ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

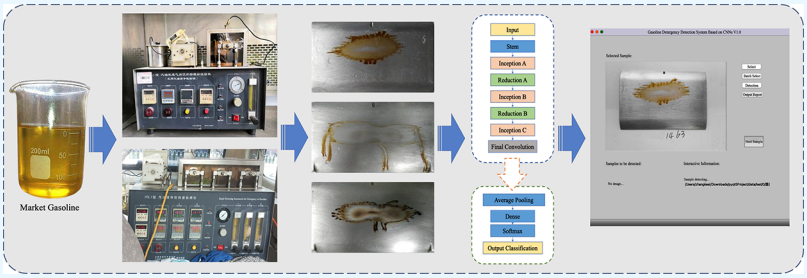
Improvement of a Rapid Method of Detecting Gasoline Detergency Based on the Image Recognition

Rongshuo Zhang, Rencheng Zhu,* Ming Jia, Yujie Pang, Bowen Zhang, Xiaofeng Bao, and Yunjing Wang*

Cite This: *ACS Omega* 2023, 8, 34134−34145

Read Online

ACCESS | Metrics & More | Article Recommendations



ABSTRACT: The detergency of motor gasoline is closely related to vehicle exhaust emissions and fuel economy. This paper proposed an improved method for the rapid detection of gasoline detergency based on the deposit images of test gasoline on aluminum plates produced by a multichannel gasoline detergency simulation test (MGST). The detection algorithm system was structured to recognize the deposit plate images by computer vision based on the convolutional neural networks (CNNs). Compared with the traditional simulation test, the improved MGST method resulted in significant reductions in fuel consumption, cost, and test time. The performance of three transfer learning models (Inception-ResNet-V2, Inception-V3, and ResNet50-V2) and a customized CNN was evaluated in the detection algorithm system, and their detection accuracies reached 94, 94, 88, and 82%. Inception-RsNet-V2 was selected due to its higher accuracy and better robustness. Based on the model interpretation, it is evident that the model undergoes feature extraction from the sediment deposits on the deposit plate. Subsequently, it employed the acquired deposit features to accurately detect gasoline samples that failed to meet detergency standards. This approach was proved to be effective in enhancing the detection process and ensuring reliable results for gasoline detergency evaluation. It is beneficial to environmental protection regulators for managing market gasoline detergency and urban mobile source pollution. In addition, a deposit plate image database should be established to further improve the detection model performance during the environmental regulation.

1. INTRODUCTION

In recent years, global vehicle ownership has increased rapidly, driven by the continuous promotion of urbanization.[1] It leads to a substantial amount of carbon dioxide (CO_2) emissions and also releases harmful pollutants into the atmosphere, such as hydrocarbons (HC), nitrogen oxides (NOx), carbon monoxide (CO), particulate matter (PM), ozone (O_3) and benzene, and polycyclic aromatic hydrocarbons. Gasoline vehicles are one of the major contributors to CO_2 and pollutant emissions.

Fuel quality is a critical factor affecting vehicle emissions, with gasoline detergency serving as a pivotal environmental attribute for evaluation. Vehicle gasoline detergency refers to the performance of vehicle gasoline to reduce or prevent deposits in the engine fuel line, intake valve, and combustion chamber.[2,3] It is an important index that reflects the tendency of vehicle gasoline to generate carbon deposits during use and the potential for removing carbon deposits that have already been generated.[4−6] Adding detergent additives to gasoline is an economical and effective way to improve the detergency of gasoline and keep engines clean.[5,7,8] Previous studies have substantiated that the incorporation of appropriate detergent additives into gasoline can effectively cleanse engine carbon deposits and uphold engine performance, ultimately enhancing fuel efficiency and mitigating emissions.[2,9]

The incorporation of appropriate detergent additives not only leads to the reduction of criteria pollutants emissions but also diminishes emissions of aromatics, alkanes, and carbonyls.

Received: July 24, 2023
Accepted: August 29, 2023
Published: September 7, 2023

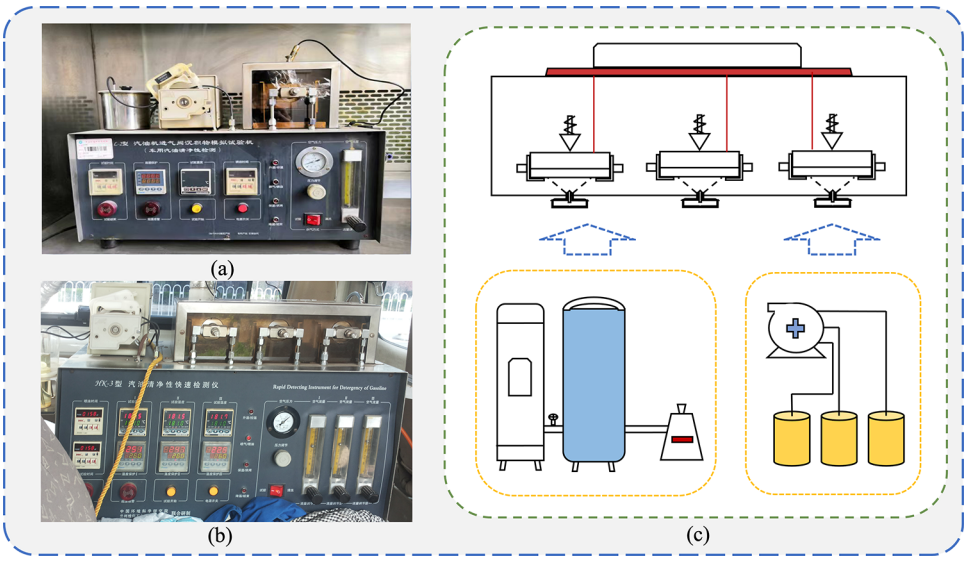


Figure 1. TST equipment and vehicle-mounted MGST (a, b); the MGST technical schematic (c), including the multichannel gasoline detergency tester, compressor, and oil pump.

Sebastian[10] et al. found that fouling on injector tip surfaces and nozzles in the case of direct injection engines correlated well with increased particulate numbers and mass (PN/PM) emissions, consistent with the results reported by Jiang[11] et al. and Houser[12] et al. However, if the detergent containing fuel was used from the beginning of the test, the engine injectors could be cleaned and resulted in decreased particulate emissions. Zhu[13] et al. evaluated 10 different gasoline detergents and selected one with good performance. After adding detergent-containing fuel and driving 1.2×10^4 km, they found that the mass of deposits on the engine intake valve was reduced by 6.8 wt % and that the vehicle emissions of HC, CO, and NOx were obviously decreased by approximately 14.3, 13.6, and 16.7 wt %. Jin[14] et al. found that 9, 8, and 7 of the 14 tested detergent additives contributed to the reduction of CO, HC, and NOx emissions, and most of the tested additives reduced fuel consumption by 0.9 to 3.5 v.%. Improving the gasoline detergency could effectively reduce the pollutant emissions of vehicles.

In principle, the method of detecting the vehicle gasoline detergency should closely emulate real-world driving conditions. Globally, bench tests have been used for gasoline detergency detection, including Ford 2.3 L,[15] M111,[16] and M102.[17] The Ford 2.3 L method was the same as the ASTM D6201 method (Standard Test Method for Dynamometer Evaluation of Unleaded Spark-Ignition Engine Fuel for Intake Valve Deposit Formation). The detergency was evaluated by measuring the increase in the mass of the intake valve deposit (IVD) and combustion chamber deposits (CCDs). The M111 method was equal to the CEC F-20-98 (deposit-forming tendency on intake valves). After the test, the IVD and CCD were collected with a special tool, and the evaluation of the deposits had to be completed within 16 h. This committee also developed the M102 method, referred to as the gasoline engine IVD test method CEC F-05-93 (inlet valve cleanliness in the MB M 102 E engine). The advantage of the bench test methods was that the test results were more representative of the real-world gasoline detergency, but the disadvantages were that they were time-consuming, complicated procedures with high costs.

For these reasons, several new methods have been invented to investigate gasoline detergency. Silva[9] et al. applied infrared spectroscopy to detect whether detergent was added. Rodrigues[18] et al. used near-infrared hyperspectral imaging to identify four detergent and dispersant additives and to detect their concentrations in gasoline. All the methods mentioned above could detect whether detergent had been added to the gasoline, but they did not evaluate whether the gasoline detergency was up to standard in real-driving conditions. A simulation test of the IVD of the gasoline engine method was developed in GB/T 37322-2019 (State Administration for Market Regulation, Standardization administration, 2019). The test gasoline was sprayed with constant pressure and flow of air in a forced-air hood onto an aluminum deposit collector that had been weighed and heated to a specific test temperature. A gasoline film was formed on the surface of the deposit plate, which was continuously heated, baked, and flushed by the atomized test gasoline. Over time, deposits gradually formed on the collector surface. The collector was weighed before and after the test, and the increase was the deposit weight applied to evaluate the detergency of the gasoline. The traditional simulation test (TST) is the current test method widely used in China, which can correct the test gasoline detergency, and the deposition morphology on the plate also reflects some other information. Each test consumes approximately 300 mL gasoline in 2 h and greatly reduces fuel consumption and time compared to bench tests. However, as the gasoline quality gradually improves and the weight gain of deposits gradually decreases, the misjudgment percentage of using the weighing method to evaluate the gasoline detergency cannot be ignored. Moreover, promoting the use of the TST method on a large scale is not convenient due to the influence of experimental operation and measurement accuracy.

With the development of computer hardware, deep learning[19] has made a big impact in the field of image recognition, and it has also been widely used in environmental fields.[7,20] A fine-tuned VGG16 neural network was applied to classify microplastics with an accuracy of 98.33%.[21] Apeksha[22] et al. found that deep learning models could predict air quality more accurately than traditional machine learning models,

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

which was consistent with the findings of ref 23. Li[24] et al. customized a novel CLSTMA model for monitoring water quality in paper industry wastewater treatment systems. Lu[25] et al. customized a deep learning model to classify recyclable waste. Although all of these attempts achieved good performances, no study has applied the deep learning method to gasoline detergency detection to date.

In this paper, a new method containing a multichannel gasoline detergency simulation test (MGST) and a gasoline detergency detection model based on convolutional neural network (CNN) is proposed, which mainly makes the following contributions:

(1) A set of gasoline detergency detection methods is established that can rapidly make deposit plates and detect test gasoline detergency in the field.
(2) A high-accuracy gasoline detergency detection model is built based on CNN.
(3) A new solution for evaluating the detergency of vehicle gasoline is proposed, which is fast, low-cost, and wide applicability.

2. MATERIALS AND METHODS

2.1. Acquisition of Experimental Data. Through years of practical experience, the vehicle-mounted MGST equipment was designed mainly consisting of a car, generator, compressor, oil pump, and multichannel gasoline detergency tester. It has been verified that the testing instrument has a correlation coefficient of 0.94 with the simulation test of IVD of the gasoline engine method.[26] The equipment can produce three samples at the same time, including 92#, 95#, and 98# gasoline deposition samples. The test equipment and technical schematic are shown in Figure 1.

The steps of the deposit plate production are as follows:

1) Polish and grind the deposit plates until the surface is smooth and free of dirt, fix the deposit plates in the inspection hood, clamp them to the heating equipment, and heat them to 180−190 °C.
2) Compress the compressed air to 0.12 MPa, control the flow at 700 L/h, and supply air to the three nozzles, while the metering pump supplies gasoline evenly to the three nozzles at a rate of 4 mL/min. Gasoline and air are fully mixed in the nozzles and sprayed onto the deposit plates at the test temperature.
3) Spray the mixture of gasoline and air continuously for 10 min, followed by heating the deposit plate between 215 and 220 °C and baking for 5 min to oxidize the plate surface to form a gasoline film.
4) After cooling, remove the collector and place it into the image acquisition device. The images are captured and passed to the detection system deployed with the trained deep learning networks for detergency detection.

According to the standard GB/T 37322-2019, using the TST method, the limit value of deposit plate weight gain was 2.0 mg. The gasoline detergency was considered unqualified when the weight gain exceeded 2.0 mg. In the past, MGST was a qualitative analysis method in which the staff subjectively judged whether the test gasoline detergency was qualified or not. The test gasoline samples that were judged to be unqualified or disputed were further tested by the TST method. The test results of the TST method were recorded as the label of the test gasoline detergency.

2.2. Deposit Plate Image Processing. The deposit plate image dataset was divided into two different sets: 80% of the data were used as the training set, and the test set accounted for 20%.[27] There were also parts of deposit plate images tested in past years that were obtained from the Academy of Environmental Sciences. Since the TST method had no special requirements for deposit plate images, the stock images were taken under a nonfixed camera and camera position, and the photo exposure, contrast, and other parameters were not uniform. In addition, the total number of deposit plate images was small. Direct training would likely cause model overfitting and lead to a lack of recognition accuracy.[23,28] To address these problems, normalization and data augmentation were taken in image preprocessing.[29] The transformed images are expanded into the original training dataset to increase the amount of training data. In this paper, image transformation includes many image manipulations from computer vision, such as random cropping, horizontal flipping, vertical flipping, random rotation, random color, and contrast enhancement.

2.3. Detergency Detection Model. CNNs are feedforward neural networks that can be seen as a series of connected layers. A common CNN is cross-stacked by convolutional, pooling, and fully connected (FC) layers,[30] combined with special network layers such as the activation function. Finally, the category probability output is obtained by the *softmax* function.[31] The loss function is a method to calculate the differences between the predicted and original values. The model is trained in a direction that makes the loss value smaller. The convolution layer extracts the features of the image by the convolution operation of the convolution kernel with the image pixel matrix. The convolution layer receives an input feature map in the shape of (image_height, image_width, image_channels), extracts features from it, and generates an output feature map in the shape of (height, width, depth). The formula for calculating the shape of the output feature map is shown in eqs 1 and 2.

$$\text{height} = \frac{\text{image\_height} - \text{kernel\_size} + 2P}{\text{stride}} + 1 \quad (1)$$

$$\text{width} = \frac{\text{image\_width} - \text{kernel\_size} + 2P}{\text{stride}} + 1 \quad (2)$$

where P is the number of paddings. This paper used zero padding to independently control the shape of the output feature map, stride is the step size for the motion of the convolution kernel, and kernel_size is the shape of the kernel. A single kernel can only extract one kind of feature, so a series of kernels were applied to extract different features. The depth is the number of kernels.

The pooling layer can reduce the amount of network data and further extract the main features of the image, including max pooling and average pooling. The FC layer can calculate the summation of all the features extracted by the network, multiply the input from the previous layer with the weight matrix W, and add the bias vector b to perform a linear transformation. Finally, after the nonlinear activation function $f(\cdot)$ operation, the algorithm is shown in eq 3.

$$y = f(W \cdot x + b) \quad (3)$$

In this paper, we applied three activation functions, *ReLu*, tanh, and *softmax*; the first two functions increase the nonlinearity of the model, while the last one obtains the category probability output. As shown in eqs 4, 5, and 6.

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

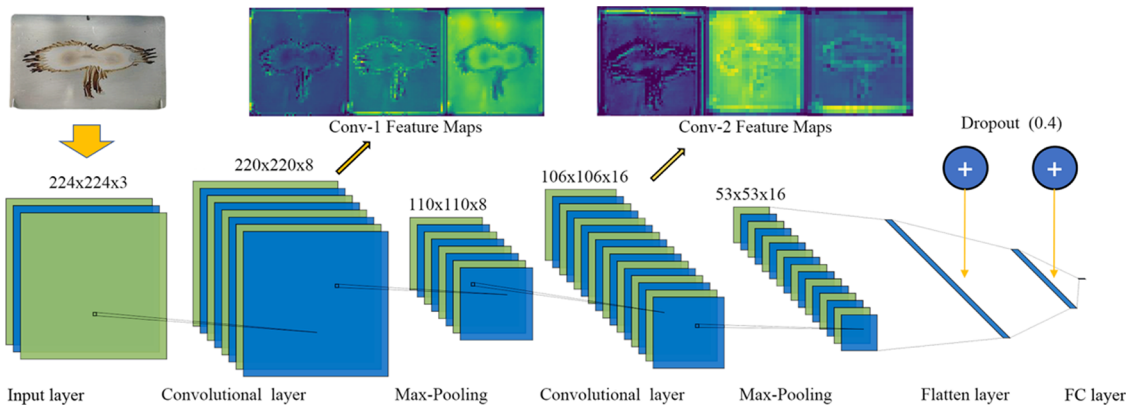


Figure 2. Structure of the customized CNN detection model.

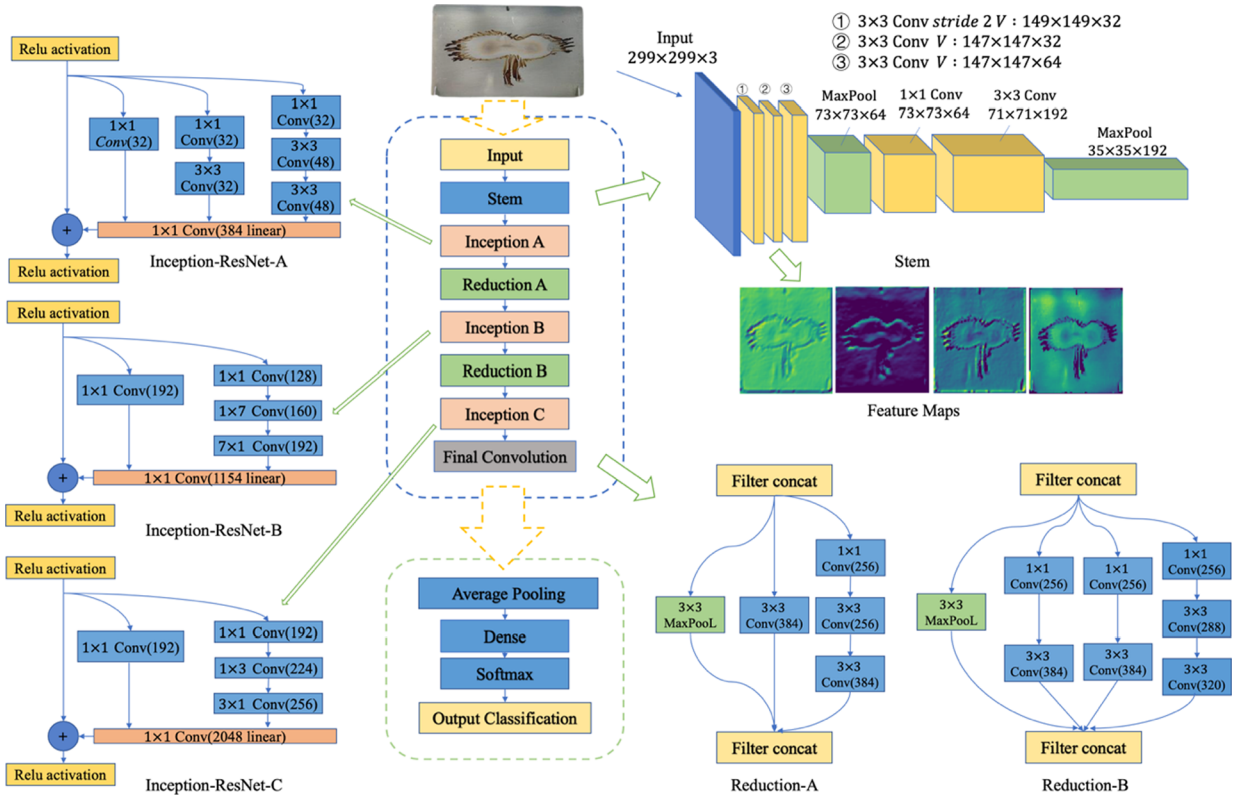


Figure 3. Structure of the Inception-ResNet-V2 networks.

$$ReLU(x) = \begin{cases} x, \text{ if } x \geq 0 \\ 0, \text{ if } x < 0 \end{cases} \quad (4)$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \quad (5)$$

$$softmax(z_j) = \frac{e^{z_j}}{\sum_{i=1}^{n} e^{z_i}} \quad (6)$$

Considering the small amount of training data, it could not achieve a good training effect. Both customized CNNs and transfer learning (TL) were used in this paper for comparison to select the model with the best performance for deployment based on evaluation metrics. The customized CNNs had two convolutional layers (size of 5 × 5, 64 outputs and size of 5 × 5, 128 outputs), and each layer was followed by a max pooling layer 2 × 2; a flattened layer one-dimensional multidimensional input vector and followed by a dropout layer with a dropout rate of 0.4;[32] an FC layer with 50 units followed by a dropout layer with a dropout rate of 0.4; and an FC layer with 2 units followed by *softmax* activation function to output the probabilities of one image under each category. The maximum value was taken as the final classification result.[33] The dropout layers were designed to prevent overfitting. The structure of the customized CNN detection model is shown in Figure 2.

TL could counter the problem of limited size of the training samples because of using pretrained models to extract features to gain better performance.[34] In this paper, we selected three pretrained models that are widely used, including (a) ResNet50-V2,[35] (b) Inception-V3,[36] and (c) Inception-Resnet-V2.[37] The hyperparameters of each model were determined after multiple experiments. The customized CNNs trained all parameters, and the pretrained models that

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

Table 1. TST Test Results

numbering	increase (mg)	class	numbering	increase (mg)	class
01	2.9 ± 0.1	unqualified	14	4.4 ± 0.1	unqualified
02	0.2 ± 0.1	qualified	15	6.6 ± 0.1	unqualified
03	0.1 ± 0.1	qualified	16	1.4 ± 0.1	qualified
04	0.1 ± 0.1	qualified	17	2.8 ± 0.1	unqualified
05	0.4 ± 0.1	qualified	18	1.1 ± 0.1	qualified
06	0.1 ± 0.1	qualified	19	0.1 ± 0.1	qualified
07	0.3 ± 0.1	qualified	20	0.3 ± 0.1	qualified
08	0.4 ± 0.1	qualified	21	0.4 ± 0.1	qualified
09	0.1 ± 0.1	qualified	22	1.4 ± 0.1	qualified
10	8.2 ± 0.1	unqualified	23	1.5 ± 0.1	qualified
11	2.7 ± 0.1	unqualified	24	0.2 ± 0.1	qualified
12	2.3 ± 0.1	unqualified	25	2.4 ± 0.1	unqualified
13	1.6 ± 0.1	qualified			

were only used for feature extraction trained the customized FC layers, leaving other layers frozen. By comparing the classification performance metrics of four models, Inception-ResNet-V2 showed the best performance in this paper.

Inception-ResNet-v2 networks include the Inception block and the ResNet block, and the structure of the networks is shown in Figure 3. The structure of the backbone networks was constructed by the stem networks, including five groups of Inception-ResNet-A blocks, a reduction-A reduction layer, 10 groups of Inception-ResNet-B blocks, a reduction-B reduction layer, and 5 groups of Inception-ResNet-C blocks, followed by a global average pooling layer, a dense layer, and the *softmax* function.

The inception partial structure split the large convolutional kernels into small convolutional kernels in series; for example, the large convolutional kernel of 7 × 7 was split into small convolutional kernels of 1 × 7 and 7 × 1. This asymmetric structure enabled the extraction of more multilevel structural and diversity features. As the core part of the networks, there were three types of Inception-ResNet blocks in the combinatorial network, and the formulas of the three types of blocks were calculated as eq 7.

$$\begin{cases} h(x^{i+1}|\delta) = g(F(h(x^{i}|\delta), \delta = A, B, C \\ L(l, n) = g\left(\sum_{j=1}^{m} L(l, j) \times W_n^l + b_n^l\right) \end{cases} \tag{7}$$

In the above equation, $h(x^{i+1} \mid \delta)$ denotes the i-th block feature map of Inception-ResNet-δ ($\delta = A, B, C$), g denotes the activation function of the network, F denotes the linear transformation function to compute the output features of the multiplexed convolution in the block, $L(l, j)$ denotes the l-th output feature in the j-th layer convolutional network, and W_n^l and b_n^l denote the n-th parameter in the l-layer network.

2.4. Performance Evaluation. To evaluate the performance of the models, based on the confusion matrix, a scientific evaluation index system was established, including the following:

(a) TP (true positives): correctly characterized qualified deposit plate images.

(b) TN (true negatives): correctly characterized unqualified deposit plate images.

(c) FP (false positives): incorrectly characterized unqualified deposit plate images.

(d) FN (false negatives): incorrectly characterized qualified deposit plate images.

Accuracy is the most commonly used classification performance metric, and a greater value means that a larger proportion of observations are predicted correctly. Precision is similar to accuracy, but the difference is that precision evaluates the effectiveness of only one category while accuracy evaluates the overall classification results. Recall is the proportion of every positive observation that is TP. The calculation formulas are shown in eqs 8, 9 and 10.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{10}$$

It is desirable to score high on both precision and recall, but they represent a paradoxical pair of metrics. F_β_score is found to balance precision and recall, and F_β_score with $\beta = 1$ is the F_1_score that is a kind of average used for ratios referred to as the harmonic mean. F_1_score is the most popular adopted metric in binary classification tasks, where the relative contributions of both accuracy and recall are equal. The calculation formulas are as follows in eqs 11 and 12.

$$F_\beta\text{\_score} = \frac{(1 + \beta^2) \times \text{precision} \times \text{Recall}}{\beta^2 \times (\text{precision} + \text{Recall})} \tag{11}$$

$$F_1\text{\_score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

3. RESULTS AND DISCUSSION

3.1. Experimental Results of Detergency Testing. In this study, a total of 25 test samples of market gasoline were selected from Beijing and surrounding areas. Each sample was subjected to the TST and MGST methods with a total of 50 deposit plates. Before the development of the detection model, the detergency of gasoline sample could only be determined qualitatively by human experts based on the morphological characteristics of the deposits. Although subjective judgment can rapidly screen out unqualified gasoline samples, qualitative assessment methods are plagued by inherent uncertainties and present significant challenges for widespread implementation

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

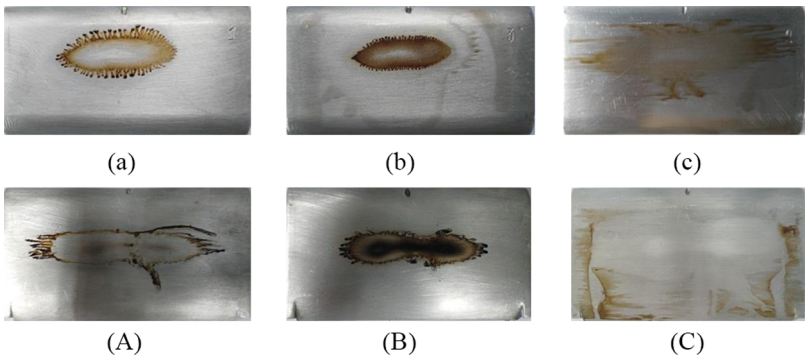


Figure 4. MGST deposit plate images (a−c); the TST deposit plate images (A−C).

and applicability in a larger context. The deposit plate weight gain obtained from the TST was accurately measured to determine whether the detergency weight exceeded the 2.0 mg limit. The deposit plates obtained from the identical set of gasoline samples using the MGST method were subjected to the same detection result as the TST method. The MGST was significantly improved in terms of time, gasoline consumption, and test costs. Compared with the TST method, the most common method, the cost and gasoline consumption of MGST were reduced by 87.5 and 80.0% v.%. Since the MGST could test three gasoline samples at the same time, the time consumption would be reduced by 92.6% if all channels worked simultaneously. Even if only one sample was tested, the time consumption could also be reduced by 77.8%. If the complete detection cycle were considered, the time consumption could be further reduced because the model detection time could be ignored.

The experimental results are shown in Table 1. According to the statistical findings, the measured weight increments of the deposit plates ranged from 0.1 to 6.6 mg, indicating substantial variations in the gasoline detergency within the same city area. Out of the total samples, 17 gasoline samples complied with the specified limit, resulting in a qualification rate of 68.0%. This finding suggests that a majority of the gasoline samples demonstrate satisfactory detergency according to the relevant standards. However, a considerable proportion of the gasoline samples still fall short of meeting the established criteria.

Three sets of deposit plate images were selected to represent different situations for presentation. The results of the comparative testing of selected samples No. 1, 15, and 18 are shown in Figure 4. The first group deposit plate with a weight gain of 2.9 mg exceeded the limit of the test result. However, the center of the deposition was diffuse, and such a situation might be caused by adding insufficient detergent additives or not adding detergent additives as needed. The 15th group deposit plate with a weight gain of 6.6 mg drastically exceeded the limit, in which deposition was greater without obvious spread caused by not adding the detergent additives. The 23rd group test result was qualified with a weight gain of 1.5 mg, for which deposition was clear diffusion. The analysis of deposit characteristics from the comparative images of the deposit plates in the graph indicates a significant correlation between the morphological features obtained through both preparation methods.

In addition, 113 deposit plate images retained by the TST method test were obtained from the Academy of Environmental Sciences, including five qualified gasoline deposit plate images and 108 unqualified. Since there was no need for image recognition in the past test method, most of the deposit plates with qualified test results were washed and reused without obtaining photos. Therefore, most of the previous images were unqualified deposit plate photos. A total of 163 images of deposit plates were obtained in this study, of which 39 were qualified and 124 were unqualified. A total of 130 deposit plate images were included in the training dataset containing 101 unqualified deposit plate images and 29 qualified deposit plate images. In the test dataset, there were 23 unqualified samples and 10 qualified samples.

3.2. Image Preprocessing and Model Training. We applied the Python library to convert 2D images to three-channel images and then resized the images to the same size of 224 × 224 or 299 × 299 px based on the different neural networks. All pixel values were divided by the maximum pixel value to values between 0 and 1. The algorithm is shown in eq 13.

$$x' = \frac{x - X_{\min}}{X_{\max} - X_{\min}} \tag{13}$$

In this paper, x' is the obtained value, x is the pixel value, $X_{\max}$ is 255, and $X_{\min}$ is 0. Then, we used three methods (random cropping, horizontal flipping, and random rotation) to expand the training dataset. The number of training dataset images increased from 130 to 520. Table 2 lists the final number of images in two different categories used for the training and test datasets.

Table 2. Final Image Numbers of the Training and Test Datasets

sets	qualified	unqualified	total
train	116	404	520
test	10	23	33
total	126	427	553

In this study, the Python 3.8 programming language and TensorFlow 2.4.1 framework were utilized as the backend tools. All experiments were performed on a computer equipped with an Intel Core I5-10400F processor operating at 2.90 GHz, 16 GB RAM, and an NVIDIA GeForce GTX 1660 graphics card.

The optimal hyperparameters of each model were finally determined through a series of experiments, as shown in Table 3. Serious overfitting occurred when training the customized CNNs, so early stopping was applied when the model performance was best with epochs = 20 to improve the generalization performance of the model.[38] The vanishing

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

Table 3. Hyperparameters of Different Models

models	LR	batch_size	epochs	activation	optimizer
Inception-ResNet-V2	0.001	16	81	Tanh	SGD
Inception-V3	0.001	16	75	Tanh	SGD
ResNet50-V2	0.001	10	79	Tanh, *ReLu*	SGD
customized CNN	0.001	8	20	*ReLu*	SGD

gradient occurred in ResNet50-V2 networks using only tanh as the activation function, which led to the failure to improve the accuracy of the model. Therefore, when designing the FC layers of the ResNet50-V2 model, a combination of the tanh activation function followed by the *reLu* activation function was added to solve the vanishing gradient while ensuring that all features were learned.

3.3. Performance of the Detection Model. Figure 5 shows the confusion matrix for the model classification result. Figure 5a–d represents Inception-ResNet-V2, Inception-V3, ResNet50-V2, and customized CNNs. The rows of the matrixes represent the predicted classes of deposit plate images, and the columns of the matrixes represent the true categories of deposit plate images. The on-diagonal values in Figure 5a–d indicate the number of correctly classified images, and the off-diagonal values present the number of misclassified images. For example, in Figure 5a, 23 deposit plate images of the unqualified category and eight qualified images were correctly classified, and two images at qualified classes were incorrectly identified. Due to the obvious distinction of the deposition, all four models performed well on the unqualified classes. However, different situations occurred in the detection of 10 qualified deposit plate images because of the small number of qualified training images. The Inception-ResNet-V2 and Inception-V3 models were able to classify eight images correctly, while ResNet50-V2 and customized CNNs were correct in classifying six and four samples. Therefore, increasing the number of training images to enhance the accuracy of the network will be part of our future work, such as the development of deposit plate image databases.

Based on the results of the confusion matrixes and the previously developed evaluation metrics, eqs 8–12, Table 4 presents the performance of each model. The classification accuracy of the Inception-ResNet-V2 and Inception-V3 models on the test set was 94%, compared with 88% for ResNet50-V2 and 82% for customized CNNs. Using small convolution

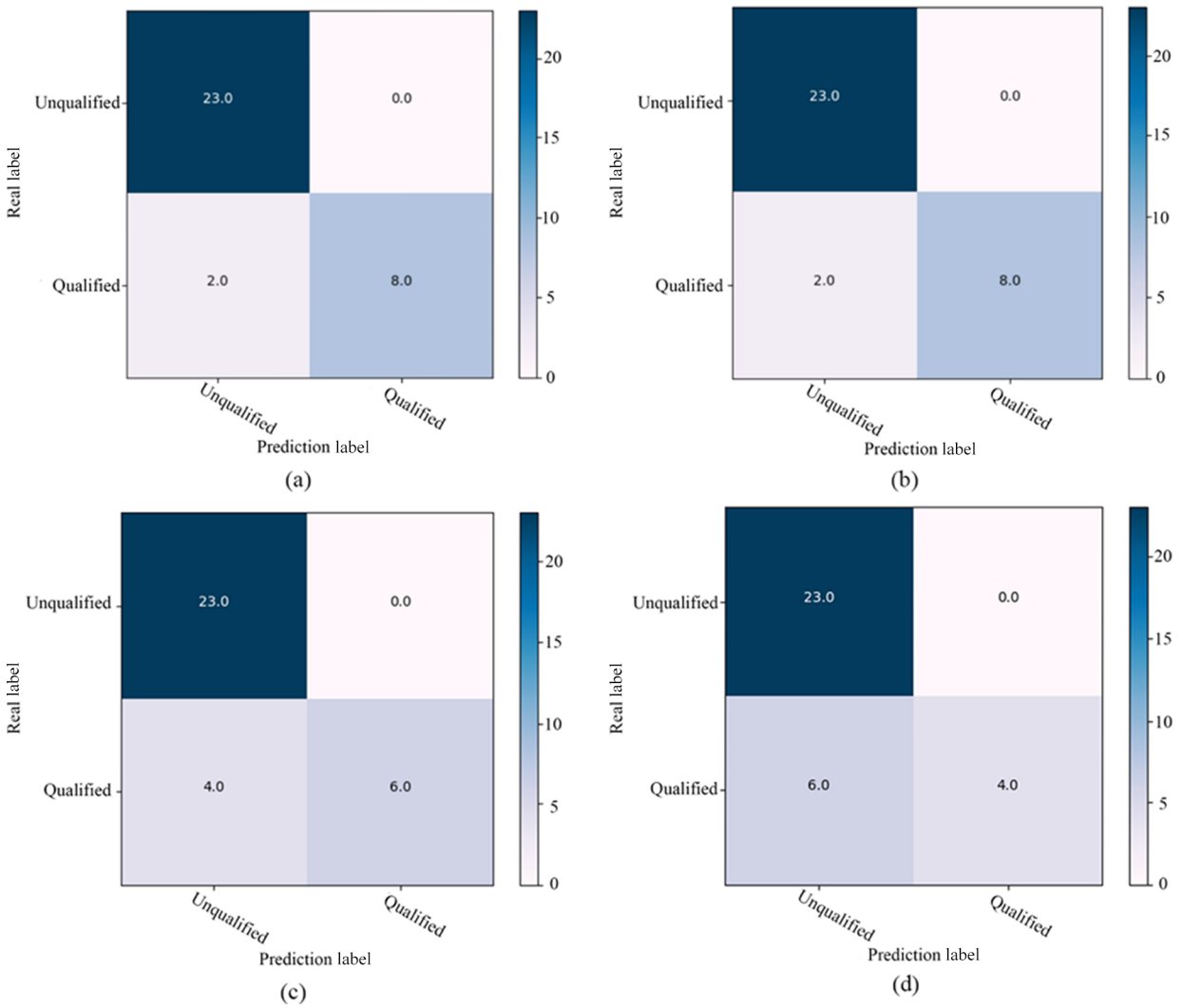


Figure 5. Confusion matrixes of the deposit plate with four methods. The ground truths of the gasoline labels are listed on the vertical axis, while estimations are listed on the horizontal axis. (a) Inception-ResNet-V2, (b) Inception-V3, (c) ResNet50-V2, and (d) customized CNNs.

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134–34145

Table 4. Performance of Four Models

models	sorts	precision	recall	F_1_score	accuracy	times
Inception-ResNet-V2	unqualified	0.92	1.00	0.96	0.94	33.38 min
	qualified	1.00	0.80	0.89		
Inception-V3	unqualified	0.92	1.00	0.96	0.94	11.64 min
	qualified	1.00	0.80	0.89		
ResNet50-V2	unqualified	0.85	1.00	0.92	0.88	22.11 min
	qualified	1.00	0.60	0.75		
customized CNNs	unqualified	0.79	1.00	0.88	0.82	2.23 min
	qualified	1.00	0.40	0.57		

kernels instead of large ones, Inception-ResNet-V2 and Inception-V3 achieved better results, and Inception-V3 cost less time than Inception-ResNet-V2 because of the simpler structure. However, time cost was not the most important evaluation metric, and it should be further validated which model performed better.

The loss values and accuracy of the four models were visualized in the form of line graphs during the training and validation process, as shown in Figure 6. From Figure 6b,B,c,C, both the Inception-V3 and ResNet50-V2 models had a larger change in the loss curve, showing less stability, and the two models performed better on the training set than on the validation set with slight overfitting. Figure 6d,D shows that the customized CNNs were not suitable for small dataset classification. Although the early stopping technique was used, the model still exhibited severe overfitting and poor robustness. As the number of epochs increased, the loss values of Inception-ResNet-V2 steadily declined, and the accuracy gradually increased. The validation accuracy was higher than that of the training set, as illustrated in Figure 6a,A. Although the training time of the Inception-ResNet-V2 model took longer, it had better robustness and generalization ability that was more suitable to deploy as the final gasoline detergency detection model.

The reason for false positive detection errors was also worth noting, and the comparison of the two methods on the 23rd test gasoline sample results is shown in Figure 7a,b, in which the MGST results were misjudged. The weight gain of the 23rd TST deposit plate was 1.5 mg so that the 23rd MGST deposit plate image obtained a qualified label. However, after comparing with other photos of the deposit plates labeled qualified and unqualified, the amount of deposition of the 23rd MGST result was indeed greater, which might be due to some mistakes in the process of producing the deposit plates. Therefore, parallel experiments should be carried out immediately if the gasoline detergency was judged to be unqualified in the actual detection process using the MGST method. If the two test results were unqualified, it was necessary to conduct further TST for final detection. Figure 7c shows a previous deposit plate with a weight gain of 1.9 mg, which was very close to the limit value. Considering only the morphological characteristics of the deposition, it is difficult to determine the gasoline detergency when the weight gain of the deposit plate is close to the limit value. This problem gradually decreases with the increase in training data, but it is difficult to fundamentally solve it.

3.4. Model Interpretation. Deep learning models are widely regarded as black-box models due to their ability to extract features and make predictions in a manner that is difficult for humans to comprehend.[39] Interpreting the features on which the model relies to determine gasoline detergency in a reliable manner is of paramount importance.

In this study, an activation feature map and heatmap of activation values were applied to analyze the model detection process. The first channel of each convolutional layer was chosen to visualize, resulting in a total of 780 feature maps. As shown in Figure 8a, these feature maps show the hierarchical abstraction of input deposit plate image by the model at different layers and illustrate the model performance of feature extraction. From the illustrations, it is evident that the convolutional feature maps extracted by the model exhibit a progressively abstract and sophisticated trend as the network depth increases. In the shallow layers of the feature maps, the model could proficiently extract low-level characteristics such as deposit edges, colors, and textures, accomplishing an initial level of image processing. As the depth of the model increases, the model learns higher-level feature representations in the deeper layers, and the feature maps progressively exhibit more abstract features. This hierarchical feature representation enables the model to better comprehend the intrinsic structure of the input deposit plate images, facilitating accurate detection and analysis of gasoline detergency.

In Figure 8b, the average activation value for each layer in the model is depicted as a heatmap, with each small grid serving as a representation. Notably, the last four small grids have been left empty and are filled with zeros for illustrative purposes. The heatmap reflects the significance of deposit features extracted by the model at different levels, where higher average activation values correspond to features of greater importance, whereas lower values indicate features that the model pays less attention to or is less sensitive to. As depicted in the figure, in the initial shallow layer, the model primarily learns low-level features, such as sediment edges, color, and texture, resulting in relatively higher average activation values. However, as the model depth increases, there is a subsequent decrease in the average activation value, indicating the gradual abstraction of deposit characteristics. Notably, the average activation value of the final model layer experiences an increase, implying a heightened focus on the final abstraction. This observation suggests that the model makes decisions in gasoline detergency detection based on both graphical features of the deposit and the deep-level features extracted by the model.

The activation feature maps provide insights into the process of feature extraction for deposit identification by the model. On the other hand, the class activation heatmap aids in identifying the precise deposit plate image features that drive the model's detection results.[40] This combined analysis facilitates a comprehensive understanding of the model's decision-making process and the crucial features guiding its performance in gasoline detergency detection. The regions

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

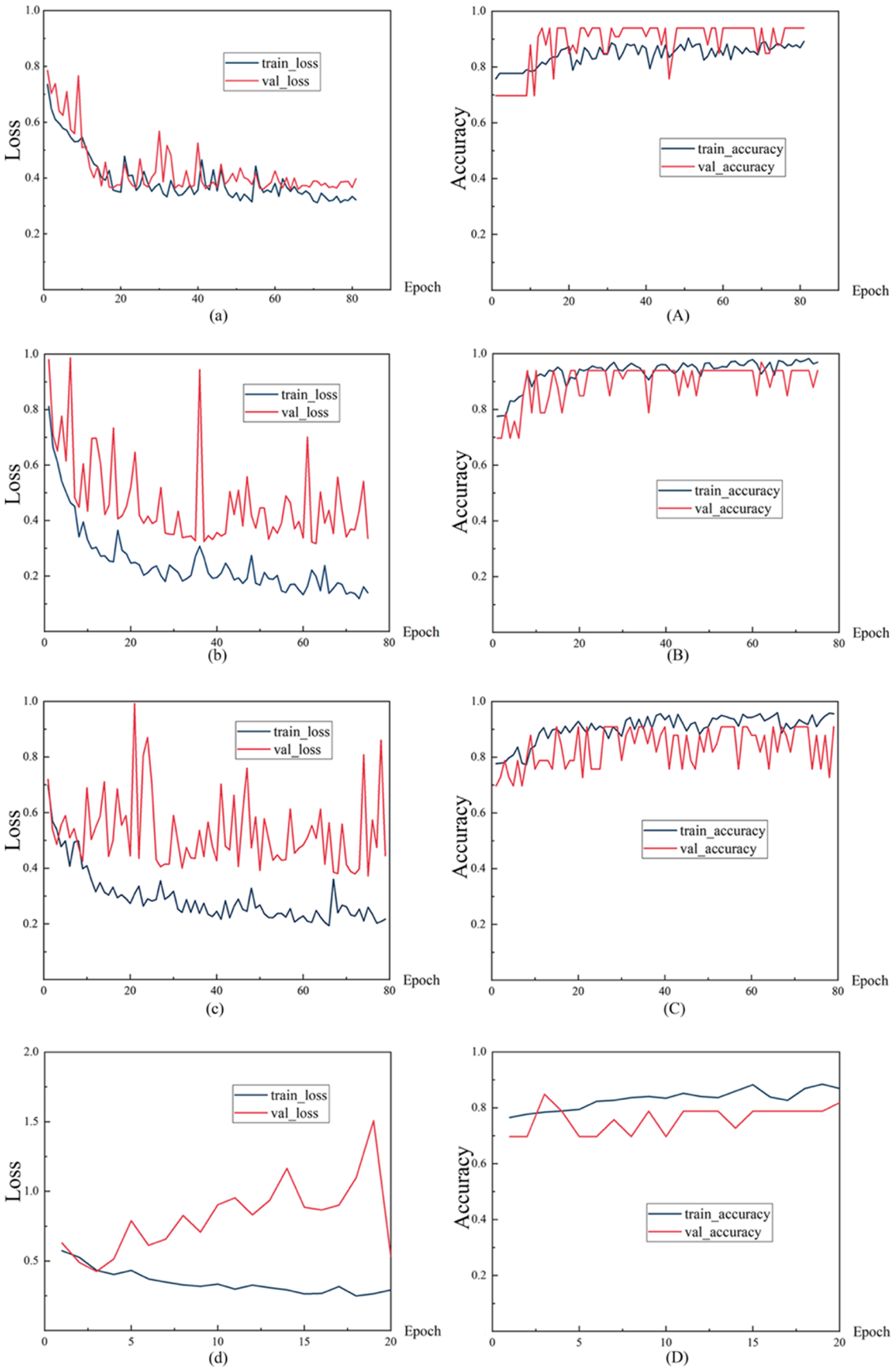


Figure 6. Both the training and validation loss values and accuracy of the four models. The horizontal coordinate indicates the number of epochs. (a, A) Inception-ResNet-V2, (b, B) Inception-V3, (c, C) ResNet50-V2, and (d, D) customized CNNs.

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

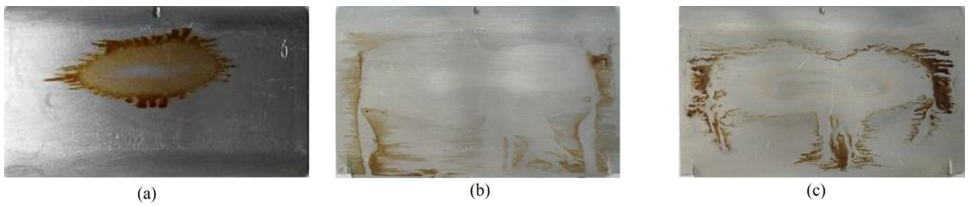


Figure 7. (a) 23rd test result of MGST; (b) 23rd test result of TST; (c) previous deposit plate image with a weight gain of 1.9 mg.

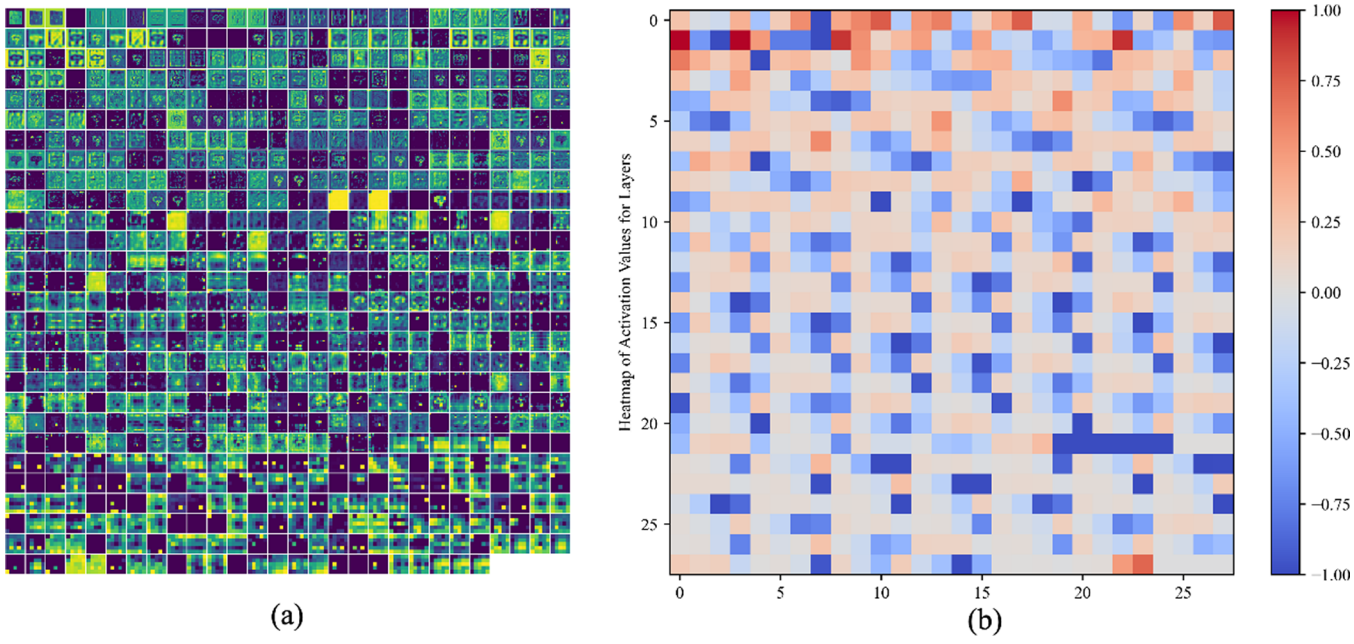


Figure 8. Activation feature maps (a) and the heatmap of activation values (b) for all layers.

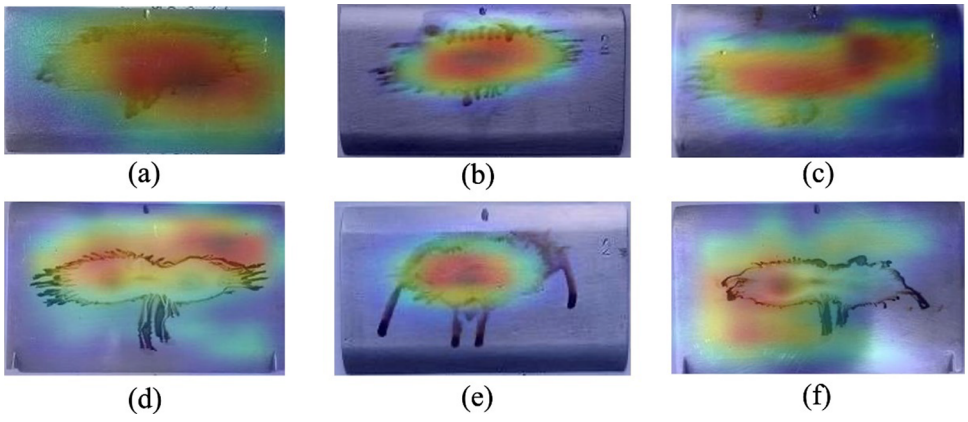


Figure 9. Class activation heatmap of typical deposit plates, including the training set data (a–c), along with the testing set data (d–f).

with colors leaning toward red in the heatmap represent the deposit features on which the model relies to judge the detergency of the gasoline sample. More red color indicates that the model places greater emphasis on these features, which have a stronger impact on determining whether the gasoline sample is qualified. As shown in Figure 9, the red-colored regions are mostly concentrated around the area of deposit, covering aspects such as the quantity, contours, and center of the deposit, et al. This closely resembles the way human experts analyze the morphological features of deposits manually. Consequently, it is evident that the model makes detergency judgments based on specific morphological characteristics of the deposit. This similarity to manual analysis further reinforces the model's ability to accurately assess gasoline detergency by focusing on the precise shape-related features of the deposit.

4. CONCLUSIONS AND FUTURE WORK

In this paper, an improved method for the rapid detection of gasoline detergency was proposed, which used MGST to produce deposit plates and applied a deep learning model to determine. The features of the deposit plate images were extracted by the CNNs to evaluate the gasoline detergency in

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145

the field. The main conclusions drawn from the study are as follows:

Compared with traditional methods, the MGST method achieved good performance in reducing the test time and fuel consumption. It took 77.8% less time to generate a single sample and 92.6% less time to generate multiple samples simultaneously. In addition, the MGST method reduced 87.5 and 80% of cost and gasoline consumption. The detection efficiency was greatly improved, and the exhaust emissions were also significantly reduced in the detection process.

The gasoline detergency detection system was established based on the deep learning. Four different models, including three TF models (Inception-ResNet-V2, Inception-V3, ResNet50-V2) and a customized CNN, were compared in this paper. Their accuracies were 94, 94, 88, and 82%. Compared to the Inception-V3 model, the Inception-ResNet-V2 model had better robustness and generalization ability to perform more effectively in detection. After the model interpretation, it is found that the model can detect the gasoline detergency by recognizing the deposit characteristics of the deposition plate. The regulatory efficiency of the regulatory authorities was greatly improved when the system was deployed on personal computers.

The detection scheme of the new method based on practical experience is as follows:

1) The environmental regulatory staff first takes a 1-L sample of gasoline when arriving at the gas station and then takes 60 mL from the sample for the MGST.
2) This batch of gasoline is deemed to be qualified if the test result of MGST is satisfactory. The staff should immediately take a 350 L gasoline sample from the same gasoline gun if the result is unsatisfactory, and another 60 mL gasoline will be taken for the second MGST.
3) The results of the first and second MGST are compared. If the results are different, it means that the detergency gasoline provided by the gas station is doubtful. The samples are taken back to the laboratory and the small and large samples of gasoline are subjected to the TST separately.
4) The laboratory test report is produced based on the TST results.

Good results were achieved on the existing dataset, but the accuracy can still be improved when the deposit plate weight gain is close to the limited value with the update of vehicle gasoline. It could be widely applied for improving the quality of gasoline detergency due to the advantages of low cost, reduced time consumption, high accuracy, and convenient operation.

The current detection efficiency of the model for gasoline detergency is still insufficient. In the future, the establishment of a database of deposit plate images is proposed to continuously update the version of the detection network. This effort is crucial in ensuring the model's effectiveness in gasoline detergency testing and its potential for broad adoption in practical applications.

AUTHOR INFORMATION

Corresponding Authors

Rencheng Zhu − *School of Ecology and Environment, Zhengzhou University, Zhengzhou 450001, China; State Environmental Protection Key Laboratory of Vehicle Emission Control and Simulation, Chinese Research Academy of Environmental Sciences, Beijing 100012, China;* orcid.org/0000-0001-5455-1850; Email: zhurc@zzu.edu.cn

Yunjing Wang − *State Environmental Protection Key Laboratory of Vehicle Emission Control and Simulation, Chinese Research Academy of Environmental Sciences, Beijing 100012, China*; Email: wangyj@craes.org.cn

Authors

Rongshuo Zhang − *School of Ecology and Environment, Zhengzhou University, Zhengzhou 450001, China*

Ming Jia − *State Environmental Protection Key Laboratory of Vehicle Emission Control and Simulation, Chinese Research Academy of Environmental Sciences, Beijing 100012, China*

Yujie Pang − *School of Ecology and Environment, Zhengzhou University, Zhengzhou 450001, China*

Bowen Zhang − *School of Ecology and Environment, Zhengzhou University, Zhengzhou 450001, China*

Xiaofeng Bao − *State Environmental Protection Key Laboratory of Vehicle Emission Control and Simulation, Chinese Research Academy of Environmental Sciences, Beijing 100012, China; National Engineering Laboratory for Mobile Source Emission Control Technology, Tianjin 300399, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c05350

Author Contributions

R.Z.: methodology, writing—original draft. R.Z.: conceptualization, writing—review and editing, and funding acquisition. Y.P.: visualization. M.J.: project administration. B.Z.: software. X.B.: resources. Y.W.: resources, writing—review and editing.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Key Specialized Research and Development Program in Henan Province (Grant No. 212102310524), the National Natural Science Foundation of China (Grant No. 51808507), the Open Research Fund of State Environmental Protection Key Laboratory of Vehicle Emission Control and Simulation, Chinese Research Academy of Environmental Sciences (No. VECS2022K05), and the Fundamental Research Funds for the Central Public-interest Scientific Institution (No. 2022YSKY-05).

REFERENCES

(1) Olakunle, O.; Kasypi, M.; Adel, G.; Saira, A.; Lai, F. C. Sustainable transition towards greener and cleaner seaborne shipping industry: Challenges and opportunities. *Clean. Eng. Technol.* **2023**, *13*, No. 100628.

(2) Zand, A. D.; Bidhendi, G. N.; Pezeshk, H. The influence of deposit control additives on exhaust CO and HC emissions from gasoline engines (case study: Tehran). *Transp. Res. D Transp. Environ.* **2007**, *12*, 189−194.

(3) Barnes, J. R.; Stephenson, T. *Influence of combustion chamber deposits on vehicle performance and tailpipe emissions*, 1996; p 962027.

(4) Karpov, S. Improving the environmental and performance properties of automotive gasolines. Detergent additives. *Chem. Technol. Fuels Oils* **2007**, *43*, 173−178.

(5) Santos, A. P. F.; Silva, K. K. D.; Dweck, J.; DAvila, L. A. Quantification of detergent-dispersant additives in gasoline by thermogravimetry. *Thermochim. Acta* **2019**, *681*, No. 178400.

(6) Ziwica, K. H. *Intake valve deposits — fuel detergency requirements revisited*, 1987; p 872117.

(7) Shifa, Z.; Kai, Z.; Majid, B.; April, G.; Baikun, L.; Xingmao, M.; Jason, R. Z.; Joshua, S.; et al. Machine learning: new ideas and tools in environmental science and engineering. *Environ. Sci. Technol.* **2021**, *55*, 12741−12754.
(8) Daud, S.; Hamidi, M. A.; Mamat, R. A review of fuel additives' effects and predictions on internal combustion engine performance and emissions. *AIMS Energy* **2022**, *10*, 1−22.
(9) Silva, M. P. F.; Brito, L. R. E.; Honorato, F. A.; Paim, A. P. S.; Pasquini, C.; Pimentel, M. F. Classification of gasoline as with or without dispersant and detergent additives using infrared spectroscopy and multivariate classification. *Fuel* **2014**, *116*, 151−157.
(10) Sebastian, H.; Yannis, H.; Alexander, T.; Christopher, C.; Roger, C.; Kit, G. T.; Paul-Benjamin, R.; Marc, S.; Michael, R. Injector Fouling and its impact on engine emissions and spray characteristics in gasoline direct injection Engines. *SAE Int. J. Fuels Lubr.* **2017**, *10*, 287−295.
(11) Jiang, C.; Xu, H.; Srivastava, D.; Ma, X.; Dearn, K.; Cracknell, R.; Krueger-Venus, J. Effect of fuel injector deposit on spray characteristics, gaseous emissions and particulate matter in a gasoline direct injection engine. *Appl. Energy* **2017**, *203*, 390−402.
(12) Houser, K. R.; Crosby, T. A. *The Impact of Intake Valve Deposits on Exhaust Emissions*; SAE International, 1992, 922259.
(13) Zhu, R.; Bao, X.; Jia, M.; Yue, X.; Liu, Z.; Wang, B. Study on impact of gasoline detergent on vehicle emissions and its detergency. *J. Environ. Eng. Technol.* **2016**, *6*, 307−313. (In chinese).
(14) Jin, B.; Wang, M.; Zhu, R.; Jia, M.; Wang, Y.; Li, S.; Bao, X. Evaluation of additives used in gasoline vehicles in China: fuel economy, regulated gaseous pollutants and volatile organic compounds based on both chassis dynamometer and on-road tests. *Clean Technol. Environ. Policy* **2021**, *23*, 1967−1979.
(15) ASTM International. ASTM D5500: Standard test method for vehicle evaluation of unleaded automotive spark-ignition engine fuel for intake valve deposit formation. 2003.
(16) CEC. The coordinating European council for the development of performance tests for fuels, lubricants and other fluids. *Deposit Forming Tendency on Intake Valves*, 1998, F-20-98.
(17) CEC. The coordinating European council for the development of performance tests for fuels, lubricants and other fluids. *Inlet Valve Cleanliness in the MB M102 E Engine*, 1993, F-05-93.
(18) Rodriguese Brito, L.; Da Silva, M. P. F.; Rohwedder, J. J. R.; Pasquini, C.; Honorato, F. A.; Pimentel, M. F. Determination of detergent and dispensant additives in gasoline by ring-oven and near infrared hypespectral imaging. *Anal. Chim. Acta* **2015**, *863*, 9−19.
(19) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436−444.
(20) Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6999−7019.
(21) Bin, S.; Medhavi, P.; Dian, Y.; Jihui, Y.; Zhengyu, L.; David, P.; Thomas, P.; Kelsey, S.; Elodie, P.; Dwayne, M. R. J.; et al. Automatic quantification and classification of microplastics in scanning electron micrographs via deep learning. *Sci. Total Environ.* **2022**, *825*, No. 153903.
(22) Apeksha, A.; Durga, T. A hybrid deep learning framework for urban air quality forecasting. *J. Cleaner Prod.* **2021**, *329*, No. 129660.
(23) Zhang, Q.; Fu, F.; Tian, R. A deep learning and image-based model for air quality estimation. *Sci. Total Environ.* **2020**, *724*, No. 138178.
(24) Li, X.; Yi, X.; Liu, Z.; Liu, H.; Chen, T.; Niu, G.; Yan, B.; Chen, C.; Huang, M.; Ying, G. Application of novel hybrid deep leaning model for cleaner production in a paper industrial wastewater treatment system. *J. Cleaner Prod.* **2021**, *294*, No. 126343.
(25) Lu, G.; Wang, Y.; Yang, H.; Zou, J. One-dimensional convolutional neural networks for acoustic waste sorting. *J. Cleaner Prod.* **2020**, *271*, No. 122393.
(26) Bao, X.; Zhang, D.; LI, R.; YUE, X.; Huang, X. *Method and equipment for rapid test of gasoline detergency* (In chinese).
(27) Liguo, W.; Lu, Z.; Xian, L.; Jianjie, F.; Aiqian, Z. SepPCNET: Deeping learning on a 3D surface electrostatic potential point cloud for enhanced toxicity classification and its application to suspected environmental estrogens. *Environ. Sci. Technol.* **2021**, *55*, 9958−9967.
(28) Chen, H.; Geng, L.; Zhao, H.; Zhao, C.; Liu, A. Image recognition algorithm based on artificial intelligence. *Neural Comput. Appl.* **2022**, *34*, 6661−6672.
(29) Takahashi, R.; Matsubara, T.; Uehara, K. Data Augmentation Using Random Image Cropping and Patching for Deep CNNs. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *30* (9), 2917−2931.
(30) LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278−2324.
(31) Grave, E.; Joulin, A.; Cissé, M.; Grangier, D.; Jégou, H. *Efficient softmax approximation for GPUs*, 2016, abs/1609.04309.
(32) Smirnov, E. A.; Timoshenko, D. M.; Andrianov, S. N. *Comparison of regularization methods for ImageNet classification with deep convolutional neural networks*: Jeju Island, Korea, 2013; pp 93−98.
(33) Wenqian, D.; Jiahui, Z.; Wenxia, L.; Zhengdong, L.; Huaping, W.; Xi, H. Efficient recognition and automatic sorting technology of waste textiles based on online near infrared spectroscopy and convolutional neural network. *Resour., Conserv. Recycl.* **2022**, *180*, No. 106157.
(34) Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A Comprehensive survey on transfer learning. *Proc. IEEE* **2021**, 43−76.
(35) He, K.; Zhang, X.; Ren, S.; Sun, J. *Deep residual learning for image recognition*, 2015, abs/1512.0338
(36) Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. *Rethinking the inception architecture for computer vision*, 2015, abs/1512.00567.
(37) Szegedy, C.; Ioffe, S.; Vanhoucke, V. *Inception-v4, Inception-ResNet and the impact of residual connections on learning*, 2017, abs/1602.07261.
(38) Liang, H.; Zhang, S.; Sun, J.; He, X.; Huang, W.; Zhuang, K.; Li, Z. *DARTS+: Improved differentiable architecture search with early stopping*, 2019, abs/1909.06035.
(39) Wen, X.; Xie, Y.; Wu, L.; Jiang, L. Quantifying and comparing the effects of key risk factors on various types of roadway segment crashes with LightGBM and SHAP. *Accid. Anal. Prev.* **2021**, *159*, No. 106261.
(40) Ramprasaath, R. S.; Michael, C.; Abhishek, D.; Ramakrishna, V.; Devi, P.; Dhruv, B. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2019**, arXiv:1610.02391v4.

https://doi.org/10.1021/acsomega.3c05350
ACS Omega 2023, 8, 34134−34145