Research on Shellfish Recognition Based on Improved Faster RCNN

Yiran Feng[†], Sang-Yun Park[††], Eung-Joo Lee[†††]

ABSTRACT

The Faster RCNN-based shellfish recognition algorithm is introduced for shellfish recognition studies that currently do not have any deep learning-based algorithms in a practical setting. The original feature extraction module is replaced by DenseNet, which fuses multi-level feature data and optimises the NMS algorithm, network depth and merging method; overcoming the omission of shellfish overlap, multiple shellfish and insufficient light, effectively solving the problem of low shellfish classification accuracy. In the complexifier test environment, the test accuracy was improved by nearly 4%. Higher testing accuracy was achieved compared to the original testing algorithm. This provides favourable technical support for future applications of the improved Faster RCNN approach to seafood quality classification.

Key words: shellfish, deep learning, Faster RCNN, recognition

1. INTRODUCTION

Most seafood currently relies on mechanical and manual sorting. To reduce labour requirements and improve processing efficiency, computer vision can be applied to the non-contact calculation and measurement of seafood. This can improve processing efficiency and calculation accuracy without damaging the seafood being inspected. Related research has been conducted in the field of shellfish identification. Yang Jingyao et al.[1] used an extreme learning machine to achieve shellfish classification and identification, but the algorithm was difficult to achieve rapid high-volume inspection; Yang Mei et al.[2] studied scallop identification based on BP neural networks, but were unable to solve problems such as deformation and occlusion; Li Hongjie et al.[3] relied on computer vision tech-

nology to achieve seafood classification and quality assessment. However, the image processing time was relatively long in the recognition process; Xi, Rui et al.[4] proposed a computer vision deep learning method, which enabled the recognition rate of potato bud-eye to reach 96.32%, which greatly improved the detection accuracy although the recognition object was single; in 2013, Costa, C[5] applied computer vision to automatically classify size, sex and skeletal abnormalities of grouper, based on the least squares modelling multi-element technique combining image analysis and contour morphology was applied to the classification process of live fish.

This research innovatively introduced a Faster R-CNN network[6] for the recognition, localization and detection of four shellfish species based on previous research to address the lack of deep

※ Corresponding Author: Eung-Joo Lee, Address: 428, Sinseon-ro, Nam-gu, Busan, Korea, TEL: +82-53-629-1143, FAX: +82-, E-mail: ejlee@tu.ac.kr
Receipt date: May 20, 2021, Approval date: May 26, 2021
[†] Dept. of Information Communication Engineering, Tongmyong University, Dept of Mechanical Engineering and Automation, Dalian Polytechnic University, China
(E-mail: 345509139@qq.com)

[††] Dept. of Education Innovation Support Team, Silla University
(E-mail: sypark@silla.ac.kr)
[†††] Dept. of Information Communication Engineering, Tongmyong University

learning algorithms for shellfish recognition in real environments[7–9]. Based on the features of various shellfish, the Faster R–CNN framework was modified by replacing the feature extraction method and merging means, and collecting and constructing shellfish datasets in real environments, which finally solved the problem of shellfish recognition and localisation in occlusion, light and multi–target environments, and effectively improved the recognition rate of shellfish.

2. MATERIALS AND METHODS

2.1 Faster RCNN Architectures

Faster R–CNN has been widely used in recent years[10], by the combination of RPN and Fast R–CNN[11], and this study optimizes the architecture method. The three networks of Faster R–CNN are: basic network feature extraction, RPN (region candidate network) and test network. The colour, contour and texture of shellfish are all deep abstract features of shellfish. This network, after being trained, can extract the deep–level features of the input shellfish images, and after DenseNet extracts the shellfish feature maps at different scales and fuses them to achieve accurate shellfish recognition and classification. The process of shellfish recognition is shown in Fig. 1.

2.2 Improved Faster RCNN

While deeper networks allow deeper data information to be extracted, parameters inevitably increase as the network deepens[12]. This poses a number of problems for network optimisation and experimental hardware. The use of DenseNet as a feature–fetching network helps to address these issues, as the sample size of the dataset built for the shellfish classification and detection algorithm is small and network training tends to lead to overfitting[13].

As a novel network architecture, DenseNet borrows ideas from ResNet[14]. The most intuitive difference between the two architectures is the different transfer functions of the different network modules.

$$x_1 = H_l(x_{l-1}) + x_{l-1} \tag{1}$$

$$x_l = H_l([x_0, x_1, \cdots, x_{l-1}]) \tag{2}$$

The transfer function of ResNet described by (1) and (2) shows that the layer $l-1$ output of the network is equal to the non–linear change in the layer L output plus the layer $l-1$ output. In contrast, the layer 1 output of a DenseNet block is the set of all non–linear transformations of the previous layer's output.

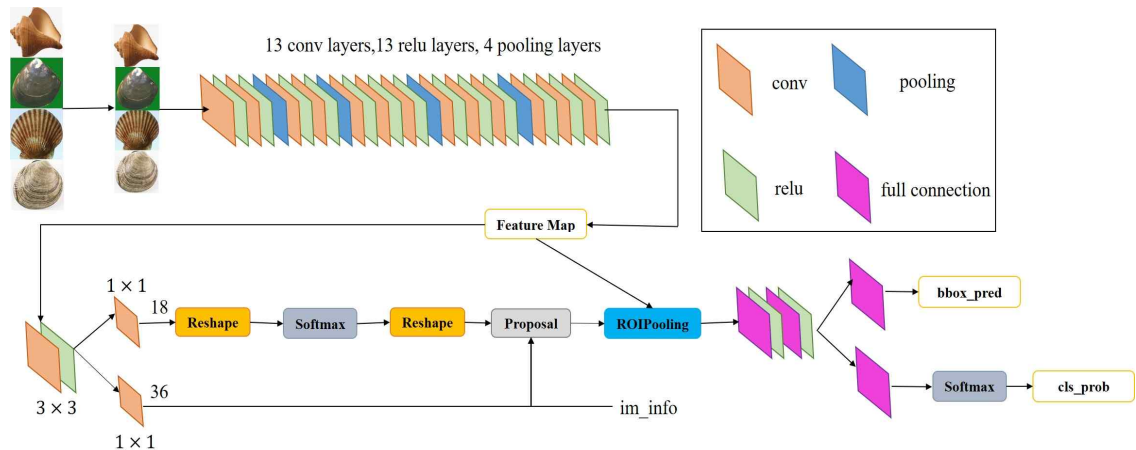The convolution in each network module (Dense Block) is interconnected[15]. h stands for each in–

Fig. 1. Faster RCNN algorithm steps.

Table 1. DenseNet structure parameters.

Structure	Parameter
Layers	7*7 conv, stride 2
Pooling	3*3 max pool, stride 2
Dense Block(1)	$\begin{bmatrix} 1*1\,conv \\ 3*3\,conv \end{bmatrix} *6$
Transition Layers(1)	1*1 conv
	2*2 average pool, stride 2
Dense Block(2)	$\begin{bmatrix} 1*1\,conv \\ 3*3\,conv \end{bmatrix} *12$
Transition Layers(2)	1*1 conv
	2*2 average pool, stride 2
Dense Block(3)	$\begin{bmatrix} 1*1\,conv \\ 3*3\,conv \end{bmatrix} *24$
Transition Layers(3)	1*1 conv
	2*2 average pool, stride 2
Dense Block(4)	$\begin{bmatrix} 1*1\,conv \\ 3*3\,conv \end{bmatrix} *16$
Transition Layers(4)	1*1 conv
	2*2 average pool, stride 2
Classification Layers	7*7 global average pool
	1000D fully-connected, softmax

put is operated with a k-dimensional 3×3 kernel using Batch Norm (batch normalisation) and ReLU (excitation function). k is the depth of the output features. the depth of DenseNet is 32[16]. For DenseNet, the aggregated connections in the mid-dle of its control module allow for a reasonable use of both shallow and deep layers which are charac-terised by ensuring their high efficiency and sig-nificantly reducing their diversity and computation.

Four 121-layer Dense Blocks are used to form the feature extraction network, and after removing the fully connected and classification layers, the RPN and RoI (Region of Interest) pooling layers are connected to complete target identification and localisation. The main parameters of the Dense Block architecture are listed in Table 1. lists the parameters of the four-DenseBlock architecture.

The front-end feature extractor and the end re-pressor of the Faster R-CNN detection algorithm were modified in order to achieve shellfish classi-fication detection in a realistic environment[17]. The algorithm steps are shown in Fig. 2.

2.3 Algorithm Example

The process of extracting features from shellfish is visualised and manipulated, with the original scallop image going through a filter, also known as convolution, to filter the image, then activation, and finally pooling to achieve a reduction in param-eters, using scallops as an arithmetic example, as shown in Fig. 3, where the granularity of the fea-ture map is enhanced as the depth of the network
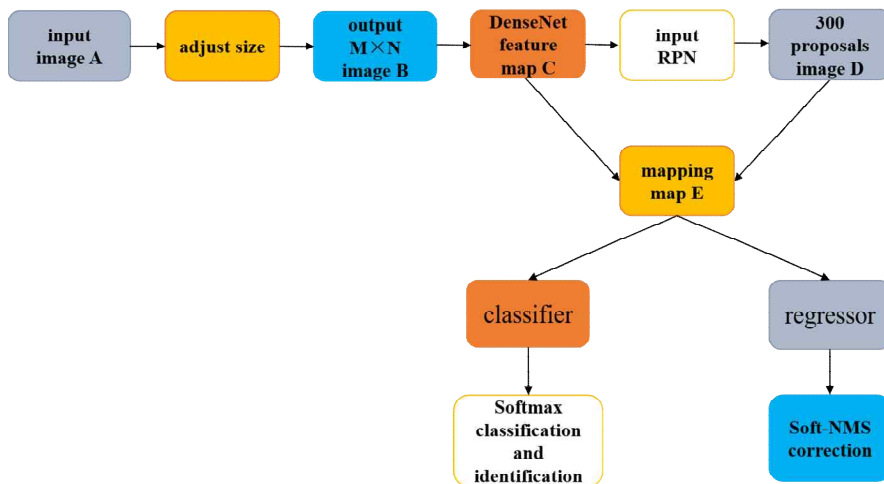

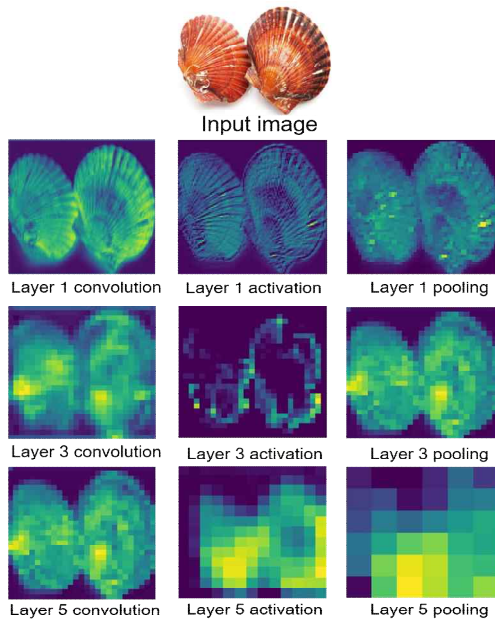Fig. 2. The process of proposed algorithm.

Fig. 3. Feature Maps of the convolution layer.

deepens.

3. EXPERIMENTS AND DATA PROCESSING

3.1 Experimental Conditions and Data Processing

Data features included varying light intensities, occlusions, complex backgrounds, and multiple targets to ensure that the detection model covered common real-life shellfish. In addition, 50% of the dataset was mirrored and the other 50% was panning extended to label the detected shellfish using LabelImg software. After extension, this data includes 8463 images, of which 90% is the training set and 10% are the detection set.

3.2 Comparative Analysis of Experimental Results

The algorithms were measured using the original algorithm and the optimised algorithm for convolution, relying on the shellfish dataset. The results of the various algorithms for shellfish detection are presented in Table 2, and the results of the shellfish diversity tests are shown in Table 3.

Based on the detection results, the Faster R-CNN using ResNet had a mAP value of over 77% in various shellfish detection. Fig. 4 shows some of the detection results. In Fig. 4(a), we note the obvious shellfish features and sufficient illumination, so the model can achieve significant detection performance. The scallop in Fig. 4(b) is partially obscured and contains a variety of shellfish with better detection results. The rainbow in Fig. 4(c) is missed. It is clear that ResNet is less adaptive in complex scenarios.

4. CONCLUSION

A deep learning algorithm for shellfish recognition is proposed in order to solve the problems of traditional shellfish recognition algorithms under different lighting, different backgrounds and different overlapping conditions. The algorithm is based on an improved Faster R-CNN, by optimising the algorithm and network architecture in order

Table 2. Test results of different network models.

Network	Conch	Scallop	Clam	Mussel
ResNet	71.8	76.2	81.6	79.8
DenseNet	75.1	80.3	86.2	82.9
DenseNet+SoftNMS	75.1	80.3	88.3	84.6

Table 3. Test results of difficult sample.

Network	Multi-object	Complex Backgrounds	Illumination Effect	Occlusion Effect
Original Algorithm	76.5	74.9	65.8	64.6
Improved Algorithm	78.9	76.8	72.2	73.4

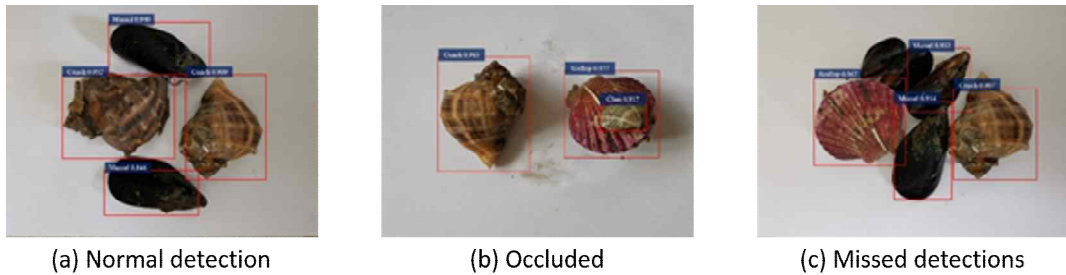(a) Normal detection   (b) Occluded   (c) Missed detections

Fig. 4. ResNet-Faster R-CNN detection result.

to improve the accuracy of shellfish detection. At the same time, the merging strategy is optimised with Soft-NMS instead of the original algorithm, thus improving the accuracy of detection. In addition, a dataset of four common shellfish species was created in conjunction with production realities. The detection algorithm studied is able to fulfil the needs of seafood processing enterprises for shellfish classification and identification, and has a better detection performance by nearly 4% improvement in detection accuracy than the traditional model under complex conditions such as insufficient light and overlap. In the next research, we will continue to optimise the algorithm and apply deep learning to the classification and quality detection of other seafood products.

REFERENCE

[1] J.Y. Yang, H.J. Li, and X.H. Tao, "Shellfish Recognition Based on Gabor Transformation and Extreme Learning Machine," *Journal of Dalian Polytechnic University*, Vol. 32, No. 4, pp. 310-312, 2013.

[2] M. Yang, H.L. Wei, and S.G. Hua, "A Scallop Image Recognition Method Based on a Neural Network," *Journal of Dalian Ocean University*, Vol. 29, No. 1, pp. 70-74, 2014.

[3] H.J. Li, X.H. Tao, and X.Q. Yu, "Application of Computer Vision Technology on Quality Evaluation of Seafood," *Journal of Food and Machinery*, Vol. 28, No. 4, pp. 154-156, 2012.

[4] R. Xi, K. Jiang, W.Z. Zhang, Z.Q. Lv, and J.L. Hou, "Recognition Method for Potato Buds Based on Improved Faster R-CNN," *Journal of Agricultural Machinery*, Vol. 51, No. 4, pp. 216-223, 2020.

[5] C. Costa, F. Antonucci, and C. Boglione, "Automated Sorting for Size, Sex and Skeletal Anomalies of Cultured Seabass Using External Shape Analysis," *Aquacultural Engineering,* Vol. 52, No. 7, pp. 58-64, 2013.

[6] R. Shaoqing, H. Kaiming and R. Girshick, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada,* pp. 91-99, 2015.

[7] A. Siddiqui, S. Musharraf and M. Choudhary, "Application of Analytical Methods in Authentication and Adulteration of Honey," *Food Chemistry,* Vol. 21, No. 7, pp. 687-698, 2017.

[8] X. Zou, L. Zhi and J. Shi, "Detection of Freshness Attributes of Yao Meat Based on Hyperspectral Imaging Technique," *Food Science,* Vol. 3, No. 6, pp. 65-77, 2014.

[9] M. Kamruzzaman, Y. Makino and S. Oshita, "Rapid and Non-Destructive Detection of Chicken Adulteration in Minced Beef Using Visible Near-Infrared Hyperspectral Imaging and Machine Learning," *Journal of Food Engineering,* Vol. 10, No. 7, pp. 8-15, 2016.

[10] Z. Liu, L. Yang and L. Wang, "Detection Approach Based on an Improved Faster RCNN for Brace Sleeve Screws in High-Speed

Railways," *IEEE Transactions on Instrumentation and Measurement,* Vol. 69, No. 7, pp. 4395-4403, 2019.

[11] X.R. Wu and X.Y. Ling, "Facial Expression Recognition Based on Improved Faster RCNN," *Journal of Intelligent Systems,* Vol. 4, No. 9, pp. 1-8, 2020.

[12] B. Chen, W. Chen, and X. Wei, "Characterization of Elastic Parameters for Functionally Graded Material by a Meshfree Method Combined with the NMS Approach," *Inverse Problems in Science and Engineering,* Vol. 26, No. 4, pp. 601-617, 2018.

[13] Z. Ning, F. Yiran, and E.-J. Lee, "Activity Object Detection Based on Improved Faster R-CNN," *Journal of Korea Multimedia Society,* Vol. 24, No. 3, pp. 416-422, 2021.

[14] Z. Li, Y. Lin, and A. Elofsson, "Protein Contact Map Prediction Based on ResNet and Dense Net," *BioMed Research International,* pp. 1-12, 2020.

[15] E. Zhang, B. Xue, and F. Cao, "Fusion of 2D CNN and 3D DenseNet for dynamic gesture recognition," *Electronics,* pp. 11-15, 2019.

[16] M. Peris and L. Escuder-gilabert, "Electronic Noses and Tongues to Assess Food Authenticity and Adulteration," *Trends in Food Science & Technology,* Vol. 58, pp. 40-54, 2016.

[17] M. Nan and Y. Li, "Improved Faster RCNN Based on Feature Amplification and Oversampling Data Augmentation for Oriented Vehicle Detection in Aerial Images," *Remote Sensing,* Vol. 12, pp. 25-58, 2020.

Feng Yiran

He is a doctoral candidate majoring in electronic information and Communication at Tongmyong University, South Korea. He received the B.E. degree and M.E. degree from Dalian Polytechnic University, China in 2013 and 2015 respectively. He works as a teacher at Dalian Polytechnic University and member of Korea Multimedia Association. He is mainly engaged in image recognition and robot technology research.

Sang-Yun Park

He received his B.S(1992-1995), in Physics, M,S(1996-1998), in multimedia Application from Kyungsung University, Korea, in 1996, Aug. 1998, respectively, and Ph. D.(2006-2011) in Information & Communication Engineering from Tongmyong University, Korea. in Aug. 2011. His main research interests include computer vision, neural network, pattern recognition. image processing, HCI(human computer interaction) and biometrics.

Lee Eung-Joo

Chairman of the Korean Multimedia Society. He received his B. S., M. S. and Ph. D. in Electronic Engineering from Kyungpook National University, Korea, in 1990, 1992, and Aug. 1996, respectively. Since 1997 he has worked with the Department of Information & Communications Engineering, Tongmyong University, Korea, where he is currently a professor. From 2005 to July 2006, he was a visiting professor in the Department of Computer and Information Engineering, Dalian Polytechnic University, and from Dec 2018 he was appointed honorary professor of Dalian Polytechnic University, China. His main research interests include biometrics, image processing, and computer vision.