Journal of *Imaging*

MDPI

Article

A Novel Defect Inspection System Using Convolutional Neural Network for MEMS Pressure Sensors

Mingxing Deng [1,†], Quanyong Zhang [2,†], Kun Zhang [1], Hui Li [2,3,*] **, Yikai Zhang [2] and Wan Cao [4]**

[1] School of Automobile and Traffic Engineering, Wuhan University of Science and Technology, Wuhan 430065, China
[2] The Institute of Technological Sciences, Wuhan University, Wuhan 430072, China
[3] School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China
[4] Wuhan FineMEMS Inc., Wuhan 430075, China
* Correspondence: li_hui@whu.edu.cn
† These authors contributed equally to this work.

Abstract: Defect inspection using imaging-processing techniques, which detects and classifies manufacturing defects, plays a significant role in the quality control of microelectromechanical systems (MEMS) sensors in the semiconductor industry. However, high-precision classification and location are still challenging because the defect images that can be obtained are small and the scale of the different defects on the picture of the defect is different. Therefore, a simple, flexible, and efficient convolutional neural network (CNN) called accurate-detection CNN (ADCNN) to inspect MEMS pressure-sensor-chip packaging is proposed in this paper. The ADCNN is based on the faster region-based CNN, which improved the performance of the network by adding random-data augmentation and defect classifiers. Specifically, the ADCNN achieved a mean average precision of 92.39% and the defect classifier achieved a mean accuracy of 97.2%.

Keywords: MEMS pressure sensor; defect detection; convolutional neural network

Citation: Deng, M.; Zhang, Q.; Zhang, K.; Li, H.; Zhang, Y.; Cao, W. A Novel Defect Inspection System Using Convolutional Neural Network for MEMS Pressure Sensors. *J. Imaging* **2022**, *8*, 268. https://doi.org/10.3390/jimaging8100268

Academic Editors: KC Santosh, Ayush Goyal, Djamila Aouada, Aaisha Makkar, Yao-Yi Chiang, Satish Kumar Singh and Alejandro Rodríguez-González

Received: 25 August 2022
Accepted: 27 September 2022
Published: 30 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A microelectromechanical system (MEMS) is miniaturized, compact, inexpensive, easy to integrate, and compatible with the standard process [1]. It has been widely used in industrial productions [2]. The quick development of electronic portable, wearable, or implantable devices increases the need to develop very small-sized integrated MEMS sensors [3].

MEMS pressure sensors have prominent advantages in weight and power consumption due to their characteristics [4]. Basov [5] proposed a mathematical model of a high-sensitivity pressure sensor with a novel electrical circuit utilizing a piezosensitive transistor differential amplifier with a negative feedback loop. The circuit based on a vertical NPN and a horizontal differential transistor amplifier was used to analyze and optimize the pressure sensitivity and output stability of a PNP transistor. Yu et al. [6] presented a piezoresistive absolute micro-pressure sensor, which was of great benefits for altitude location. Based on the ANSYS simulation results, the surface stress and deflection were analyzed. The sensor equation was established. Xu et al. [7] systematically analyzed and discussed the influence of the bending-stiffness distribution of the diaphragm on the stress-concentration characteristics of the pressuresensor chip. In order to improve the application range of the pressure sensor, Basov [8] designed a pressure sensor suitable for an ultra-low pressure range and optimized its structure. Based on the theoretical model and experimental characteristics of an ultra-high-sensitivity MEMS pressuresensor chip, Basov [9] proposed a novel circuit. A pressure-sensitive differential amplifier was adopted in the circuit and had a negative feedback loop based on two bipolar junction transistors.

However, some nonelectrical faults may occur in MEMS. Defect detection plays an essential role in reliability evaluation [10]. The method of deep learning is superior to traditional machine learning and has been widely used in various fields [11]. Because of its wide application, various deep-learning neural networks have been promoted. Lecun et al. [12] proposed a suitable network architecture that used a gradient-learning algorithm to synthesize complex decision surfaces. The classification of high-dimensional patterns such as handwritten characters was realized, and various handwriting-recognition methods were compared. Convolutional neural networks show the ability to surpass all other technologies. Simonyan et al. [13] proposed a classic visual-geometry group network to promote the development of deep learning. The performance of neural networks is improved by adding neural-network layers, but the cost is a serious shortcoming. More specifically, a deeper network means more parameters, which can easily cause overfitting, high computational expenses, and long training cycles. Parameters are related to the fully connected layer. Therefore, replacing the fully connected layer with sparse connections is the key to solving these problems [14]. The occurrence of the residual structure increases the number of network layers and improves the accuracy without the phenomenon of gradient dispersion and explosion. Some classical neural networks have been proposed. Krizhevsky et al. [15] first proposed the AlexNet neural network using the ReLu function as the activation function of neurons. By visualizing the middle layer result of AlexNet, Zeiler et al. [16] improved the model by reducing the convolution kernel. Szegedy et al. [17] combined the initial structure with the residual connection to improve the overall effect of the network. Huang et al. [18] proposed a dense convolutional neural network (CNN) that connects each layer with all previous layers. Focusing on the channel dimension of convolutional neural networks, Hu et al. [19] proposed a novel architectural unit, termed the "Squeeze-and-Excitation" (SE) block, by explicitly modeling the interdependencies between channel relationships to adaptively recalibrate the channel-wise eigenresponse. Howard et al. [20] proposed the mobileNet network, which is an efficient, lightweight, and streamlined network. The separable convolution was used in the depth direction of the network, and two hyperparameters, width multiplier and resolution multiplier, were used to balance the model speed and precision. Zhang et al. [21] proposed the ShuffleNet network, which was a computationally efficient CNN architecture specially designed for mobile devices with very limited computing power. The new architecture utilized two new operations, pointwise group convolution and channel shuffling, which greatly reduced computational cost while maintaining accuracy. Xie et al. [22] proposed a simple, highly modular network structure for image classification. The network was constructed by repeating a building block that aggregates a set of transitions with the same topology. Sandler et al. [23] proposed MobileNetV2 by removing the nonlinear activation layer behind the small-dimensional output layer. The inverted residual structure was used to ensure the expressiveness of the model and alleviate the degradation of features. The accuracy rate of the MobileNetV2 was 78% but the reduction of the numbers of network parameters reduced the difficulty of the deployment. Based on the traditional convolution-module design, Mehta et al. [24] proposed an efficient spatial-pyramid convolution module (ESP Module) that could reduce model computation, memory, and power consumption. The applicability of terminal devices was improved and the accuracy rate of the ESP Module was 85%. On the premise of reducing the number of parameters, the detection effect of the network was maintained. Based on Shufflenet V1, Ma et al. [25] proposed Shufflenet V2 by reducing the degree of network fragmentation. The number of groupings and element-by-element operations improves the speed and accuracy of the network. Mehta [26] proposed an EspNetv2 that replaced the original EspNet pointwise convolutions with group pointwise convolutions. The accuracy rate of the network was 89%. This network had good robustness and few parameters, which had a significant impact on the application.

Currently, the mainstream target-detection networks include the Faster Region-based Convolutional Neural Network (Faster RCNN) and You-Only-Look-Once (YOLO). The main difference between the two networks was that YOLO had a faster detection speed,

whereas Faster RCNN had more accurate detection results [27,28]. He et al. [29] classified the defects of the steel surface with an accuracy of 99.67%; the MAP of defect inspection was 82.3%. However, the structure of the network was complex and unstable, and the MAP was still far from the industrial requirements. Some deep-learning algorithms were proposed for industrial inspections, but could not be fully adapted to the universal defects inspection. Therefore, in order to obtain better detection, a defect detection algorithm with low time cost needs to be developed to meet the defect characteristics of different products. Aiming at the defects in the manufacturing processes of the MEMS pressure sensor, accurate detection CNN (ADCNN) based on Faster RCNN is proposed in this paper.

First, the devices for packaging the MEMS pressure sensor and the packaging process of the pressure sensor are introduced. Chip damage, chip scratches, glue surface, gold-wire bonding, and aluminum-wire bonding are the defects that likely occur in the chip-packaging processes, which may cause failure of the performance of the MEMS pressure sensor. The types of defects for MEMS pressure sensors are provided by the MEMS pressure-sensor manufacturer. Second, an ADCNN network is proposed for MEMS pressure-sensor defect detection. The ADCNN is based on the Faster RCNN network, which improves the accuracy by using random-data augmentation and defect classifiers. Then, the training method and experimental results of the model are presented. Finally, the ADCNN is used to detect the defects of the MEMS pressure sensor.

2. System Overview

Figure 1a shows the MEMS pressure-sensor packaging devices. A high-temperature curing device is been widely used for curing the glue used in the production process. An automatic production line is used for chip packaging. Figure 1b shows the mechanical device of the automatic production line. The device mainly includes three workstations. The padding workstation is used to pad raw materials. The gluing workstation is used to apply the glue. The bonding workstation is used to bond gold and aluminum wires. The MEMS pressure-sensor-chip packaging achieves automation by streamlining these workstations. Figure 2 shows the MEMS pressure-sensor packaging process. It consists of three main steps. First, a machine applies glue to the Al_2O_3 substrate and pastes the MEMS pressure-sensor chip. The process requires that the chip not have any damage or scratches. Second, the bonding workstation bonds the gold wire to connect the chip to the circuit on the Al_2O_3 substrate. Gold-wire bonding breaks may occur in the process. The next step is to reapply the glue and paste the shied adhesive. Then, the gluing workstation pours potting adhesive into the shied ring and bonds the aluminum wire. Glue-surface wrinkles and aluminum-wire bonding breaks may occur, resulting in MEMS pressure-sensor failure. Finally, the finished products are encapsulated.

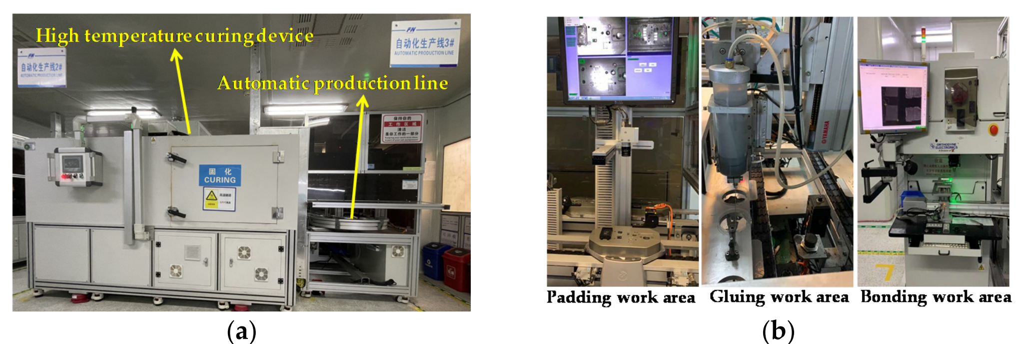

Figure 1. MEMS pressure-sensor packaging equipment. (**a**) High-temperature curing device and automatic production line. (**b**) Mechanical devices of automatic-production-line work area.

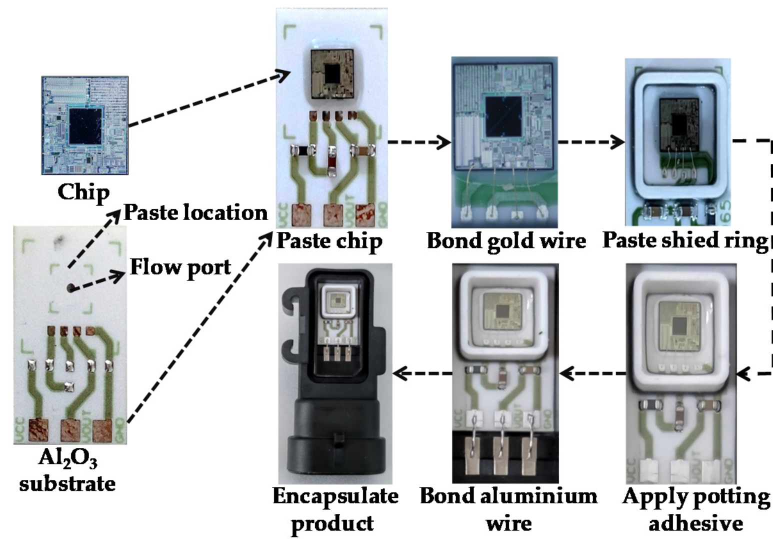

Figure 2. Packaging process of the MEMS pressure sensor.

Figure 3 shows the sample defects of the MEMS pressure sensor. The image datasets, each with a resolution of 1000 × 800, were provided by Wuhan FineMEMS Inc. It contains five kinds of defects: chip scratch, chip damage, glue-surface wrinkle, broken gold-wire bonding, and broken aluminum-wire bonding. A chip scratch or chip damage could damage the internal circuits of the MEMS pressure-sensor chip. A glue-surface wrinkle could affect the air tightness of the MEMS pressure sensor. Broken gold-wire bonding and broken aluminum-wire bonding could lead to an open circuit, which could cause the performance failure of the MEMS sensor.

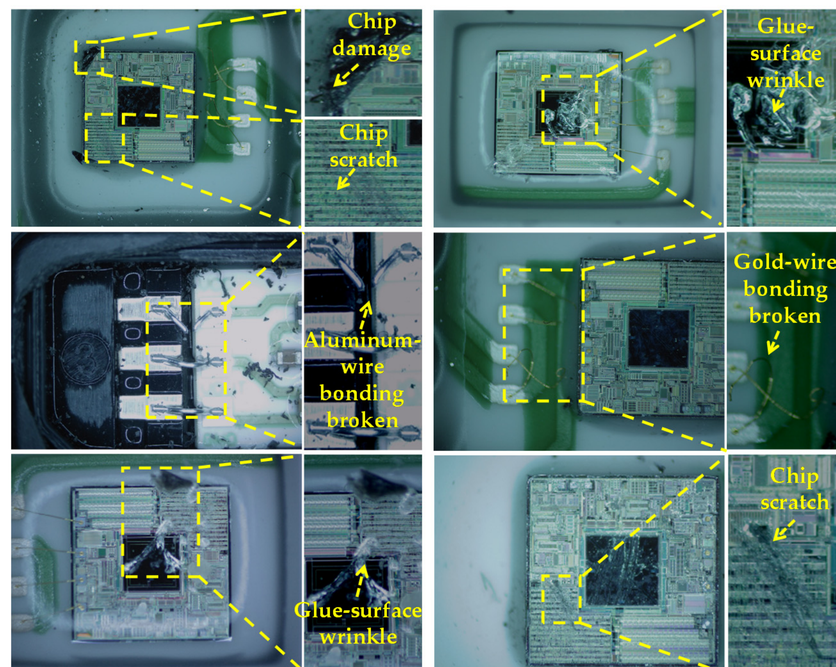

Figure 3. Five kinds of defect images of the MEMS pressure sensor from direct observation.

3. Proposed CNN

Defect detection is very important in industrial engineering. The purpose of this paper aims to determine all the defective objects in the image, and classify and locate them.

3.1. Improved Network Framework

The ADCNN is based on traditional Faster RCNN target detection. It improves the network to inspect the MEMS pressure-sensor chip-packaging process by changing the traditional convolution connection and adding defect classifiers and a random-data-augmentation module.

Traditional fast RCNN uses CNN to extract features and preset anchors points to obtain feature-layer and proposed anchor points. Then, the network obtains the location and classification. Traditional Faster RCNN extracts features and uses feature layers and suggested anchors to preset the anchor in the detection anchor to obtain the feature layer and suggested anchor of the target. However, the defects in MEMS pressure-sensor chips are different in size and scale. The number of defect images is limited. Traditional Faster RCNN cannot detect these different scale defects in a single image.

Figure 4 shows the detail of the ADCNN. The ADCNN comprises a detection location and classifier parts. The preliminary improvement changes the traditional convolution connection by two structures called skip and deep blocks in the detection part. The function of the skip block is to change the size of features extracted from defective images. It contains two repeated 1×1 convolutions, where strides are doubled. One of the convolutions is followed by a 3×3 convolution and 1×1 convolution, where strides are single. Each convolution is followed by batch-normalization layers and ReLu activation functions to prevent gradient explosion and vanishing. Features are put into two 1×1 convolutions, and then the convolved features are added. The deep block increases the number of network layers. It contains two repeated 1×1 convolutions and a 3×3 convolution. Strides of convolutions are single in the deep block. Extracted features and inputs are directly added. Two structures make the network better extract the details of defect images.

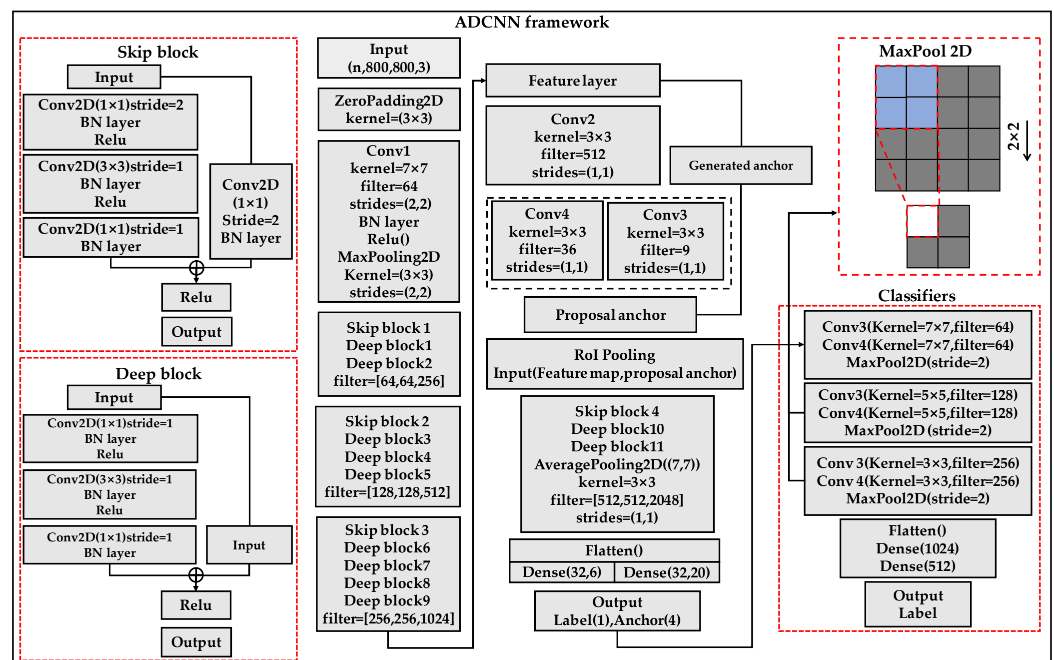

Figure 4. Framework of the ADCNN.

The second improvement is that the ADCNN adds defect classifiers to the traditional Faster RCNN. The improved traditional network is used to first detect the possible locations of smaller-scale defects. According to these locked positions, the defect classifier identifies further defects. The classifier part contains two repeated 7×7 convolutions, 5×5 convolutions, and 3×3 convolutions. Each repeated convolution is followed by max-pool layers. Classifiers can more effectively extract the features of detection targets by

using convolutions of different sizes. The defect classifiers can improve the accuracy of the whole network detection.

3.2. Workflow of the Network

Figure 5 shows the workflow of the ADCNN. The implementation of target detection depends on a special detection dataset containing expensive manual annotations. Defect images are obtained from the MEMS pressure-sensor manufacturer and labeled. A total of 6707 defect images are taken from the production line of Wuhan FineMEMS Inc. The labeled data are augmented in each training epoch. The defect images are input to the ADCNN, and these images are processed to adapt to the input form of the network. This process is called data encoding.

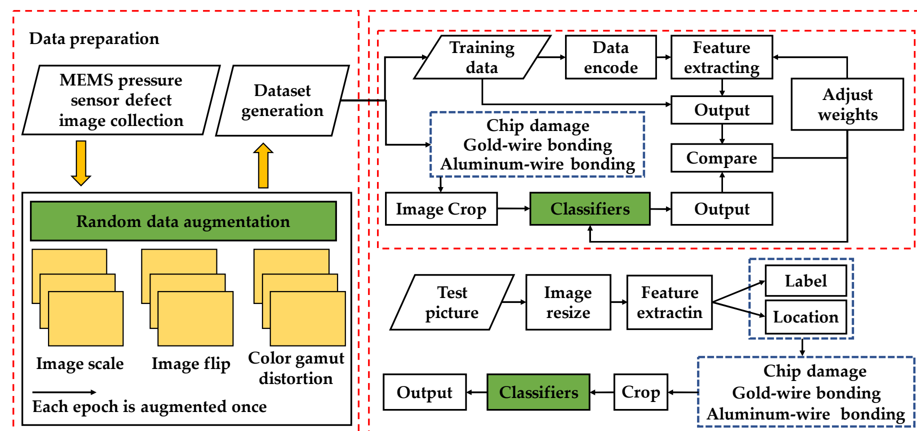

Figure 5. Workflow of the ADCNN for dealing with the defect images.

The detection-location part of the ADCNN obtains defect images and outputs location and classification. The classifier part of the ADCNN obtains the output in the detection-location part. The network deals with different kinds of defects in different ways. For large-scale defects such as glue-surface wrinkling and chip scratching, the detected locations and classifications are directly output by the network. For gold-wire-bonding and aluminum-wire-bonding defects, whether the bonding effect is qualified needs to be determined by classifiers. Small-scale defects for chip damage need to be further identified by the classifier. By detecting the position information of the positioning part as the special input of the classifier, the defect images of gold-wire bonding, aluminum-wire bonding, and chip damage are processed by the network to obtain the classification results. The results of the classification and location in the detection are output.

4. Training and Validation

4.1. Data Augmentation

Figure 6 shows the distribution of the defect datasets. The distribution is uneven and the number of defect images is limited. Random-data augmentation is applied to solve these problems. Random-data augmentation includes image scale, flip, and color-gamut distortion. The data-augmentation parameters are random in each training epoch. Thus, it significantly enhances data diversity and effectively avoids overfitting. Figure 7 shows the results from random-data augmentation. The size-scaling range of the defect image was between 0.9 and 0.5 of the original size. The flip angle was between 0 and 360 degrees. Gamut distortion was used to generate a new image by randomly adjusting the saturation, brightness, and contrast of the original images. The original image size, color gamut, and background were changed to obtain more defect images and to improve network robustness.

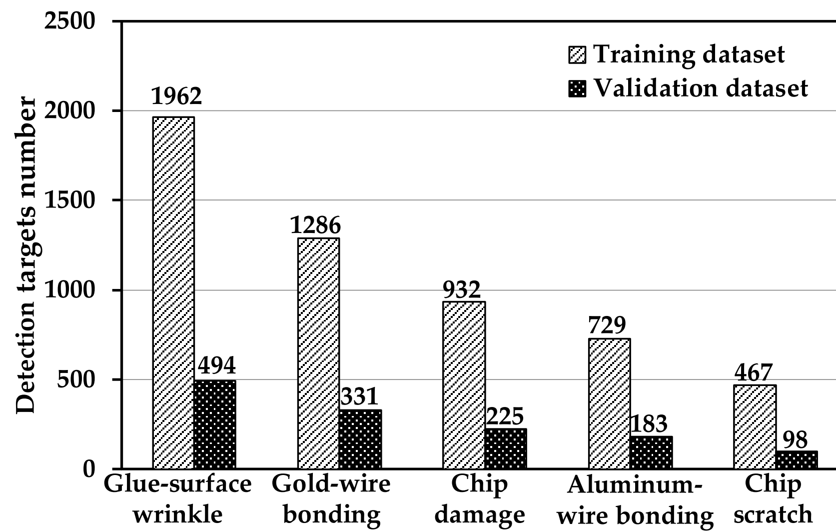

Figure 6. Numbers of five detection targets of MEMS pressure sensors.

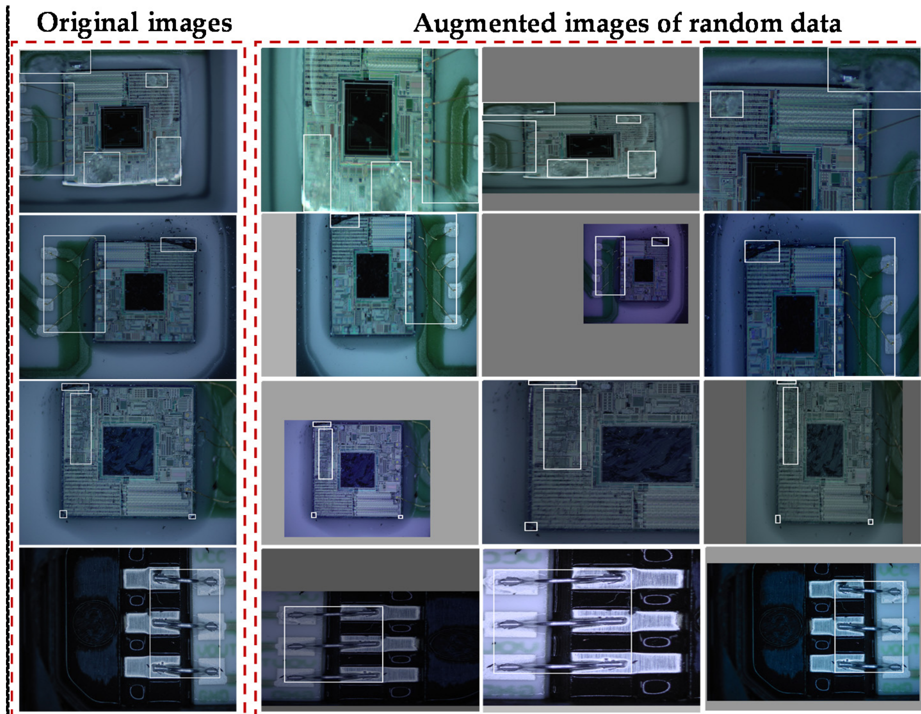

Figure 7. Data augmentation of five kinds of defect images of the MEMS pressure sensor in Figure 3.

4.2. Training Method

The training process of the ADCNN is divided into three parts. First, the network is trained to obtain the proposal anchors. Then, the network is trained to obtain the detection location and the first classification result using the proposal anchors. Finally, the classifiers are trained to obtain the detection results.

The network is optimized by the gradient descent with the loss function as the objective function. The loss function includes classification loss and regression loss. The L_{cls1} classification loss is used to train parameters, which determines whether the preset anchors contain targets on MEMS pressure sensors. Conversely, the L_{reg} regression loss is used to

train parameters, which adjusts the location of the preset anchor on the MEMS pressure sensor. It is described as follows.

$$L(p_i, t_i) = \frac{1}{N_{cls}} L_{cls1} + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

$$L_{cls1}(p_i, p_i^*) = \sum_i -log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (2)$$

$$L_{reg} = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & x \geq 1 \; or \; x \leq -1 \end{cases} \quad (3)$$

where i is the anchor index of the batch and p_i is the predicted probability of the target classification. t_i is the predicted coordinate-adjustment value and t_i^* is the true coordinate-adjustment value. λ is the weight, which balances classification loss and regression loss, and x is the value of the difference between t_i and t_i^*. p_i^* is a binary indicator, indicating whether the anchor contains the real detection target, and N_{cls} is the number of classification-training batches. N_{reg} is the number of regression-training batches. The adjustment parameters of the preset anchors are obtained by training networks, and the proposal anchors are obtained by adjusting parameters. Using the coincidence degree between the proposed and true anchors eliminates the proposal anchors with a low coincidence degree.

Then, the network is trained to obtain the coordinate-adjustment information of the proposal anchors and the classification results of detection targets in the detection-location part. The loss function is the same as in the previous loss function except for L_{cls2}. The L_{cls2} loss is used to train parameters, which determines the classification results of the target on MEMS pressure sensors. It is described as follows.

$$L(y_{ij}, t_i, p_{ij}, p_i) = L_{cls2} + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (4)$$

$$L_{cls2}(y_{ij}, p_{ij}) = -\frac{1}{N_{cls}} \sum_{i=1}^{N_{cls1}} \sum_{j=1}^{M} y_{ij} \log(p_{ij}) \quad (5)$$

where M is the number of possible classifications. y_{ij} is a binary indicator, indicating whether the classification j is the real classification of input instance x_i. p_{ij} is the probability that the network predicts that the input instance x_i belongs to classification j.

The final step is to train classifiers for gold-wire bonding, aluminum-wire bonding, and chip damage. Classifier training is the same as the above training, and the loss function is the L_{cls2} function.

4.3. Training Results

Figure 8a shows the loss before data augmentation. There was over-fitting in training. Because the data set was very small, the network could not achieving good training results. Figure 8b shows the loss after data augmentation. The problem of overfitting was solved.

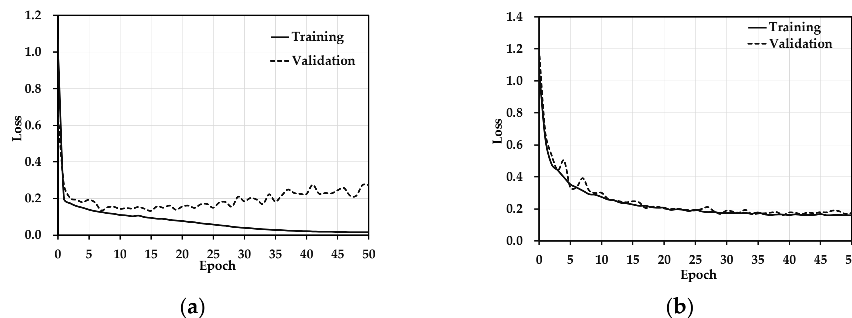

Figure 8. The MEMS pressure sensor detects the total loss of the positioning network. (**a**) Loss before data augmentation. (**b**) Loss after data augmentation.

5. MEMS Defect Detection

Figure 9 shows the defect-detection results of the MEMS pressure sensor. The numbers are the probability of the ADCNN detecting the different defects. The ADCNN could detect gold-wire bonding, glue-surface wrinkles, aluminum-wire bonding, chip damage, and chip scratches in the packaging process of the MEMS pressure-sensor chip.

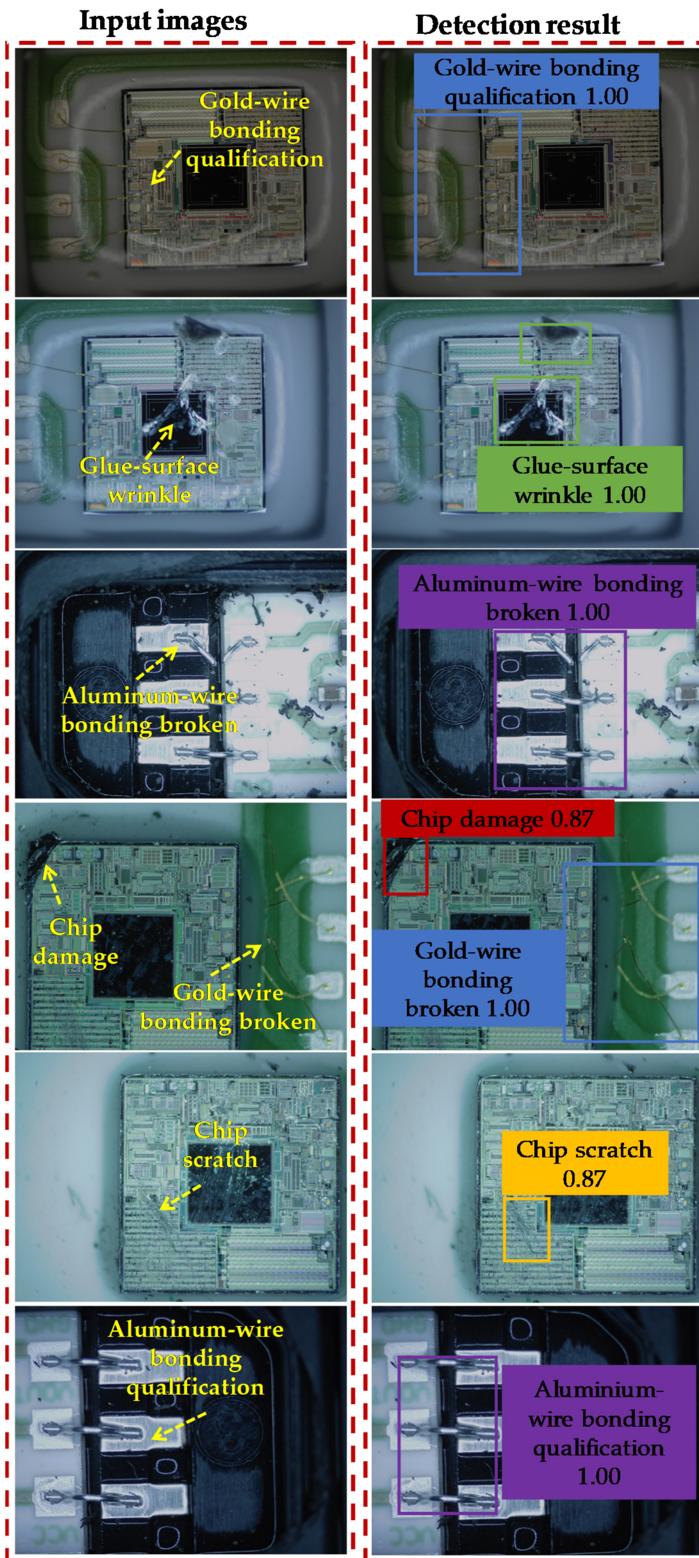

Figure 9. Defect-detection results of the MEMS pressure sensor based on the ADCNN.

Average accuracy (AP) is used to evaluate the results of testing experiments. AP is a good trade-off between accuracy and recall, which are two important testing indexes. These indexes are defined as follows.

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$AP = \frac{Precision + Recall}{2} \tag{8}$$

where TP, FP, and FN represent the number of true positives, false positives, and false negatives, respectively. The results of the defect detection were obtained for the MEMS pressure-sensor dataset. They include the mean average precision (MAP) and the accuracy in the detection location and classifier. Table 1 shows the target-detection statistics before network optimization. The MAP only achieved 89.6%, and the AP of chip damage only achieved 65.8%. The APs of chip scratches, gold-wire bonding, and aluminum-wire bonding were high.

Table 1. Statistics of target-detection results without network optimization.

Detection Classes	*Recall* (%)	*Precision* (%)	*AP* (%)	MAP (%)
Chip scratch	100.0	95.2	98.2	89.6
Chip damage	94.5	60.9	65.8	
Gold-wire bonding	99.7	98.2	99.7	
Glue-surface wrinkles	82.0	98.8	84.3	
Aluminum-wire bonding	100.0	100.0	100.0	

In order to obtain high performance of defect identification and classification, the Adam optimizer was used with a learning rate of 0.05 and a batch number of 2. These parameters were determined based on the characteristics of the dataset. Table 2 shows the target-detection statistics after network optimization. The MAP achieved 92.4%. The random-data augmentation greatly alleviated the problem of fewer defects in the MEMS pressure sensor and the MAP of network was improved by optimization of the network. Recall of chip damage was 80%, but the precision was only 75.8%. Other defects scattered the attention of the network, which may have caused low precision of chip-damage detection. The network trained classifiers to solve this problem. Figure 10a shows the loss of classifiers. As shown in Figure 10a, the performances were excellent because overfitting did not occur in training. Figure 10b shows the accuracy of classifiers, where the average accuracy of classifiers was 97.2%. Since there is currently no research that combines deep learning and MEMS pressure-sensor chip-defect detection, it was compared with the research on steel-surface detection using deep learning. He et al. [29] classified the defects of a steel surface with an accuracy of 99.67%, and the MAP of defect inspection was 82.3%. Therefore, the low accuracy of chip damage was effectively solved by the classifiers. In addition, the stability of the network was improved by locating the aluminum-wire-bonding area and the gold-wire-bonding area, and checking the defects by classifiers.

Table 2. Statistics of target-detection results with network optimization.

Detection Classes	*Recall* (%)	*Precision* (%)	*AP* (%)	MAP (%)
Chip scratch	96.1	86.8	92.3	
Chip damage	80.0	75.8	71.5	
Gold-wire bonding	98.3	82.7	98.8	92.4
Glue-surface wrinkles	88.1	93.5	95.4	
Aluminum-wire bonding	100.0	100.0	100.0	

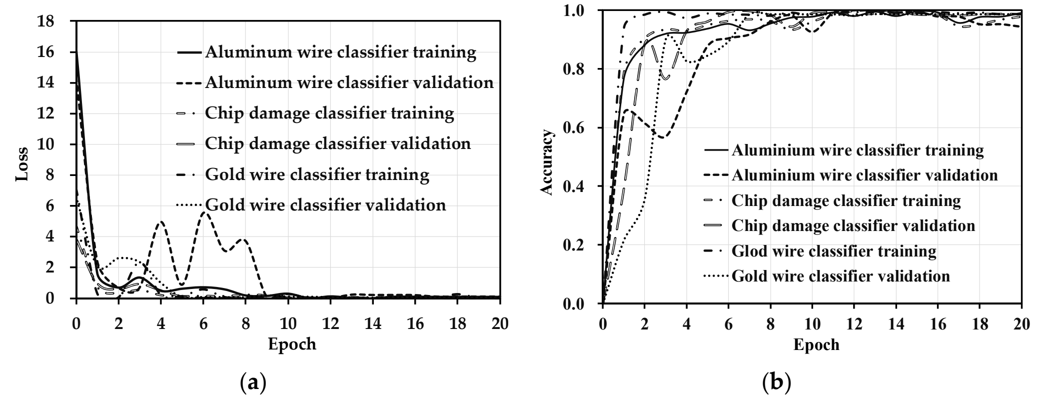

Figure 10. MEMS pressure-sensor detection-classification network. (**a**) Loss. (**b**) Accuracy.

It can be seen from Table 3 that the average detection accuracy of the network used in this work was 92.4%, and compared with the Faster RCNN, YOLOv3, YOLOv4 networks, the average detection-accuracy values were improved by 2.6%, 9.4%, and 1.1%, respectively. The detection time of a single defect picture was 68 ms, which meets the requirements of defect detection for a production line. Compared with other networks, the training results of the ADCNN were more accurate, and the detection time of a single image met the test requirements.

Table 3. Comparison of detection results of different networks.

Network	*AP*/%					MAP /%	Single-Picture Detection Time/ms
	Chip Scratch	Chip Damage	Gold-Wire Bonding	Glue-Surface Wrinkles	Aluminum-Wire Bonding		
Faster RCNN	90.5	89.6	93.5	85.7	89.9	89.8	51
YOLOv3	85.4	79.9	83.0	87.6	81.8	83.0	45
YOLOv4	92.8	82.7	95.5	93.2	92.4	91.3	42
ADCNN (This work)	92.3	91.2	98.8	95.4	98.4	92.4	68

6. Conclusions

In this paper, a defect-detection system for MEMS pressure-sensor chip packaging is proposed. The system can obtain the specific category and detailed location of the defect. The ADCNN improves on the Faster RCNN framework with a skip block and a deep block. Additionally, it adds a defect classifier and data augmentation to improve the accuracy of the network. The experiments of the ADCNN on the MEMS pressure-sensor defect dataset show that the network achieved 92.4% MAP for the defect detection. These results demonstrate a high level in the industry. Future studies will focus on the following two directions. Firstly, because of the expensive manuals and symbols in the detection data set, a generative countermeasure network should be developed. Secondly, CNN is used for feedback control of the MEMS pressure-sensor chip package so as to obtain a more intelligent system.

Author Contributions: Conceptualization, H.L. and K.Z.; methodology, K.Z. and Y.Z.; software, Q.Z., K.Z. and Y.Z.; validation, H.L. and M.D.; formal analysis, H.L., K.Z. and W.C.; investigation, Q.Z. and K.Z.; resources, W.C.; data curation, W.C., H.L. and K.Z.; writing—original draft preparation, M.D., Q.Z. and K.Z.; writing—review and editing, M.D., Q.Z. and H.L.; supervision, H.L. and M.D.; project administration, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key R&D Program of China under grant No. 2019YFB1704600 and the Hubei Provincial Natural Science Foundation of China under grant No. 2020CFA032.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ahmad, F.; Baig, A.; Dennis, J.; Hamid, N.; Khir, M. Characterization of MEMS comb capacitor. *Microsyst. Technol.* **2020**, *26*, 1387–1392. [CrossRef]
2. Zhang, Y.; Li, B.; Li, H.; Shen, S.; Li, F.; Ni, W.; Cao, W. Investigation of potting adhesive induced thermal stress in MEMS pressure sensor. *Sensors* **2021**, *21*, 2011. [CrossRef] [PubMed]
3. Hafez, N.; Haas, S.; Loebel, U.; Reuter, D.; Ramsbeck, M.; Schramm, M.; Horstmann, J.; Otto, T. Characterisation of MOS transistors as an electromechanical transducer for stress. *Phys. Status Solidi* **2019**, *216*, 1700680. [CrossRef]
4. Gao, T.; Sheng, W.; Zhou, M.; Fang, M.; Luo, F.; Li, J. Method for fault diagnosis of temperature-related MEMS inertial sensors by combining hilbert-huang transform and deep learning. *Sensors* **2020**, *20*, 5633. [CrossRef] [PubMed]
5. Basov, M. Development of high-sensitivity pressure sensor with on-chip differential transistor amplifier. *J. Micromechanics Microeng.* **2020**, *30*, 065001. [CrossRef]
6. Yu, Z.; Zhao, Y.; Sun, L.; Tian, B.; Jiang, Z. Incorporation of beams into bossed diaphragm for a high sensitivity and overload micro pressure sensor. *Rev. Sci. Instrum.* **2013**, *84*, 015004. [CrossRef]
7. Xu, T.; Lu, D.; Zhao, L.; Jiang, Z.; Wang, H.; Guo, X.; Zhao, Y. Application and optimization of stiffness abruption structures for pressure sensors with high sensitivity and anti-overload ability. *Sensors* **2017**, *17*, 1965. [CrossRef] [PubMed]
8. Basov, M.; Prigodskiy, D. Investigation of high-sensitivity piezoresistive pressure sensors at ultra-low differential pressures. *IEEE Sens. J.* **2020**, *20*, 7646–7652. [CrossRef]
9. Basov, M. Ultra-high sensitivity MEMS pressure sensor utilizing bipolar junction transistor for pressures ranging from −1 to 1 kPa. *IEEE Sens. J.* **2020**, *21*, 4357–4364. [CrossRef]
10. Asgary, R.; Mohammadi, K.; Zwolinski, M. Using neural networks as a fault detection mechanism in MEMS devices. *Microelectron. Reliab.* **2007**, *47*, 142–149. [CrossRef]
11. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
12. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
13. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
14. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 8–10 June 2015; pp. 1–9.
15. Krizhevsky, A.; Sutskever, I.; Hinton, G. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2017**, *60*, 84–90. [CrossRef]
16. Zeiler, M.; Fergus, R. *Visualizing and Understanding Convolutional Neural Networks*; Springer International Publishing: Cham, Switzerland, 2013; pp. 818–833.
17. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; pp. 4278–4284.
18. Huang, G.; Liu, Z.; Maaten, L.; Weinberger, K. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 22–25 July 2017; pp. 2261–2269.
19. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 2011–2023.
20. Howard, A.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.

21. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 6848–6856.
22. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 22–25 July 2017; pp. 5987–5995.
23. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L. MobileNetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 4510–4520.
24. Mehta, S.; Rastegari, M.; Caspi, A.; Shapiro, L.; Hajishirzi, H. ESPNet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 552–568.
25. Ma, N.; Zhang, X.; Zheng, H.; Sun, J. ShuffleNet v2: Practical guidelines for efficient CNN architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 116–131.
26. Mehta, S.; Rastegari, M.; Shapiro, L.; Hajishirzi, H. ESPNetv2: A light-weight, power efficient, and general purpose convolutional neural network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Los Angeles, CA, USA, 16–20 June 2019; pp. 9182–9192.
27. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *39*, 1137–1149. [CrossRef] [PubMed]
28. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 29–30 June 2016; pp. 779–788.
29. He, Y.; Song, K.; Meng, Q.; Yan, Y. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 1493–1504. [CrossRef]