: Conceptualization, C.-J.L.; methodology, C.-J.L. and Y.-C.L.; software, Y.-C.L. and H.-Y.L.; data curation, Y.-C.L.; writing—original draft preparation, C.-J.L., Y.-C.L. and H.-Y.L.; funding acquisition, C.-J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Technology of the Republic of China, grant number MOST 109-2221-E-167-027.

Conflicts of Interest: The authors declare no conflict of interests regarding the publication of this paper.

References

1. Erik, H.; Boon, K.L. Face detection: A survey. *Comput. Vis. Image Underst.* **2001**, *83*, 236–274. [[CrossRef](#)]
2. Yang, M.H.; Kriegman, D.J.; Ahuja, N. Detecting faces in images: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 34–58. [[CrossRef](#)]
3. WenYi, Z.; Rama, C.; Jonathon, P.; Azriel, R. Face recognition: A literature survey. *ACM Comput. Surv.* **2003**, *35*, 399–458. [[CrossRef](#)]
4. Saatci, Y.; Town, C. Cascaded classification of gender and facial expression using active appearance models. In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR06), Southampton, UK, 10–12 April 2006; pp. 393–398. [[CrossRef](#)]
5. Mäkinen, E.; Raisamo, R. An experimental comparison of gender classification methods. *Pattern Recognit. Lett.* **2008**, *29*, 1544–1556. [[CrossRef](#)]
6. Singh, V.; Shokeen, V.; Singh, B. Comparison of feature extraction algorithms for gender classification from face images. *Int. J. Eng. Res. Technol.* **2010**, *2*, 61034767.
7. Eidinge, E.; Enbar, R.; Hassner, T. Age and Gender Estimation of Unfiltered Faces. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 2170–2179. [[CrossRef](#)]

8. Duan, M.; Li, K.; Yang, C.; Li, K. A hybrid deep learning CNN–ELM for age and gender classification. *Pattern Recognit. Lett.* **2018**, *275*, 448–461. [[CrossRef](#)]
9. Kumar, A.; Kim, J.; Lyndon, D.; Fulham, M.; Feng, D. An Ensemble of Fine-Tuned Convolutional Neural Networks for Medical Image Classification. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 31–40. [[CrossRef](#)]
10. Ševo, I.; Avramović, A. Convolutional Neural Network Based Automatic Object Detection on Aerial Images. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 740–744. [[CrossRef](#)]
11. Ponce, J.M.; Aquino, A.; Andújar, J.M. Olive-Fruit Variety Classification by Means of Image Processing and Convolutional Neural Networks. *IEEE Access* **2019**, *7*, 147629–147641. [[CrossRef](#)]
12. Ding, C.; Tao, D. Trunk-Branch Ensemble Convolutional Neural Networks for Video-Based Face Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 1002–1014. [[CrossRef](#)] [[PubMed](#)]
13. Rasti, P.; Uiboupin, T.; Escalera, S.; Anbarjafari, G. Convolutional Neural Network Super Resolution for Face Recognition in Surveillance Monitoring. In Proceedings of the 2016 International Conference on Articulated Motion and Deformable Objects, Palma de Mallorca, Spain, 13–15 July 2016; pp. 175–184.
14. Levi, G.; Hassner, T. Age and gender classification using convolutional neural networks. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, USA, 7–12 June 2015; pp. 34–42.
15. Hou, J.; Wang, S.; Lai, Y.; Tsao, Y.; Chang, H.; Wang, H. Audio-Visual Speech Enhancement Using Multimodal Deep Convolutional Neural Networks. *IEEE Trans. Emerg. Top. Comput. Intell.* **2018**, *2*, 117–128. [[CrossRef](#)]
16. Ozbulak, G.; Aytar, Y.; Ekenel, H.K. How transferable are CNN-based features for age and gender classification. In Proceedings of the 2016 International Conference of the Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany, 21–23 September 2016. [[CrossRef](#)]
17. Esteban, R.; Sherry, M.; Andrew, S.; Saurabh, S.; Yutakaleon, S.; Jie, T.; Quoc, V.L.; Alexey, L.K. Large-scale evolution of image classifiers. In Proceedings of the ICML'17: 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; Volume 70, pp. 2902–2911. [[CrossRef](#)]
18. Benteng, M.; Xiang, L.; Yong, X.; Yanning, Z. Autonomous deep learning: A genetic DCNN designer for image classification. *Neurocomputing* **2020**, *379*, 152–161. [[CrossRef](#)]
19. Zoph, B.; Le, Q.V. Neural Architecture Search with Reinforcement Learning. *arXiv*, **2016**; arXiv:1611.01578.
20. Yang, W.H.; Tang, Y.S. Design optimization of cutting parameters for turning operations based on the Taguchi method. *J. Mater. Process. Tech.* **1998**, *84*, 122–129. [[CrossRef](#)]
21. Ballantyne, K.N.; van Oorschot, R.A.; Mitchell, R.J. Reduce optimisation time and effort: Taguchi experimental design methods. *Forensic Sci. Int. Genet. Suppl. Ser.* **2008**, *1*, 7–8. [[CrossRef](#)]
22. Harlapur, M.D.; Mallapur, D.G.; Rajendra, K. *AIP Conference Proceeding*; AIP Publishing: New York, NY, USA, 2018; Volume 1943.
23. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
24. Alex, K.; Ilya, S.; Geoffrey, H. ImageNet classification with deep convolutional neural networks. *Neural Inf. Process. Syst.* **2012**, *60*, 84–90. [[CrossRef](#)]
25. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9. [[CrossRef](#)]
26. Simonyan, K.; Zisserman, A. Very deep convolutional neural network based image classification using small training sample size. In Proceedings of the 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 3–6 November 2015; pp. 730–734. [[CrossRef](#)]
27. Kaiming, H.; Xiangyu, Z.; Shaoqing, R.; Jian, S. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
28. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [[CrossRef](#)]
29. Islam, M.M.; Tasnim, N.; Baek, J.-H. Human Gender Classification Using Transfer Learning via Pareto Frontier CNN Networks. *Inventions* **2020**, *5*, 16. [[CrossRef](#)]

30. Lin, C.J.; Lin, C.H.; Jeng, S.Y. Using feature fusion and parameter optimization of dual-input convolutional neural network for face gender recognition. *Appl. Sci.* **2020**, *10*, 3166. [[CrossRef](#)]
31. Lin, C.J.; Jeng, S.Y.; Chen, M.K. Using 2D CNN with Taguchi parametric optimization for lung cancer recognition from CT images. *Appl. Sci.* **2020**, *10*, 2591. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).