Research Article

Machine Vision-Based Object Detection Strategy for Weld Area

Chenhua Liu, Shen Chen, and Jiqiang Huang

School of Mechanical Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China

Correspondence should be addressed to Chenhua Liu; liuchenhua023@163.com

Received 27 February 2022; Revised 18 March 2022; Accepted 21 March 2022; Published 11 April 2022

Copyright © 2022 Chenhua Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For the noisy industrial environment, the welded parts will have different types of defects in the weld area during the welding process, which need to be polished, and there are disadvantages such as low efficiency and high labor intensity when polishing manually; machine vision is used to automate the polishing and achieve continuous and efficient work. In this study, the Faster R-CNN object detection algorithm of two-stage is used to investigate the relationship between flops and the number of network parameters on the model by using a V-shaped welded thick plate as the research object and establishing the workpiece dataset with different lighting and angles, using six regional candidate networks for migration learning, comparing the convergence degree of different Batch and Mini-Batch on the model, and exploring the relationship between flops and the number of network parameters on the model. The optimal learning rate is selected for training to form a weld area object detection network based on the weld plate workpiece under few samples. The study shows that the VGG16 model is the best in weld seam area recognition with 91.68% average accuracy and 25.02 ms average detection time in the validation set, which can effectively identify weld seam areas in various industrial environments and provide location information for subsequent automatic grinding of robotic arms.

1. Introduction

The rapid development of welding has promoted the progress of related industries. The weld seam quality directly affects the structural use performance and the product life of the product [1]. Still, the weld seam, after welding, inevitably produces defects such as spatter, weld tumor, leakage, and porosity. Manual grinding of the weld area is required to eliminate welding defects, but there are disadvantages such as subjectivity and low efficiency. And the current welding is a large V-shaped welded thick plate; an automated weld area detection is needed to be introduced to find the location of the weld area [2]. At present, the detection of the weld area at home and abroad is mainly based on traditional and deep learning methods. Traditional image weld region extraction algorithm identifies to determine the weld seam location by constructing the critical points of the image and descriptor to the image to give feature information, such as ORB (Oriented FAST and Rotated BRIEF) [3], Image Hu moment invariant features [4], and AdaBoost weak learning [5]. Laser weld area detection and 3D laser reconstruction of the weld area are widely used in current research [6]. Such methods are extracted manually, and their poor generalization and low detection accuracy are unavoidable drawbacks.

Given weld seam forming results from the nonlinear formation of multiple welding parameters, convolutional neural networks (CNN) with nonlinear properties are also gradually applied to the welding industry [7], including object detection of weld seam areas. The mainstream object detection methods are mainly divided into one-stage detection algorithms represented by SSD and YOLO series based on regression analysis and two-stage detection algorithms defined by Faster R-CNN series based on candidate regions according to the detection process [8]. The former detection algorithm pursues more speed, while the latter also pursues more accuracy.

Although various object recognition methods have achieved significant results in workpiece recognition, the V-shaped weld plate workpiece samples belong to a small number of instances, which cannot meet the deep network training requirements, and there is a lack of weld seam region extraction algorithms under the few-sample weld plate workpiece. For V-shaped weld plate workpiece, such few-sample data contains very little labeled data, mainly
based on the traditional classical mature object detection method, using migration learning few-sample learning strategy [9]; based on fine-tuned migration learning of the implementation of the whole system, large-scale dataset for learning the source domain model used the model parameters to initialize the object domain model, and later on, the small-scale workpiece dataset is used for fine-tuned recognition. For the accuracy of weld seam detection, the Faster R-CNN network in the two-stage model is applied as the object recognition framework. The fine-tuned migratory learning of the learning source domain is performed using the VGG16 network. It is also tested in the test set with an accuracy rate that meets the needs of industrial use.

2. Object Detection of Weld Seam with Few Samples

2.1. Creation of Weld Area Dataset. Migration learning of the network model is performed on the dataset ImageNet, and the Intel RealSense D435i RGB-D camera is used to capture the V-shaped weld plate workpiece with weld seam, and each view of the weld plate workpiece is shown in Figure 1. Using the eye-in-hand calibration strategy [10] to find the coordinate conversion relationship, as shown in Figure 2, the coordinates of the weld area are converted to the robot coordinates under the robot coordinates to realize the vision-based automated welding and grinding work [11].

Collect images of weld plate workpieces under backlight, normal light, and multiple angles, 1000 images in total. In order to improve the network feature learning and training speed, the image size is adjusted to 400 × 300 pixels. In order to learn the workpiece features at each angle and improve the network overfitting phenomenon [12], for high-frequency images under unbalanced lighting, different variance values and Gaussian filter noise with Gaussian kernel are added to them, and the data enhancement scheme of a random level, pretzel noise attack, arbitrary angle rotation, and random clipping is used to expand the dataset to 5000 samples [13] because the workpiece samples are few samples’ data. Following the image data format of the neu-dataset (this is a dataset of steel plate surface curves produced by Northeastern University in China), the LabelMe annotation tool is used to export to XML, ensuring that each annotated border has only one weld feature. As shown in Figure 3, 80% of the images from the dataset were randomly selected as the training set, and 10% of the validation and test sets were divided, respectively.

2.2. Migration Learning Model Building Based on Welded Plate Workpiece Dataset. The Faster R-CNN [14] in the two-stage model is used to build a workpiece-oriented object recognition network with fewer samples, using the pre-trained weights and bias information on the ImageNet dataset, freezing all parameters except the fully connected layer, and modifying the Softmax classifier to train the weld seam features of the weld plate workpiece. A simplified schematic of the network structure is shown in Figure 4.

The network model uses shared convolutional layers to extract weld region image conv features’ map [15]. The network model uses a shared convolutional layer to extract the weld region image feature map, which is fed to the RPN (Region Proposal Networks) and the ROI (Region of Interest), respectively. The feature extraction layer consists of 13 convolutional layers, 13 ReLu activation layers, and four pooling layers; the RPN network determines whether the feature map belongs to the foreground or background by the Softmax classifier, where the RPN works as shown in Figure 5.

A 3 × 3 mask is used to slide the window motion on the feature map, and the position of the mask center corresponding to the original map is used as the center point (Figure 5(a)). Nine anchors with different scale aspect ratios are generated in the feature map (Figure 5(b)), and each anchor is assigned to the corresponding class label (positive label: foreground weld area and negative label: background area). The border regression algorithm obtains the weld area bounding box values and output to the ROI pooling layer for dimensionality reduction. The input of the ROI pooling layer is the feature map generated after the last convolution layer and the candidate region box are generated by the RPN layer,
and the final output is the ROI feature map. Finally, the fully connected layer and the Softmax classifier determine whether the candidate region is the weld region and output the exact location of the bounding box [16].

Six pretrained convolutional neural networks VGG16 [17], VGG19 [18], Googlenet [19], Resnet50 [20], Alexnet [21], and Lenet [22] on the neu-dataset are used as RPNs for migration learning so that the Faster R-CNN model first obtains the underlying feature weights of the images and then migrates the learning of these feature information to the task of weld region recognition, achieving the goal that the model can recognize accurate weld regions with a small number of artifacts.

Initial training was conducted using SGD (stochastic gradient descent) with learning rate set to 0.1, momentum factor set to 0.9, maximum Epochs set to 30, and dropout set to 0.1; the six feature region candidate networks were trained in turn, respectively. During the training process, the loss function $L$ is as in equations (1), (2), and (3):

$$L([p_i], [u_i]) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i,$$  

(1)

$$L_{cls}(p_i, p_i^*) = -\log[p_i p_i^*] + (1 - p_i)(1 - p_i^*),$$  

(2)

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*),$$  

(3)

Figure 3: Image data enhancement: (a) original image, (b) Gaussian filtered image, (c) image with 'salt and pepper,' (d) image with random rotation, (e) image with random flap, and (f) image with random cut.

Figure 4: Schematic diagram of Faster R-CNN network structure.
where $t_i = (t_x, t_y, t_w, t_h)$ is the coordinate value and bias of the bounding box, $t_i^*$ denotes the predicted actual coordinate value and bias, $N_{cls}$ is the number of classification samples, $\{P_i\}$ and $\{u_i\}$ are the output values, $\{P_i^\ast\}$ denotes the predicted probability of the weld region, and $P_i^\ast$ denotes the anchor point discriminant value, and when its value is 1, the anchor point is a positive label indicating identification to the weld region [23]; when its value is 0, it is a negative label indicating identification to the weld plate region. $L_{cls}$ denotes the classification layer loss function, and $N_{reg}$ denotes the number of regression samples. Since $L_1$ loss has high noise immunity and insensitivity to numerical fluctuations compared with $L_2$ loss, it can achieve better results in finding the border value of the target classification, so $L_1$ loss is used, as shown in (1) [24]:

$$R = \text{smooth}_{L_1}(x) = \begin{cases} 
0.5x^2 & |x| < 1 \\
|x| - 0.5 & \text{others}
\end{cases}$$  \hfill (4)

Iterative training and testing were performed on professional graphics workstations with the training and testing environments, as shown in Table 1.

2.3. Evaluation Method for the Performance of Weld Area Identification Models. Precision $P$ (Precision), Recall $R$ (ReCall), and Average Precision $AP$ (Average Precision) [25] are used as evaluation metrics for object detection in the weld area to assess the degree of merit of the model. The calculation is shown in equations (5), (6), and (7):

$$P = \frac{TP}{TP + FP},$$ \hfill (5)

$$R = \frac{TP}{TP + FN},$$ \hfill (6)

$$AP = \frac{1}{0} \int P.R.dR.$$ \hfill (7)

In (2), the larger the AP is, the better its recognition of weld regions is. TP is the number of correctly classified positive samples, FP is the number of incorrectly classified positive samples, and FN is the number of incorrectly classified negative samples. For each identified weld region, the identification score of the weld region is expressed as a confidence level. The precision-recall curves are plotted [26]. The precision responds to how accurately it predicts the positive samples, the recall responds to how well the classifier covers the positive samples, and the average precision is the integral of the precision-recall curve.

This study ensures that different candidate zone networks are tested and analyzed in a unified hardware-software environment to ensure a consistent FLOPS (Floating-Point Operations Per Second) [27]; the metric is used to measure to estimate the execution performance of the computing platform. The number of parameters of the RPN is only related to the network structure, and the memory occupation of the model is approximately four times the memory occupation of the number of parameters.
number of parameters [28] for the convolutional and fully connected layers is calculated as shown in equation (8), (9), and (10):

$$\text{ParasConv} = n \times (h \times w \times c + 1),$$  
$$\text{ParasFull} = \text{Weight}_{in} \times \text{Weight}_{out},$$  
$$\text{ParasNum} = \sum_{i=1}^{j} (\text{ParasConv}_{i} + \text{ParasFull}_{i}),$$  

where c is the number of input channels, n is the number of output channels, h is the height of the convolutional layer, and w is the width of the convolutional layer. The pooling layer does not need to calculate the number of parameters. A low number of parameters prevents the model from reaching the weld area features, making the model underfit. A high number of parameters will cause the model to occupy too much memory space, and the memory access cost (MAC) will increase. To measure the model complexity of the candidate area network, the number of floating-point operations, FLOPs, is introduced.

To compute FLOPs, this study assumes convolution is implemented as a sliding window [29] and that the non-linearity function is computed free. For the specified convolution kernel, it is calculated as in equation (11).

$$\text{FLOPs} = 2HW(C_{in}K^2 + 1)C_{out},$$  

where H, W, and C_{in} are height, width, and number of channels of the output feature map, K is the kernel width (assumed to be symmetric), and C_{out} is the number of output channels [30]. For fully connected layer, it is calculated as shown in (7):

$$\text{FLOPs} = (2D_{in} - 1)D_{out},$$  

where D_{in} is the input dimensionality and D_{out} is the output dimensionality. The performance parameters of each RPN are shown in Table 2.

### 3. Result and Analysis

#### 3.1. Recognition Results’ Weld Areas by Different Feature Extraction Networks

This study uses a V-shaped welded steel plate with dimensions of 30.00 cm × 17.00 cm × 0.50 cm, a V-shaped opening angle of 45°, and a weld seam formed in the following welding parameters: the steel plate material is mild steel Q215, the welding method used is melt electrode gas metal arc welding (GMAW) and multilayer multipass welding, the shielding gas is argon, the welding current is 200 A, the welding wire diameter is 2 mm, and welding speed is 2 mm per second. By the recognition effect of the weld seam on the same V-shaped welding plate, the convolutional neural network with the best effect is selected as the RPN, and the model is fine-tuned on this basis for comparative analysis of the recognition of the weld seam in different work scenarios.

Ensuring consistent FLOPs, the network accuracy images and training loss images are plotted using the initial training parameters set in Section 2.2, using six different RPNs in the training and validation sets, respectively. The result is shown in Figure 5.

As shown in Figure 6, the statistical training results obtained by changing different RPNs in the Faster R-CNN model are shown in Table 3, where AP (%) - T is the average precision of the training set and AP (%) - V is the average precision of the validation set.

100, 200, 300, 400, and 500 different weld plate images are taken from the weld plate part test set as input images, and their average detection times are recorded to evaluate the operational efficiency of the algorithm (Figure 7).

Resnet50 is used as the RPN with the highest average accuracy in the training and validation sets, and the fraction of recognized weld areas is up to 94.34% (Figure 8(d)). The Resnet series network provides a shortcut connection mechanism, which ensures a reasonable recognition rate even if the network depth of Resnet50 is 50 layers and the number of layers.

The increase in the number of layers leads to flops 4GFLOPs. Resnet50 average detection time is 92.03 ms. Among them, when Alexnet is used as RPN, the flops are 727MFLOPs, and the average detection time consumes the shortest 18.02 ms because the number of convolutional layers is 5, the network learning effect is not as good as the deep network Figure 8(b), and its average recognition accuracy is 75.50% (Figure 8(a)). In the training process, Alexnet accuracy function and loss function gradually converge, and the loss function curve in the validation set fails to approximate the loss function curve in the training set, so the result weld candidate box has more rear view parts in the weld region candidate box (weld plate region). VGG19 can identify the weld area more completely (Figure 8(f)) with 20GFLOPs and 20,483,904 parameters, as shown in Table 2, resulting in a memory occupation of 548 M and a
Table 2: Performance parameters of RPNS.

| Region proposal networks | Number of network layers | Number of parameters/memory size (MB)       | Flops    |
|--------------------------|--------------------------|---------------------------------------------|----------|
| VGG16                    | 16                       | 62,378,344/(528 MB)                         | 16 GFLOPs|
| VGG19                    | 19                       | 20,483,904/(558 MB)                         | 20 GFLOPs|
| ResNet50                 | 50                       | 25,000,000/(78.1 MB)                        | 4 GFLOPs |
| Alexnet                  | 11                       | 60,965,128/(233 MB)                         | 727 MFLOPs|
| Googlenet                | 22                       | 6,990,270/(51 MB)                           | 2 GFLOPs |
| LeNet                    | 7                        | 431,080/(0.07 MB)                           | 2.32 MFLOPs|

Figure 6: Continued.
Figure 6: Training curves of different RPN. (a) Training curves of LeNet, (b) Training curves of GoogleNet, (c) Training curves of VGG16, (d) Training curves of VGG19, (e) Training curves of ResNet50. (f) Training curves of AlexNet.

Table 3: Six types of RPN network AP values.

| RPN   | VGG16 | VGG19 | Resnet50 |
|-------|-------|-------|----------|
| AP (%) -T | 91.55 | 85.34 | 94.34    |
| AP (%) -V | 86.45 | 70.05 | 84.5     |
| RPN   | Alexnet | Googlenet | LeNet   |
| AP (%) -T | 75.50 | 84.55 | 65.23    |
| AP (%) -V | 66.50 | 73.60 | 45.60    |

Figure 7: Runtime with different numbers of test sets.
A bloated model with a detection time of 70.50 ms. The worst performance is LeNet, which has the shortest average detection time, occupies 2.32 MFlops, and has the smallest number of parameters among the six networks.

Because its network layer is 7, it cannot learn the weld region’s image features in-depth and is more suitable for dealing with scenarios requiring shallow network work and can only identify some features in the weld region under this study (Figure 8(c)).

The average accuracy of VGG16 is slightly lower than that of Resnet50 at 91.55%. Since VGG16 uses several consecutive 3x3 convolutional kernels instead of the larger convolutional kernels in Alexnet compared to AlexNet, for a given receptive field, the performance of the small convolutional kernels with stacking is better than that of the large convolutional kernels because the multilayer nonlinear layers can increase the network depth to ensure learning of more complex patterns, the flops’ value of VGG16 is lower than that of VGG19 in the same series, and the number of parameters is increased to ensure the learning effect and reduce the memory occupied memory space. The average elapsed time is lower, only 25.02 ms, and the confidence level of the identified weld region is up to 91.30%, and the entire weld region is framed (Figure 8(e)). Therefore, the Faster R-CNN model based on the VGG16 candidate region network is selected as the model to identify the weld region.

3.2. Optimization of Models. As shown in the Result of VGG16 in Figure 5, the Faster R-CNN based on VGG16 for weld area recognition object detection network achieves 91.55% accuracy in the training set, and the accuracy in the validation set cannot converge to the same level as the test set. And in order to reduce the loss value to convergence in both training set and test set, different learning rates are chosen to be adjusted, and different numbers of Epochs are added, and Mini-Batch is more suitable for fewer sample data set models, so the optimization of the model is added on this basis. In order to make a comparison with the initial training parameters in Section 2.2, the size of the Mini-Batch is chosen to be 128. the time cost of training. As shown in Table 4.
The plasticity of the VGG16-based Faster-RCNN weld region detection model is verified by adjusting different Epoch and learning rates, and then, the optimal network model is obtained, and the results are shown in Table 5.

In Table 5, using VGG16 as RPN for migration learning at different learning rates, the model's accuracy gradually improves with increasing training times and starts to converge in both the training and validation sets. Reducing the learning rate of the network model can effectively improve the average accuracy of the overall model network at the same number of training sessions. Among them, when the learning rate is 0.1 and 0.01, the network converges very slowly, increasing the time to find the optimal value and not achieving optimal training accuracy. When the learning rate is 0.0001 and Epochs is 30, the accuracy of the network on top of the training set reaches 95.75%, and as Epochs keep increasing, the training accuracy keeps converging, but the growth is slow, and the average accuracy of 45 iterations is lower than before, and if we use continued iterative training, it will increase the algorithm is time complexity, but it does not bring noticeable improvement of training accuracy. When the learning rate is 0.00001, the effect is not as good as the learning rate of 0.0001 because the model hovers around the optimal value, does not converge, and appears to be overfitting. Each learning rate showed an overall decreasing trend after Epochs of 40, which was due to overfitting the network during training. Therefore, in this study, VGG16 is used as the RPN, and the Faster R-CNN with two-stage is trained on the weld plate workpiece dataset, keeping other parameters in Section 2.2 unchanged. The batch is changed to a Mini-Batch of size 128, and at a learning rate of 0.0001 and Epochs of 40 times, the 500 weld plates in the test set were used as the target for the Precision-Recall curve. The initial accuracy threshold was set to 0.8500 using a non-extreme suppression mechanism. The point of \[0, 0.8500\] was decremented in the Recall interval. The model determined the results more excellent than this threshold as positive samples and those less than this threshold as negative samples. As can be seen in Figure 9, as the entry is gradually reduced, more and more weld samples are predicted to be positive, and the curve reaches the \((1, 0)\) coordinate point, indicating that the target detection model can identify the weld seam in the image data as a positive sample. In equation (7), the area enclosed by the curve and the coordinate perimeter is AP. As shown in Figure 9, the area surrounded by its \(X\)- and \(Y\)-axes occupies almost the entire coordinate axis area, indicating the strong recognition performance of the model.

As shown in Figure 10, the color is similar to that of the weld plate workpiece at different angles and working environments, but the network can also identify the weld area of both weld plate workpieces with a score above 90%. As shown in Figure 11, the complex industrial environment with different angles and each large field of view is used for recognition, and there are multiple weld parts in the image, the model network is able to recognize the complete weld area, and some smaller weld areas in the lower-left corner of the image can also be accurately recognized.

In Figure 12, the model is tested in this study by placing it in different working scenarios. Figures 12(a), 12(b) and 12(c) show the weld target detection recognition results at different angles. It can be seen that the recognition accuracy is almost always above 90% and is close to the training accuracy when trained on the training set. The material color in the background of the recognition scene is similar to that of the foreground weld plate, and the object detection network can still accurately frame the weld area's location. Figures 12(d), 12(e) and 12(f) show the effect of weld target detection with the influence of the side light source at a normal overhead angle, which can identify the weld area more completely and obtain an accuracy of about 92%. Figures 12(g), 12(h) and 12(i)
show the weld seam target detection results with laser interference and multiangle recognition. The target detection network can still frame the weld seam in a noisy environment, with a maximum recognition accuracy of 93.17% with laser and multiangle interference. The experimental results in Figure 12 show the robustness of this neural network model for the identification of weld areas in different working environments.
4. Conclusion

The six networks LeNet, Resnet50, Googlenet, VGG16, VGG19, and Alexnet are selected as the RPN of the Faster R-CNN model for training using the migration learning method, and the recognition performance of the six Faster R-CNN models with migration learning is compared and analyzed, and the following conclusions can be drawn:

(1) Faster R-CNN based on VGG16 weld region detection effect is the best, when the model learning rate of 0.01 training times for 30 times, the model in the training set average accuracy of 91.55%, and the validation set accuracy of 86.45%. In the test set average detection time of 25.02 ms, in order to speed up the convergence of the network training process. In order to speed up the convergence of the network training process, the Batch is changed to Mini-Batch size set to 128, which can further improve the accuracy of the model in the training and validation sets, and the loss function convergence is reduced.

(2) When the learning rate is 0.0001, the accuracy decreases with the increase of Epochs in the training set, which is an overfitting phenomenon. After Epochs is 40, the model accuracy grows slowly to prevent the increase of time complexity; this study adopts the learning rate of 0.0001, Mini-Batch is 128, Epochs is 40 training parameters, and the identified network is applied to weld area recognition; as shown in figure, the recognition accuracy and breadth can meet the industrial requirements. The model is also able to accurately frame the weld seam in a mixture of multiplate conditions, multiangle welding conditions, and laser light source interference.

(3) From the point of view of real time, when the target detection model uses a VGG16 convolutional neural network as RPN, it can meet both the requirements of detection accuracy and detection time. The target detection model can achieve industrial recognition efficiency for weld area extraction in different environments, whether the recognition effect is under the ideal viewpoint or weld recognition effect under a complex environment. The industrial requirements for real time can be met.

Data Availability

All data are from laboratory collection and self-consumption.
Conflicts of Interest

All authors have no conflicts of interest and acknowledge the order of attribution.

Acknowledgments

The authors thank the Chinese Ministry of Science and Technology for supporting this research and the professors who participated in this paper. And thanks the second author for technical support during the collection of the dataset. This work was funded by the National Key Research and Development Program of China (2018YFB1306904).

References

[1] L. Liu, H. Chen, and S. Chen, “Quality analysis of CMT lap welding based on welding electronic parameters and welding sound,” Journal of Manufacturing Processes, vol. 74, pp. 1–13, 2022.
[2] G. Peng, B. Chang, G. Wang et al., “Vision sensing and feedback control of weld penetration in helium arc welding process,” Journal of Manufacturing Processes, vol. 72, pp. 168–178, 2021.
[3] H. Sharif and M. Holzel, “A comparison of prefilters in ORB-based object detection,” Pattern Recognition Letters, vol. 93, pp. 154–161, 2017.
[4] Z. Wu, S. Jiang, X. Zhou et al., “Application of image retrieval based on convolutional neural networks and Hu invariant moment algorithm in computer telecommunications,” Computer Communications, vol. 150, pp. 729–738, 2020.
[5] G. E. Jun-Feng and L. U. O. Yu-Pin, “A comprehensive study for asymmetric AdaBoost and its application in object detection,” Acta Automatica Sinica, vol. 35, no. Issue 11, pp. 1403–1409, 2009.
[6] K. Zhang, M. Yan, T. Huang, J. Zheng, and Z. Li, “3D reconstruction of complex spatial weld seam for autonomous welding by laser structured light scanning,” Journal of Manufacturing Processes, vol. 39, pp. 200–207, 2019.
[7] Z. Zhang, G. Wen, and S. Chen, “Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding,” Journal of Manufacturing Processes, vol. 45, pp. 208–216, 2019.
[8] Q. Wang, Lu. Zhang, Y. Li, and K. Kpalma, “Overview of deep-learning based methods for salient object detection in videos,” Pattern Recognition, vol. 104107340 pages, 2020.
[9] Xu. Zhu, K. Chen, B. Anduv, X. Jin, and Z. Du, “Transfer learning based methodology for migration and application of fault detection and diagnosis between building chillers for improving energy efficiency,” Building and Environment, vol. 200, 2021.
[10] P. Martinez, M. Al-Hussein, and R. Ahmad, “Online vision-based inspection system for thermoplastic hot plate welding in window frame manufacturing,” Procedia CIRP, vol. 93, pp. 1316–1321, 2020.
[11] S. Sharifizadeh, I. Biro, and P. Kinnell, “Robust hand-eye calibration of 2D laser sensors using a single-plane calibration artefact,” Robotics and Computer-Integrated Manufacturing, vol. 61, 2020.
[12] A. Araújo, “Polynomial regression with reduced over-fitting—the PALS technique,” Measurement, vol. 124, pp. 515–521, 2018.
[13] S. Yi and Y. Zhou, “Parametric reversible data hiding in encrypted images using adaptive bit-level data embedding and checkerboard based prediction,” Signal Processing, vol. 150, pp. 171–182, 2018.
[14] S. Ren, K. He, R. Girshick, J. Sun, and R.-C. N. N Faster, Towards Real-Time Object Detection with Region Proposal Networks, https://arxiv.org/abs/1506.01497 arXiv: 1506.01497 [cs.CV].
[15] H. Huang, H. Zhou, X. Yang, L. Zhang, L. Qi, and A.-Y. Zang, “Faster R-CNN for marine organisms detection and recognition using data augmentation,” Neurocomputing, vol. 337, pp. 372–384, 2019.
[16] L. Chen, M. Zhou, W. Su, M. Wu, J. She, and K. Hirota, “Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction,” Information Sciences, vol. 428, pp. 49–61, 2018.
[17] Z. Song, L. Fu, J. Wu, Z. Liu, R. Li, and Y. Cui, “Kiwi fruit detection in field images using Faster R-CNN with VGG16,” IFAC-PapersOnLine, vol. 52, no. Issue 30, pp. 76–81, 2019.
[18] N. Dey, Yu-D. Zhang, V. Rajinikanth, R. Pugalenthii, and N. S. Madhava Raja, “Customized VGG19 architecture for pneumonia detection in chest X-rays,” Pattern Recognition Letters, vol. 143, pp. 67–74, 2021.
[19] P. Tang, H. Wang, and S. Kwong, “G-MS2F: GoogleLeNet based multi-stage feature fusion of deep CNN for scene recognition,” Neurocomputing, vol. 225, pp. 188–197, 2017.
[20] A. Deshpande and V. Vania, “Estrela, Prashant Patavardhan, the DCT-CNN-ResNet50 architecture to classify brain tumors with super-resolution, convolutional neural network, and the ResNet50,” Neuroscience Informatics, vol. 1, no. Issue 4, 2021.
[21] A. Unnikrishnan, S. V., and S. K P, “Deep AlexNet with reduced number of trainable parameters for satellite image classification,” Procedia Computer Science, vol. 143, pp. 931–938, 2018.
[22] S. Lin, L. Cai, X. Lin, and R. Ji, “Masked face detection via a modified LeNet,” Neurocomputing, vol. 218, pp. 197–202, 2016.
[23] P. Wang, Y. Lei, Y. Ying, and H. Zhang, “Differentially private SGD with non-smooth losses,” Applied and Computational Harmonic Analysis, vol. 56, pp. 306–336, 2022.
[24] D. Cheng, Y. Gong, X. Chang, W. Shi, A. Hauptmann, and N. Zheng, “Deep feature learning via structured graph Laplacian embedding for person re-identification,” Pattern Recognition, vol. 82, pp. 94–104, 2018.
[25] Z. Avdeeva, E. Grebenyuk, and S. Kovriga, “Combined approach to forecasting of manufacturing system target indicators in a changing external environment,” Procedia Computer Science, vol. 159, pp. 943–952, 2019.
[26] A. Berger and S. Guda, “Threshold optimization for F measure of macro-averaged precision and recall,” Pattern Recognition, vol. 102, 2020.
[27] Z. Guo, Y. Xiao, W. Liao, P. Veeelaert, and W. Philips, “FLOPs-efficient filter pruning via transfer scale for neural network acceleration,” Journal of Computational Science, vol. 55, 2021.
[28] P. Molchanov, S. Tyree, T. Karras, and T. Aila, “Pruning Convolutional Neural Networks for Resource Efficient Transfer Learning,” vol. 3, 2016.
[29] C. R. Brodie, A. Constantin, and A. Lukas, “Flops, Gromov-Witten invariants and symmetries of line bundle cohomology on Calabi-Yau three-folds,” Journal of Geometry and Physics, vol. 171, 2022.
[30] M. Wang, X. Fan, W. Zhang et al., “Balancing memory-accessing and computing over sparse DNN accelerator via efficient data packaging,” Journal of Systems Architecture, vol. 117, 2021.