다중 프레임 이미지 분할 기술을 사용한 알루미늄 자동차 도어 결함 감지

Ugur Ercelik, HaoYu Chen, Longfei Li, 김경백

전남대학교 인공지능융합학과

ugurercelik@hotmail.com, leochy554@gmail.com, lilongfei0712@gmail.com, kyungbaekkim@jnu.ac.kr

Aluminum Car Door Defect Detection by Using Multi-frame Image Segmentation Techniques

Ugur Ercelik, HaoYu Chen, Longfei Li, Kyungbaek Kim

Dept. of Artificial Intelligence Convergence, Chonnam National University

ugurercelik@hotmail.com, leochy554@gmail.com, lilongfei0712@gmail.com, kyungbaekkim@jnu.ac.kr

Abstract

AI-based image detection technology offers a promising solution for identifying defects in car doors, significantly improving efficiency compared to traditional human inspections. This paper introduces an advanced automatic defect detection system utilizing camera-recorded datasets and trained models to identify defects in aluminum car doors. Unlike previous models focused on aluminum castings, this is the first application specifically targeting car doors. Despite progress in defect detection research, challenges such as data imbalance, complex defect characteristics, and limited research on aluminum car doors persist. To address these issues, we propose the LKADenseNet201 model, enhancing the DenseNet201 architecture with a large kernel attention mechanism.While doing this, we focus on 3 important issues: image augmentation, channel attention and model evaluation.Our image processing process mainly include image augmentation. With image augmentation, we aimed to make data diversity suitable for the real world by obtaining data from different angles and to eliminate the imbalance between defect and normal images. This improvement boosts the model's ability to perceive contextual features and increases computational efficiency, essential for detailed spatial understanding and time-critical tasks. Our approach not only enhances operator efficiency but also moves towards automating the inspection process.

1. Introduction

In response to the needs of door surface quality inspection in the automotive manufacturing industry, traditional quality inspection methods mainly rely on manual visual inspection, which is not only inefficient, but also prone to human errors and difficult to ensure accuracy. However, in industrial production, single feature recognition is not enough to deal with multiple types of surface defects. Industrial surfaces often contain multiple defects, and using multiple machine learning models to detect the same product is neither efficient nor simple. Considering the complexity and uncertainty of industrial production processes, it is crucial to develop a model that can detect multiple types of surface defects. Firstly, during product manufacturing, defective samples are rare compared to normal samples, leading to data imbalance, which can cause overfitting and poor performance. Secondly, the defects produced during manufacturing are often complex, easily confused with normal regions in the images, and vary in shape, making it difficult for models to extract effective features. Furthermore, there is limited research on using deep learning algorithms for detecting defects in aluminum car doors, and there are no existing models for recognizing surface defects.To address these issues, this paper proposes the LKADenseNet201 model. This model enhances the DenseNet201 architecture by incorporating a large kernel attention mechanism, improving the model's ability to perceive contextual features and increasing computational efficiency, thereby improving defect detection. This approach is particularly beneficial for image-based tasks that require detailed spatial understanding and where computational resources or inference time are critical. Unlike other researches, we use the Kernel Attention in our model to enable the model to focus on the most informative features in different channels in an image. While doing this, we also create a test environment suitable for the real world by providing diversity in the images with the image augmentation methods we use in the image processing stage. Additionally, we compared our LKADenseNet201 model with pretrained models and the Channel Attention CNN model. Our aim here is to show that the Kernel Attention method we use in our model is successful compared to other models

2. Related Work

In the field of Anomaly detection, several research studies have been carried out. This section presents different research studies revolving around anomaly detection.Chalapathy et al.,[1] focused on two fold, a structured and comprehensive overview of research methods in deep learning-based anomaly detection.Furthermore ,they reviewed the adoption of these methods for anomaly across various application domains and asess their effectiveness. Xia et al.,[2] improved the Serre standard model, which can simulate the ventral visual pathway with object recognition ability, based on the latest research progress and results of simulating biological visual mechanism models in computer vision, to improve the recognition effect of surface defects on solar panels. At the same time, a pre-processing scheme combining Gaussian Laplace operator operator and adaptive Wiener filter to remove noise spots is studied, and the local Gabor Binary Pattern Histogram Sequence (LGBPHS) features are obtained through pre-processing.The experimental results show that the proposed method has an accuracy rate of %98.86 in training and %98.64 in testing, and both the false detection rate and the missed detection rate do not exceed %1.

3. Methodology

Here, while detecting defects in the car door images, we also classify whether the relevant part is defect or not on the entire image that is separated into parts. In the model we developed, we first apply a threshold value to detect defects on the aluminum car door images. Then, we divide the image into frames and perform classification defect or not a on the relevant image[Figure 2]. In the training phase, we use the LKADenseNet201 model we developed. To address the limitations of traditional automated techniques in detecting surface defects on industrial aluminum car doors, we enhanced the DenseNet201 model by incorporating a Large Kernel Attention (LKA) mechanism, resulting in the Large Kernel Attention DenseNet201 (LKADenseNet201) model. This novel approach aims to improve the accuracy and automation of surface defect detection within the automotive manufacturing industry. The Large Kernel Attention (LKA) module is integrated into the DenseNet201 model to enhance its feature extraction capabilities. The LKA mechanism comprises three key components designed to capture long-range dependencies and emphasize significant features while maintaining computational efficiency. Below, we detail the structure and mathematical formulation of the LKA module.

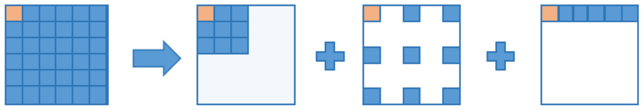

Figure 1 LKA

A. Initial Depthwise Convolution

Depthwise convolution is a type of convolution operation used in neural networks, particularly in the design of efficient models like MobileNets. It is a key component in depthwise separable convolutions, which are designed to reduce the computational cost and number of parameters in a convolutional neural network.The initial component of the LKA module is a depthwise convolution with a **5 x 5** kernel. This convolution independently processes each input channel while preserving the spatial dimensions, which enhances local feature extraction without significantly increasing computational complexity.

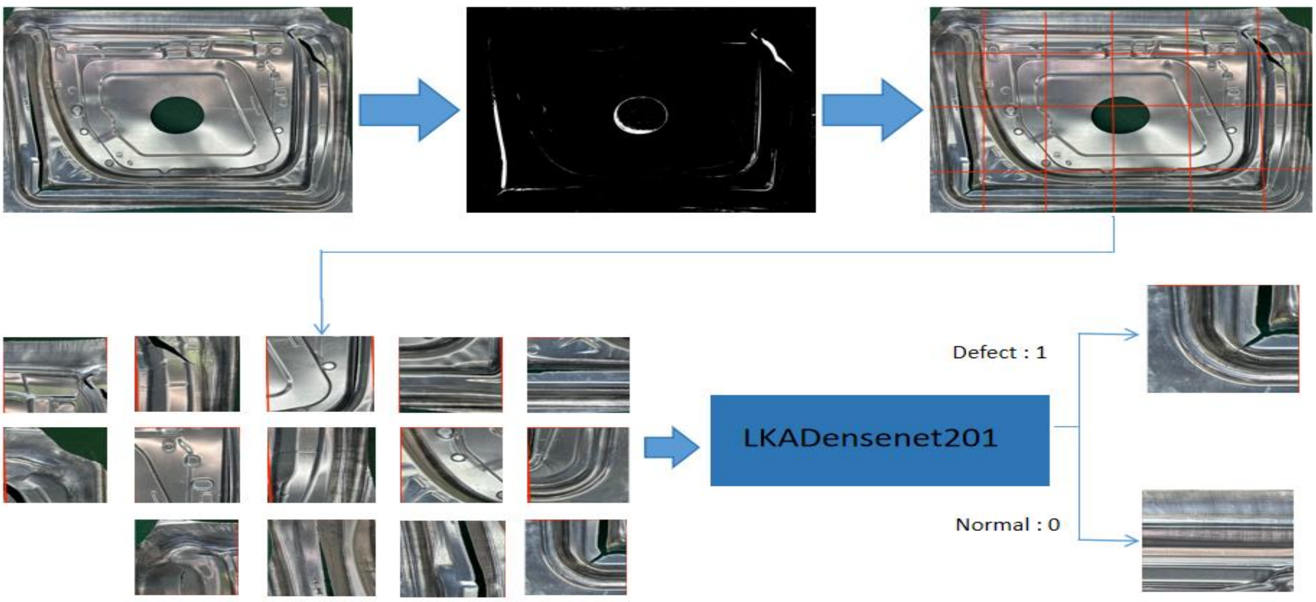

Figure 2 Architecture of Training Process

The operation can be mathematically expressed as:

$$\mathrm{Conv}_{\mathrm{depthwise}}(X) = W_{\mathrm{depthwise}} * X$$

where $X \in R^{C\times H\times W}$ represents the input feature map, $W_{\mathrm{depthwise}} \in R^{C\times 5\times 5}$ are the depthwise convolution weights, and $*$ denotes the convolution operation.

B. Dilated Spatial Convolution

Dilated Convolution is a variation of the standard convolution operation used in convolutional neural networks (CNNs). It is designed to expand the receptive field of the filter without increasing the number of parameters or the amount of computation. Dilated convolution explains gaps or "holes" between the elements of the filter, controlled by the **dilation rate**. The dilation rate *d* determines how much the filter is "spread out." Following the initial depthwise convolution, a dilated convolution with a 7 x 7 kernel and a dilation rate of 3 is applied. This operation expands the receptive field, allowing the network to capture contextual information from a larger spatial area without a substantial increase in the number of parameters. The dilated convolution operation is defined as:

$$\mathrm{Conv}_{\mathrm{dilated}}(X) = W_{\mathrm{dilated}} * X$$

where $W_{\mathrm{dilated}} \in R^{C\times 7\times 7}$ represents the dilated convolution weights, and the dilation rate d = 3 modifies the kernel to effectively cover a larger area

C. Pointwise Convolution

The final component of the LKA module is a pointwise convolution, which uses a 1 x 1 kernel to compress the channel information and create an attention map. This attention map is then element-wise multiplied with the module's original input, effectively recalibrating the feature responses. The pointwise convolution can be mathematically represented as:

$$\mathrm{Conv}_{\mathrm{pointwise}}(X) = W_{\mathrm{pointwise}} * X$$

where $W_{\mathrm{pointwise}} \in R^{C\times 1\times 1}$ are the pointwise convolution weights.
The recalibrated output Y of the LKA module is given by:

$$Y = X \odot \mathrm{Conv}_{\mathrm{pointwise}}(\mathrm{Conv}_{\mathrm{dilated}}(\mathrm{Conv}_{\mathrm{depthwise}}(X)))$$

where $\odot$ denotes the element-wise multiplication.

D. Complete LKA Operation

Combining all components, the overall attention mechanism of the LKA module can be summarized as:

$$\mathrm{Attention} = \mathrm{Conv}_{1\times 1}(\text{DW-D-Conv}(\text{DW-Conv}(F)))$$
$$\mathrm{Output} = \mathrm{Attention} \odot F$$

where $F \in RC\times H\times W$ is the input feature map, and Attention $\in RC\times H\times W$ is the attention map generated by the LKA module.

E. Advantages of LKA

Firstly, Kernel Attention emphasizes the spatial structure of the feature maps are . An extension of Kernel Attention, where larger kernels (or receptive fields) are used to capture more global context in an image. This helps in capturing more extensive spatial relationships, making it suitable for tasks where a broad spatial context is needed, like detecting anomalies or defects over large areas in images. The LKA module effectively combines the benefits of convolutional operations and self-attention mechanisms. It captures both local context information and long-range dependencies while maintaining a manageable computational cost and parameter count. This module achieves adaptivity in both spatial and channel dimensions, which is crucial for visual tasks where different channels often represent distinct object features.By integrating the LKA module with the DenseNet201 architecture, we achieve a model that not only leverages the robust feature propagation and reuse capabilities of DenseNet201 but also benefits from the enhanced feature emphasis provided by the LKA mechanism. This integrated approach significantly boosts the performance of surface defect detection in the automotive manufacturing industry. For our task, using **Large Kernel Attention (LKA)** is advantageous as it helps in capturing broader spatial patterns, which can be critical when detecting defects that might appear over larger areas of the image. **Large Kernel Attention (LKA)** are particularly effective when the spatial arrangement of features is important. It can capture complex spatial patterns across different regions of an image. Since we need to detect and classify defects in different regions of the image, we used the Large Kernel Attention method in this research.

F. Integration and Final Architecture

The integration of the Large Kernel Attention (LKA) mechanism with the DenseNet201 architecture forms the core of the LKADenseNet201 model. Initially, the DenseNet201 model, pre-trained on the ImageNet dataset, provides a robust foundation for feature extraction. By freezing the weights of the layers up to 'conv5 block32 2 conv' and fine-tuning the subsequent layers, the model retains generalized feature extraction capabilities while adapting to the specific task of defect detection. The output of DenseNet201 is then processed through the LKA module, which enhances the model's ability to emphasize significant features through a combination of depthwise convolution, dilated spatial convolution, and pointwise convolution. This results in improved discriminative power for detecting surface defects .Following the LKA module, a global average pooling layer reduces the feature maps to a manageable size, feeding into a dense layer with 1024 units activated by ReLU to capture high-level features. The architecture concludes with a sigmoid-activated dense layer for binary classification, effectively distinguishing between defective and non-defective car door surfaces. This integration strategy significantly boosts the model's accuracy and efficiency, making it a powerful tool for surface defect detection in the automotive manufacturing industry.

4. Results

The augmented dataset is then fed into the deep detection network in fixed batch sizes, significantly enhancing the dataset's richness without increasing the computational load during training. This augmentation strategy not only boosts the robustness of the model by providing more diverse training examples but also improves the overall detection performance. The experimental results demonstrate the superior performance of the proposed LKADenseNet201 model compared to several other state-of-the-art architectures in defect detection tasks. The test results of our LKADensenet201 model with pretrained models tested on 20 test images are shown in Table 1.

Table 1. Performance Comparison of LKADenseNet201 with Different Models

Model	Test Accuracy(%)	F1 Score	Confusion Matrix	
			True Positive	True Negative
VGG19	55.00	0.3903	11,0	9,0
DenseNet169	75.00	0.7251	11,0	5,4
DenseNet201	65.00	0.5809	11,0	7,2
ResNet50	90.00	0.8979	11,0	8,1
ResNet101	60.00	0.4933	11,0	8,1
LKADenseNet201	**75.00**	**0.7400**	**10,1**	**4,5**
VGG16	55.00	0.3903	11,0	9,0
DenseNet121	55.00	0.3903	11,0	9,0
Channel Attention CNN	70.00	0.6938	9,2	4,5

The augmentation methods we used on the train and validation sets and the image results we obtained as a result of augmentation are as shown in Table 2.

Table 2 . Data Augmentation

Dataset type	Methods	Result
2631 images	1. Crop 2. Rotate 3. flip	2150 samples train 461 samples validation 20 samples for test

5. Conclusion

Despite significant progress in product defect detection research, many challenges remain. Firstly, during product manufacturing, defective samples are rare compared to normal samples, leading to data imbalance, which can cause overfitting and poor performance. Secondly, the defects produced during manufacturing are often complex, easily confused with normal regions in the images, and vary in shape, making it difficult for models to extract effective features. Furthermore, there is limited research on using deep learning algorithms for detecting defects in aluminum car doors, and there are no existing models for recognizing surface defects.

To address these issues, this paper proposes the LKADenseNet201 model. This model enhances the DenseNet201 architecture by incorporating a large kernel attention mechanism, improving the model's ability to perceive contextual features and increasing computational efficiency, thereby improving defect detection. This approach is particularly beneficial for image-based tasks that require detailed spatial understanding and where computational resources or inference time are critical.

This study can be regarded as the first step in using AI to detect defects in cast aluminum car doors. Future work will involve using more comprehensive datasets to detect defects and adding tasks such as identifying defect types. The introduction of LKADenseNet201 marks a significant advancement in this field, offering a promising solution to the existing challenges and paving the way for more accurate and efficient defect detection systems.

Acknowledgements

This work was supported by the Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea government(MOTIE).(P0024554). This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00256629) grant funded by the Korea government(MSIT)

References

[1] Chalapathy, Raghavendra \& Chawla, Sanjay. (2019). Deep Learning for Anomaly Detection: A Survey.

[2] B. Xia, S. Li, N. Li, H. Li and X. Tian, "Surface Defect Recognition of Solar Panel Based on Percolation-Based Image Processing and Serre Standard Model," in IEEE Access, vol. 11, pp. 55126-55138, 2023, doi: 10.1109/ACCESS.2023.3281653.

[3] H. Chen and K. Kim, "Multi-Convolutional Channel Residual Spatial Attention U-Net for Industrial and Medical Image Segmentation," in *IEEE Access*, vol. 12, pp. 76089-76101,2024, doi: 10.1109/ACCESS.2024.3405957.

[4] HaoYu Chen, Longfei Li, Ugur Ercelik, Kyungbaek Kim. (2024-06-26). LKADenseNet201: A deep learning based defect detection model . Proceedings of the Korea Information Science Society, Jeju.

[5] LongFei Li, HaoYu Chen, Kyungbaek Kim, "SEPtDenseNet201: Car door defect detection model" in JCCI 2024