IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 January 2024, accepted 21 January 2024, date of publication 24 January 2024, date of current version 8 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358197

RESEARCH ARTICLE

Event-Based Weak Oil Impurity Detection

MING LI, HUIYU WANG, QIANGHUA CHEN, AND ZHAOXIANG ZONG
School of Electronic Information, Shanghai Dianji University, Shanghai 201306, China

Corresponding author: Zhaoxiang Zong (zongzx@sdju.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 62103256.

ABSTRACT Machine vision-based methods are usually applied to detect impurities in transparent liquids like water or alcohols. Oils have higher viscosity and lower light transmittance, so traditional vision detection methods do not work well. A weak oil impurity detection algorithm was proposed using event stream data. In this method, the binary image of the oil was firstly captured using an event camera. To reduce the noise interference from the event camera, image filtering and morphological operations were applied. Then, image algebra operations were used to remove the oil container's bottom pattern. Finally, impurity detection was performed through the YOLOv5 network. Three common edible oils serve as the experimental samples. Small flying insects, raw material fragments, metal fragments, hair strands, and tin beads of various sizes are selected as the weak impurities. Experiments were performed on a dataset containing 3000 sample images. To the best of our knowledge, existing algorithms can detect impurities with the minimum size of 0.4mm, and most of the experimental samples are transparent liquids. The proposed method can be applied to detect impurities in water and oils, and the detectable impurity size limit is 0.2mm.

INDEX TERMS Impurity detection, dynamic vision, event camera.

I. INTRODUCTION

With the development of edible oil production lines, there is a growing demand for automatic impurity detection. Traditional detection methods primarily depend on manual detection. Such methods rely on heavy human labors, resulting in inefficiency and inconsistency. Long-time manual operations can also lead to mental and visual fatigue [1]. Furthermore, detection reliability may decrease when conducted under inadequate lighting conditions or when bottles have opaque colorations. Sometimes, impurities are easily mistaken for being part of the background, making the detection more challenging. Current impurity detection methods can be divided into two categories: motion features based, and contour features based.

Contour features-based detection methods primarily depend on complex image processing algorithms. They are often employed to extract features like the edge contours and texture shapes of impurities. Zhai et al. [2] focused on the detection of impurities within grain samples. Histogram equalization and the Gauss-Laplacian operator were used to enhance the contrast between grains and impurities. Additionally, they also introduced parameters related to the area of impurities to reduce the false detection points. This method proves highly effective in detecting impurities when significant feature differences exist between the subjects and the impurities. However, it may lead to false-positives in cases that the differences are very subtle. To minimize the false detection rates, Yang et al. [3] treated sorghum as the research object. They extracted various particle features including the size, shape, color, and eccentricity. In addition, they used principal component analysis (PCA) to reduce the data dimension. Support Vector Machine (SVM) was used for classification. As to the detection of impurity in liquids, Yao et al. [4] adopted a preprocessing approach involving bilateral filtering. They applied an enhanced multi-scale wavelet transform to identify the edges of the object and employed feature classification techniques to ascertain the presence of impurities. This method simultaneously classifies impurities and bubbles, and distinguishes them according to the obtained feature information. To address the challenge of fuzzy impurities edges that may not be separated from the changing background, He et al. [5] proposed a two-stream fusion network approach. An embedding network

The associate editor coordinating the review of this manuscript and approving it for publication was Mehul S. Raval.

was applied to extract regional features. Impurities and backgrounds within each training sequence are clustered and labeled. If the features of the impurities to be detected can match those obtained by training, the impurity detection is successful.

Motion feature-based impurity detection methods generally need rotating or sudden-stopping the bottles. This makes the liquid in the bottle shake violently. Xu et al. [6] devised a machine vision-based system for liquid drug detection. The system begins by capturing images of a rotating drug bottle. Then the frame difference method is applied to locate the impurities accurately. And image recognition technology is used to identify whether there are foreign substances in the liquid. Song et al. [7] used ampoule liquid as their experimental subject. They applied an adaptive multi-threshold classification approach to extract potential impurity objects. Through the forward search process, these potential objects are assessed as genuine impurities. He et al. [8] aimed to detect liquid impurities in opaque glass bottles. The textures of the impurities are more complex. This study constructs a long-term cyclic convolutional network for classification purposes. The method goes beyond merely extracting the differences between two frames containing the impurities. Instead, it captures the continuous motion and appearance changes of the impurities across multiple frames.

Unlike liquors, liquid pharmaceuticals, and beverages, oils have distinctive properties such as high viscosity and low transparency. Motion feature-based visual detection methods are not suitable for oil products. Specifically, three factors may affect the oil impurity detection. First, the air bubbles generated during the production process may be confused with impurities. Second, the small sizes of the impurities bring challenges to feature extraction. Third, most impurities tend to gather around the bottle bottom. Its shape may be irregular and diverse, so extracting the contour of every single impurity point becomes difficult. In summary, traditional impurity detection methods are effective for liquids like alcohols and beverages. However, they cannot be directly applied to the detection of oil impurity.

The paper applies dynamic vision technology to oil impurity detection. An event camera is used to capture the images from the bottle bottom. This method ensures the full capture of impurity points. Because of the inherent device interferences and environmental factors, the captured images contain noises. Image filtering and morphological operations are applied to reduce the noises in event images. And algebraic operations are employed to reduce the bottle bottom patterns. Finally, impurities are detected using the YOLOv5 (You Only Look Once) network. The main contributions of this paper are as follows:

1) A weak oil impurity detection method is proposed. An event camera is used to collect the event stream image data, and the YOLOv5 network is used to construct the detection model.

2) A pattern template for bottle bottom is created using event stream data. This template serves to reduce the interferences during the impurity detection process.

3) A dataset of oil impurities with various bottle bottom patterns is produced and extended to 3000 images through data augmentation. Five different impurities are introduced, including small flying insects, raw material fragments, metal fragments, hair strands, and tin beads. The locations of these impurity points are manually marked.

The rest of the paper is organized as follows: Section II introduces the theoretical basis of dynamic vision and object detection algorithms. Section III details the oil impurity detection algorithm. Experimental results are given in Section IV. Finally, conclusions are drawn in section V.

II. RELATED WORKS

A. TRADITIONAL OBJECT DETECTION ALGORITHMS

Traditional object detection algorithms involve three key steps: selecting the detection window, extracting the image window, and designing the classifier. The representative algorithms are Viola-Jones, Histogram of Oriented Gradient (HOG), and DPM (Deformable Part Model). Viola-Jones [9] pioneered the use of integral image techniques to expedite Haar-like input feature calculations. AdaBoost algorithm is used for feature selection to identify key visual features. In addition, a detection cascade technique is implemented to enhance the accuracy and reduce the computational workload. The combination of HOG features and SVM classifiers has been widely used in pedestrian detection algorithms [10]. It effectively captures the appearance and shape of local objects by describing the gradient or directional edge density distribution. The algorithm first divides the image into small cells. The pixel gradient or edge direction histogram is then collected in these cells. Finally, these histograms are merged to create feature descriptors. The DPM model focuses not only on the overall features of the object to be detected, but also on the feature representation of the individual parts of that object. It can be regarded as an extension of HOG [11]. The model includes a root filter with 8×8 resolution and several component filters with a gradient resolution of 4×4. The color-based object detection approach involves several operations, which includes color segmentation, morphological processing, and feature extraction. These operations start with establishing an appropriate segmentation threshold based on the object object's color features. Subsequently, morphological processing and feature extraction are carried out. The machine learning algorithm is used to classify and identify the object.

B. OBJECT DETECTION ALGORITHMS BASED ON DEEP LEARNING

Deep learning-based object detection algorithms can be categorized into two types: the one-stage and two-stage detection methods. The two-stage detection algorithm combines the

Region Proposal Network (RPN) to suggest potential candidate regions, in order to improve the detection accuracy. The one-stage detection algorithm does not need to generate candidate boxes. The problem of locating the object boundary is directly transformed into a regression task, and thus obtains a higher detection speed [12].

In 2014, Girshick et al. introduced the RCNN (Region-based Convolutional Neural Network) network [13]. The algorithm uses candidate regions and convolutional neural networks instead of sliding windows to improve the detection accuracy. However, the redundant computation of overlapping frame features greatly slows the overall detection speed of the network. To address this issue, He et al. introduced the SPPNet (Spatial Pyramid Pooling Networks) [14]. They propose a spatial pyramid pooling layer that allows candidate boxes of arbitrary sizes. This layer generates a fixed-size feature representation map as the final output, effectively mitigating the computational redundancy problem. SPPNet significantly reduces the computation time without compromising detection accuracy. R. Girshick et al. introduced the Fast R-CNN [15] in 2015, building upon the RCNN and SPPNet. Fast R-CNN enables the training of both detectors and edge regressors with the same network configuration. It still relies on a selective search algorithm to locate regions of interest, which can slow down the detection of large real datasets. Faster R-CNN employs a RPN to generate candidate frames instead of relying on the selective search algorithm. This approach significantly enhances the detection speed. In 2017, Lin et al. introduced Feature Pyramid Networks (FPN) [16]. FPN is a network architecture that introduces top-down and horizontal connections. It constructs high-level semantic information across all layers and can accommodate various scales. Cascade RCNN [17] enhances the Faster RCNN by incorporating multiple cascade modules into the backend detector. It employs various Intersection over Union (IOU) thresholds during the training, which leads to a significant accuracy improvement for the two-stage object detection algorithm.

The representative one-stage object detection algorithms are YOLO series and SSD (Single Shot Multi-Box Detector). The YOLO network transforms object detection into a regression problem by utilizing the entire image as the network input. After passing through a neural network, it obtains the bounding box's location and its corresponding class. The SSD algorithm uses a regression approach similar to that of the YOLO model, enabling it to regress both the object's category and location within a single network. This design contributes to higher detection speeds. In the detection process, the region concept used in Faster RCNN is adopted. Combine multiple candidate regions into regions of Interest (ROIs). Objects in the image are detected by classifying and positional regression for each region.

C. EVENT CAMERA-RELATED STUDIES

The event camera is a bio-inspired sensor, which has high dynamic range, low delay, and super time resolution features. Unlike traditional cameras, the event camera doesn't provide a continuous stream of image frames at a fixed rate. On the contrary, it only reports changes in local pixel-level brightness, thus forming an event stream. Each pixel point's output is represented as an event, typically denoted as (t, x, y, p). In the event stream data, (x, y) denotes the pixel's spatial coordinates, t indicates the timestamp of the event. In addition, the polarity p means the brightness change compared to the previous sampling [18]. If the brightness increases, p equals +1 and the event polarity is positive. And p equals −1 when the brightness decreases, in which case the event polarity is negative. Figure 1 shows the comparisons of the data acquisition process between traditional cameras and event cameras as a black dot rotates on a disc [19]. When the black dot rotates with the disk, the traditional camera will lose some of the collected data and produce motion blur due to the influence of the sampling frequency. The data collected by the event camera is relatively complete, and it is not easy to produce motion blur.

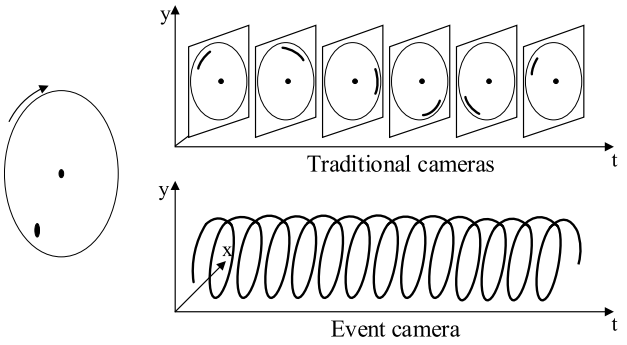

FIGURE 1. Comparisons of the data acquisition process between traditional and event cameras.

The remarkable features of the event camera have promoted its applications in various computer vision areas. It includes robot interaction, 3D reconstruction, high-speed motion estimation, and etc. Rudnev et al. [20] introduced a neural radiation field designed for a color event camera. This approach surpasses baseline rendered image quality, excels in fast motion processing, handles low illumination scenarios effectively, and minimizes data storage demands. Iaboni et al. employed an event camera to achieve real-time detection and tracking of moving robots [21]. Density-based spatial clustering is applied to handle noises and maintain stable tracking even in the absence of events. Ryan et al. [22] developed a driver monitoring system for the detection and tracking of faces and eyes using an event camera. A distinctive fully convolutional recurrent neural network architecture is proposed. This method is not limited by the specific conditions of the driving environment, making it versatile and adaptable. Since event cameras are sensitive to junction leakage current and photocurrent, they tend to amplify noises during the logarithmic conversion process. Therefore, it is necessary to perform denoising on the images captured from event cameras. Mohamed et al. [23] introduced a dynamic background

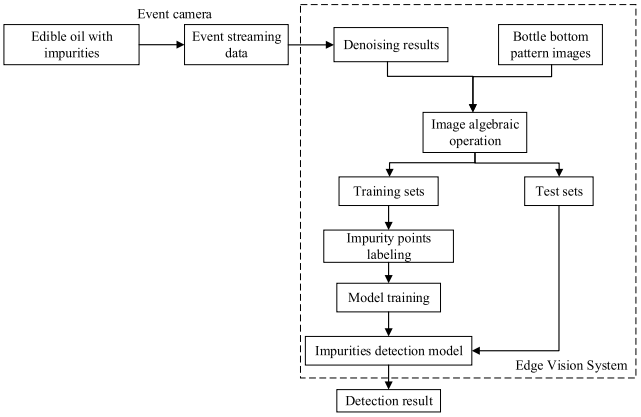

FIGURE 2. Flowchart of the edible oil impurity detection algorithm.

activity filtering method based on the K-nearest neighbors (KNN) algorithm and optical currents. This approach results in a substantial improvement in signal-to-noise ratio. R. Wes Baldwin introduced a neural network-based denoising method named EDnCNN [24], which enhances the contrast sensitivity, promising significant improvements in various applications of neuromorphic vision.

III. OIL IMPURITY DETECTION ALGORITHM

A. IMPURITY DETECTION PROCEDURE

Figure 2 shows the flowchart of our proposed edible oil impurity detection algorithm. Since the event camera is sensitive to ambient light level fluctuations and the influences by optical hardware circuits and environmental factors, the impurity images captured contain considerable noises. Therefore, it is necessary to preprocess the event stream data first. Image filtering and morphological steps are used. Secondly, the image subtraction operation is performed on the collected images containing impurities and the bottle bottom pattern. Then, the preprocessed images are divided into a training set and a test set with a ratio of 4:1. Impurities are annotated using labelme. These labeled images are used to train a YOLOv5 network, and a model is generated. Finally, the model can be used for impurity detection.

B. IMAGE PREPROCESSING

Due to the high sensitivity of the sensor to light, the output image is easier to produce noises. During the process of integral imaging, traditional cameras generally perform the smoothing operation automatically, while event cameras don't. So the images captured by event cameras have obvious noises. What is more, under low contrast conditions, the camera's own logarithmic conversion operation may even amplify the noises. For this reason, it is necessary to perform the noise reduction operation on the output event images. As shown in Figure 3, the operation is mainly composed of two parts: bilateral filtering and morphological operation.

Bilateral filtering is an edge-aware denoising technique that considers both the spatial proximity and pixel-value similarity in an image. Unlike other filtering methods that tend to blur edges when reducing noises, bilateral filtering can effectively reduce the noises while preserving local edge details and regional information. Bilateral filtering operates on images using a combination of spatial distance and Gaussian kernels. It's a nonlinear filter that effectively preserves edges by combining the features of both Gaussian and α-trimmed mean filters. This method considers information in both the spatial and value domains and its kernel is derived by multiplying spatial domain and value domain kernels [25]. The bilateral filter template weights are as follows:

$$w(i,j,k,l) = \exp\left(-\frac{(i-k)^2+(j-l)^2}{2\sigma_d^2} - \frac{\|f(i,j)-f(k,l)\|^2}{2\sigma_r^2}\right) \tag{1}$$

where σ_{d} is the standard deviation of the Gaussian kernel function on the spatial domain, which controls the weights of the pixel positions. σ_r is the standard deviation of the Gaussian kernel function on the pixel value domain, which is used to control the weights of the pixel values. $f(i, j)$ is the pixel value of the image. $f(k, l)$ is the pixel value of the coordinate point at the center of the template window.

In this paper, morphological operation is used to reduce the small noise points in the image. The opening operation consists of an erosion followed by a dilation operation. It effectively reduces small noise points in the image while preserving the shape of relatively larger objects [26]. Based on the bilateral filtering operations, the open operation further reduces the small noise generated by the event camera.

The application of bilateral filtering and morphological operations can effectively reduce most of the noises. But the pattern of bottle bottom and part of the unreduced noise are enhanced after removing the noise. This enhancement could potentially impact the accuracy of impurity detection. In this study, the enhancement of the bottom pattern's edges is reduced by subtraction in image algebraic operation.

The principle of image subtraction operation is to compare the input image with the background. The moving object is detected or segmented according to the changes in gray level and other features. This operation can be used to detect and track moving objects, reduce image backgrounds, and other tasks. Let's denote two input images of size M×N as $f(x, y)$ and $g(x, y)$, and the output image as $h(x, y)$. The pixel of each position in the image $f(x, y)$ subtracts the pixel of the corresponding position in the image $g(x, y)$ to obtain $h(x, y)$.

C. YOLOv5 OBJECT DETECTION NETWORK

The algorithm proposed in this paper uses the YOLOv5 network to create a model for detecting weak impurity points in edible oils. The model comprises three components: a backbone feature extraction layer, a feature fusion layer, and an object detection layer [27]. The network architecture is shown in Figure 4.

FIGURE 3. Flowchart of event stream image denoising. [25], [26].

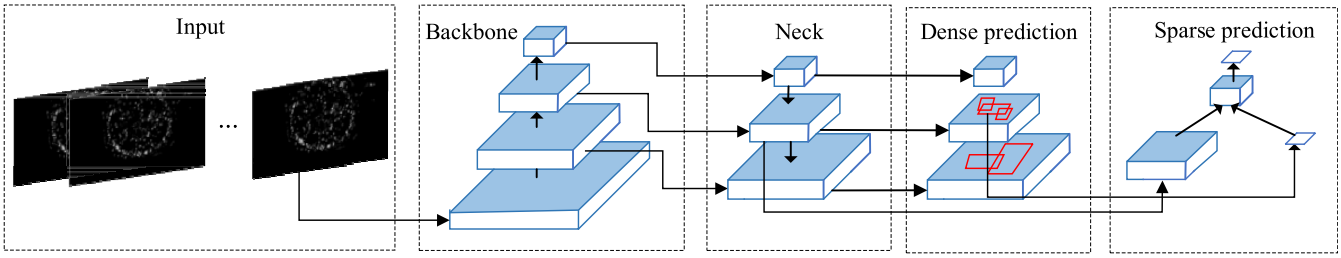

FIGURE 4. Diagram of YOLOv5 detection network. [27].

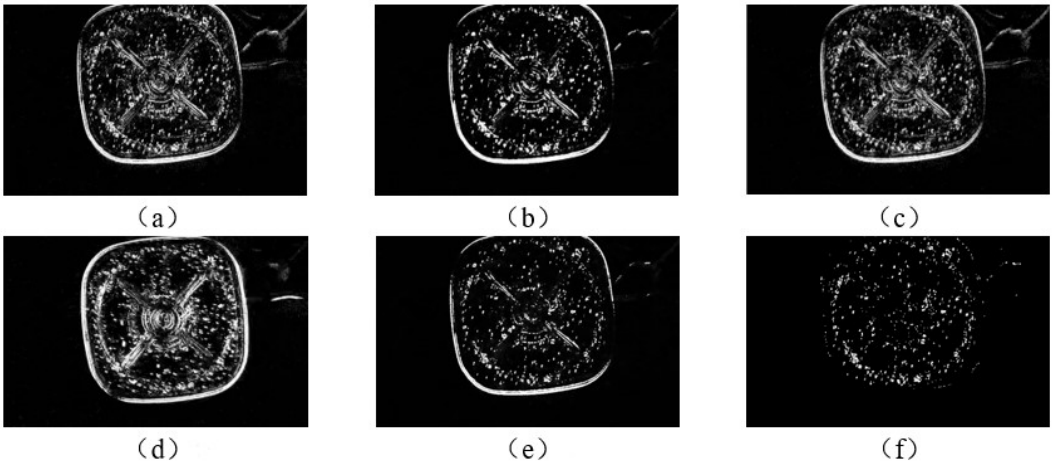

FIGURE 5. Comparative results of preprocessing. (a) the original image containing impurities. (b) the result after 3 × 3 median filtering. (c) the result after 3 × 3 mean filtering. (d) the result after bilateral filtering. (e) the result of the open operation with a 3 × 3 convolution kernel. (f) the result after removing the bottom pattern.

The backbone feature extraction layer comprises the Focus module, CSPDarknet, and SPP (Spatial Pyramid Pooling) module. First, the captured event stream data is input into the Focus module. In the Focus module, the high-resolution image is segmented into several low-resolution feature images using the slicing operation. Then these feature images are processed by the SPP module, converting them into fixed-size feature vectors. Subsequently, they are passed through the CSPDarknet53 module. The residual connection is used to enhance the feature transfer, and the skip connection is used to combine the feature maps of different scales to enhance the object detection accuracy.

To detect impurity objects of varying sizes and positions within the image, a feature fusion layer is employed. This layer uses the FPN structure, which combines feature maps from different scales using up-sampling and down-sampling operations. The top-down component accomplishes feature fusion by integrating up-sampling with coarser-level feature maps, while the bottom-up part combines different levels of feature maps using a convolutional layer. The feature fusion layer combines various impurity feature maps to produce an impurity feature image containing multi-scale information.

In the end, the previously fused impurity feature maps undergo multi-scale object detection through an object detection layer. It includes components such as the convolution layer, pooling layer, and fully connected layer. To begin with, a predefined set of bounding boxes is used within the feature map to generate candidate boxes. These region proposals are then subjected to a classification process to determine whether they correspond to impurities or not. Finally, the full connection layer is used to perform regression operation on each detection frame, and extract the location and size information of impurity points.

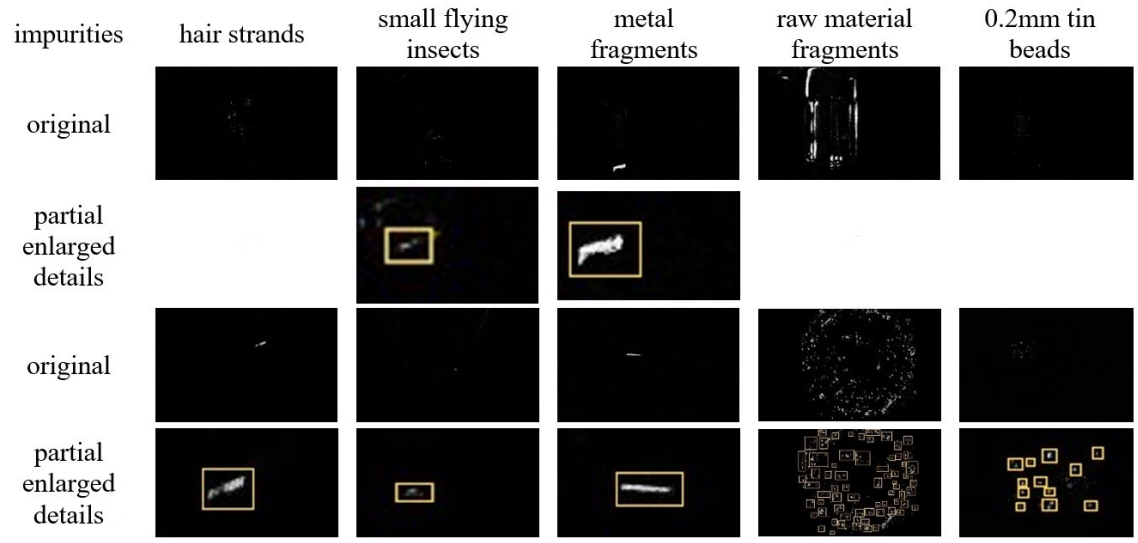

FIGURE 6. The first row shows a side view of the bottle and the third row shows a bottom view. Bottle contours may appear in side view, and the impurities cannot be seen clearly in these images. In the bottom view, impurities can be found without difficulty. For the reader's convenience, the visible impurity points are manually marked with yellow.

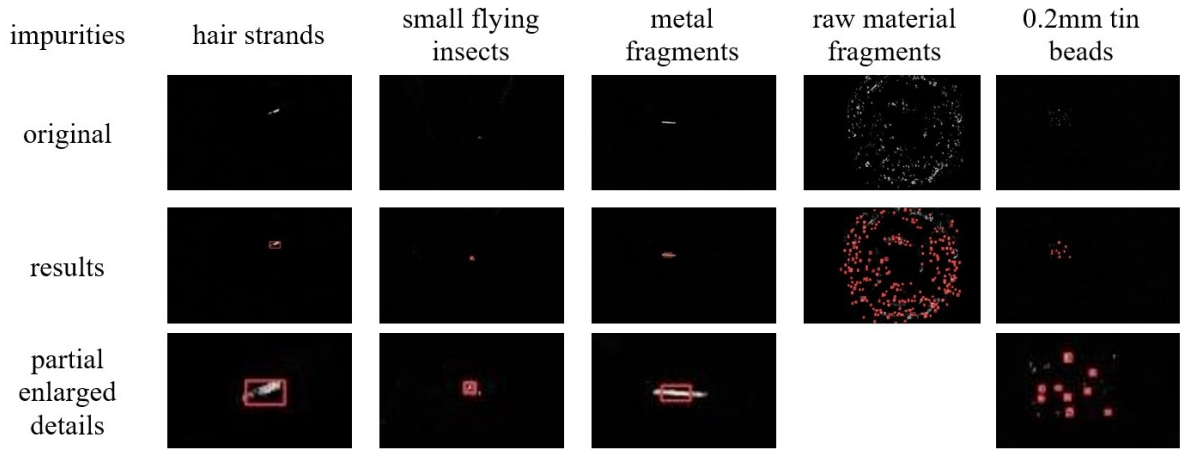

FIGURE 7. Impurity detection results for different objects.

IV. EXPERIMENTS AND RESULTS

A. DATASET CONSTRUCTION

The experiments were conducted on three commonly used oils: soybean oil, rapeseed oil, and olive oil. These oils serve to assess the performance of our proposed impurity detection algorithm. To simulate typical impurities during production, small flying insects, raw material fragments, metal fragments, hair strands, and tin beads were used, among which the diameter of tin beads varies in four sizes: 0.5mm, 0.4mm, 0.3mm, and 0.2mm. A total of 2,000 images were taken using the event camera. The dataset was further expanded to 3,000 images by data enhancement methods such as horizontal mirroring and up-down flipping. Then images were randomly divided into the training, testing, and validation sets.

B. IMAGE PREPROCESSING

Three distinct filtering methods were employed for comparisons. Figure 5 illustrates the results after several filtering operations. The median filter proves effective in eliminating the Salt & Pepper noise. However, under the circumstances of strong noises, image details will even be filtered out. Mean filter is often used to reduce uniformly distributed noises. And, this filtering technique can potentially cause edge information to become blurred and result in detail loss within the image. Bilateral filtering proves highly effective in eliminating noises from the image while preserving edge information and complicated details.

It exhibits superior denoising performance for both Salt & Pepper noise and Gaussian noise. It can be seen from the experimental results that bilateral filtering is most suitable.

The opening operation is combined with bilateral filtering to reduce noises while preserving the overall object shape and edges, as is shown in Fig. 5(e). Additionally, a morphological open operation with a 3×3 convolution kernel is applied to further remove small objects. This operation reduces the adhesion between impurity points and refines the edge of the object image.

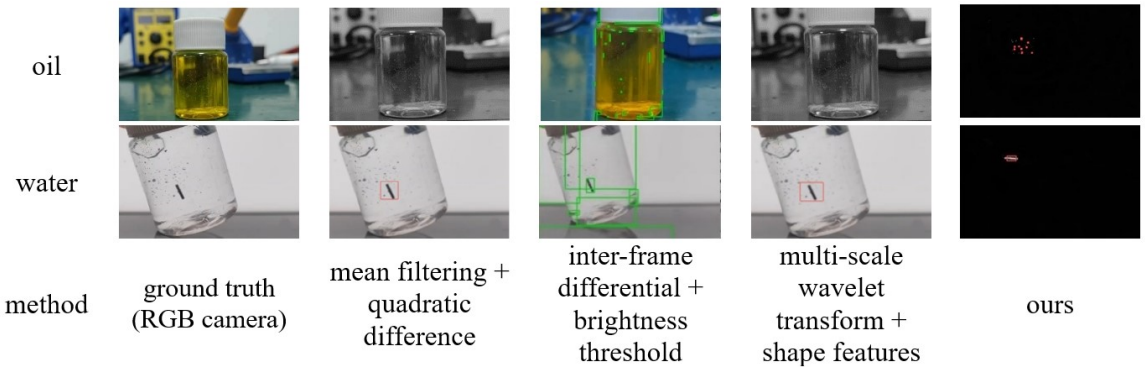

FIGURE 8. Comparison results of different methods. The use of the greyscale second difference method for the detection of impurities has more interference and has certain requirements for the bottle color. The inter-frame difference method has certain requirements for the impurities' distribution area, and the impurities attached to the bottom of the bottle cannot be detected well. It is more sensitive to brightness and has larger background interference. The method using multi-scale wavelet transform and morphological features is not able to detect the impurities in the corners of the bottle. In comparison, the detection method proposed in this paper is more effective.

TABLE 1. Comparisons of impurity detection algorithms.

Devices	Objects	Types of impurities	Size and distribution	Methods	Results
Traditional camera	Water	Black slag, Fibers, Floccule, Broken glasses	Dispersion Length: 0.5mm	Grey scale-based quadratic difference method*	Black slag: 2% White fibers: 4% Floccule: 2% Broken glasses: 10%
		Fiber	Dispersion Length: 1mm	Interframe difference method	Detection accuracy: 90%
		Black slag, White slag, Fiber	Dispersion Length: 0.4mm	Multi-scale wavelet transform + morphological features	Detection accuracy: 95%
Event camera	Water Oil	Small flying insects, Hair strands, Metal fragments, Raw material fragments, Tin beads	Aggregation Diameter: 0.2mm	Bilateral filtering + Morphological features + Background removal + YOLOv5 (ours)	Detection accuracy: 80%

* Only the missed detection rate is reported by the author, and the detection results are calculated manually by 1-missing rate.

To mitigate the influence of the bottle's bottom pattern on impurity detection, images of the oil bottle's bottom without any impurities were initially captured. The bottom image of the oil bottle without any impurities is imported into the algorithm. The pixel size of the image after morphological processing is consistent with that of the image at the bottom of the oil bottle. Next, the pattern boundaries were computed through an iterative process. After that, the pattern of the bottle bottom was extracted and subtracted from the morphology-processed image. This subtraction process removes the same regions in the two images, leaving only the differences between them. The results are shown in Fig. 5(f).

A part of the results obtained after the subtraction operation was selected as the training set to be input into the deep learning network for training. The remaining part served as the test set.

C. EXPERIMENTAL RESULTS

Three experiments were designed to evaluate the proposed impurity detection algorithm. Firstly, different shooting angles are compared. When shooting from the bottle side, the event camera can easily capture the bottle's outline. However, it faces difficulty in clearly capturing impurities within the liquid, such as hair and small particles. Shooting from the bottom solves this problem. The results are shown in Figure 6. The white pixels represent the detected objects.

Secondly, different types of impurities were chosen for comparisons, which includes small flying insects, raw material fragments, metal fragments, hair strands, and tin beads of varying sizes. The detection results are shown in Figure 7. The white pixels represent the raw data output from the event camera, and the red marked represent the detected impurity points.

Thirdly, different fluids were tested. Performance comparisons among the proposed method and traditional detection methods are performed on soya-bean oil and water respectively. The results are illustrated in Figure 8. In the first row, the detection results of each method on edible oils are shown. Tin beads with 0.2mm diameter are used. For Mean Filtering + Quadratic Difference method, it failed to detect the impurities. The Inter-Frame difference+ Brightness thresholding method is sensitive to brightness, and the impurity points detection is easily disturbed by the background environment. Only a very small part of impurities was detected. The Multi-scale Wavelet Transform + Morphological features method effectively reduces the interference of bubbles and noise. However, it still has shortcomings in detecting impurities as small as 0.2 mm. Compared with the other three algorithms, the method proposed in this paper is not affected by background image and illumination. Impurities with the diameter of 0.2 mm were detected, as is shown in red mark.

The second row shows the results of each method as is applied to water, with metal debris as the detection object. The Mean Filtering + Quadratic Difference method can successfully detect the metal debris impurities, as is indicated by the red rectangle. However, the algorithm's performance is sensitive to the color of the bottle. Its detection accuracy decreases notably when the bottle is green or brown. The Inter-Frame Differential + Brightness threshold method is sensitive to brightness changes and strictly depends on bottle sudden stop. The enhanced detection method, which combines multi-scale wavelet transform and morphological features, successfully detected the metal impurities, as is shown with a red rectangle. This method is sensitive to bottle features. If the impurities move over the bottle's edge, it disappears from the captured image and can't be detected. The algorithm in this paper also has a good effect on the detection of impurities in water, where the detection is not easily affected by bottle contour and environmental factors. The detection results are marked with red. Table 1 gives the performance comparisons among the proposed and traditional methods. Compared with other methods, our method can be applied to the detection of impurities in water and oils. Moreover, more types of impurities can be detected, and the size of detectable impurities is smaller.

V. CONCLUSION

In this paper, an event-based weak oil impurity detection algorithm is proposed. Filtering and morphological operations are used to remove the noises from the event camera. The pattern image at the bottle's bottom is extracted and removed to improve the detection accuracy. And a YOLOv5 network is applied to detect the weak impurities. Three commonly used edible oils and several impurities are used for testing. The sample impurities include small flying insects, raw material fragments, metal fragments, hair strands, and tin beads. Tin beads with different diameters are used, ranging from 0.5mm to 0.2mm at an interval of 0.1mm. Experiments were carried out, and as is compared with other detection methods, the minimum impurity size that can be detected in our method reached 0.2mm. The proposed method outperforms previous visible light-based methods in both the detection accuracy and application scenarios. Future work may be focused on combining the appearance features of the impurities with the detection trajectories.

REFERENCES

[1] Z. Shao, Y. Zhang, and W. Zhang, "Detection of impurities in bottled Baijiu based on machine vision," *J. Sichuan Univ., Natural Sci. Ed.*, vol. 56, no. 2, pp. 235–240, 2019.

[2] J. Zhai, C. Zhu, and T. Miao, "Detection of impurity within grain samples by image analysis," in *Proc. 7th Int. Symp. Syst. Softw. Rel. (ISSSR)*, Chongqing, China, Sep. 2021, pp. 50–53, doi: 10.1109/ISSSR53171.2021.00021.

[3] S. Yang, Y. Lin, Y. Li, S. Zhang, L. Peng, and D. Xu, "Machine vision based granular raw material adulteration identification in Baijiu brewing," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Kaohsiung, Taiwan, Jun. 2022, pp. 1–6, doi: 10.1109/IST55454.2022.9827757.

[4] K. Yao, P. Yang, and S. Ma, "A method for detecting impurities in liquids based on feature classification," *Semicond. Optoelectron.*, vol. 40, no. 5, pp. 719–725, 2019, doi: 10.16818/j.issn1001-5868.2019.05.022.

[5] W. He, H. Song, Y. Guo, X. Yin, X. Wang, G. Bian, and W. Qian, "A gallery-guided graph architecture for sequential impurity detection," *IEEE Access*, vol. 7, pp. 149105–149116, 2019, doi: 10.1109/ACCESS.2019.2946861.

[6] J. Xu, X. Li, and S. Shi, "Research on the application of visual detection algorithm for visible foreign bodies in large infusion," *Ind. Control Comput.*, vol. 33, no. 6, pp. 19–21, 2020.

[7] S. Song and Q. Qin, "A new algorithm for small target detection in liquid image sequence," in *Proc. Int. Conf. Intell. Control Inf. Process.*, Dalian, China, Aug. 2010, pp. 234–237, doi: 10.1109/icicip.2010.5564331.

[8] W. He, H. Song, Y. Guo, X. Wang, G. Bian, and K. Yuan, "A trajectory-based attention model for sequential impurity detection," *Neurocomputing*, vol. 410, pp. 271–283, Oct. 2020, doi: 10.1016/j.neucom.2020.06.008.

[9] A. D. Egorov, A. F. Idiyatullin, and A. D. Zakirov, "Comparison of the parametrically optimized implementation of Viola–Jones object detection method and MTCNN," in *Proc. 4th Int. Conf. Control Tech. Syst. (CTS)*, Saint Petersburg, Russia, Sep. 2021, pp. 246–248, doi: 10.1109/CTS53513.2021.9562926.

[10] B. Wang, R. Juanatas, and J. Niguidula, "Vehicle wheel hub recognition method based on HOG feature extraction and SVM classifier," in *Proc. 14th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Yogyakarta, Indonesia, Oct. 2022, pp. 195–199, doi: 10.1109/ICITEE56407.2022.9954076.
[11] A. Ali and M. A. Bayoumi, "Towards real-time DPM object detector for driver assistance," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 3842–3846, doi: 10.1109/ICIP.2016.7533079.
[12] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, and R. Qu, "A survey of deep learning-based object detection," *IEEE Access*, vol. 7, pp. 128837–128868, 2019, doi: 10.1109/ACCESS.2019.2939201.
[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Columbus, OH, USA, Jul. 2014, pp. 580–587.
[14] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015, doi: 10.1109/TPAMI.2015.2389824.
[15] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1440–1448, doi: 10.1109/ICCV.2015.169.
[16] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 936–944.
[17] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into high quality object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 6154–6162, doi: 10.1109/CVPR.2018.00644.
[18] T.-H. Wu, C. Gong, D. Kong, S. Xu, and Q. Liu, "A novel visual object detection and distance estimation method for HDR scenes based on event camera," in *Proc. 7th Int. Conf. Comput. Commun. (ICCC)*, Chengdu, China, Dec. 2021, pp. 636–640, doi: 10.1109/ICCC54389.2021.9674426.
[19] X. Zhou, Q. Liu, S. Chan, and S. Chen, "Overview of vision tracking algorithms based on event camera," *Minicomicrocomp. Syst.*, vol. 41, no. 11, pp. 2325–2332, 2020.
[20] V. Rudnev, M. Elgharib, C. Theobalt, and V. Golyanik, "EventNeRF: Neural radiance fields from a single colour event camera," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Vancouver, BC, Canada, Jun. 2023, pp. 4992–5002, doi: 10.1109/cvpr52729.2023.00483.
[21] C. Iaboni, H. Patel, D. Lobo, J.-W. Choi, and P. Abichandani, "Event camera based real-time detection and tracking of indoor ground robots," *IEEE Access*, vol. 9, pp. 166588–166602, 2021, doi: 10.1109/ACCESS.2021.3133533.
[22] C. Ryan, A. Elrasad, W. Shariff, J. Lemley, P. Kielty, P. Hurney, and P. Corcoran, "Real-time multi-task facial analytics with event cameras," *IEEE Access*, vol. 11, pp. 76964–76976, 2023, doi: 10.1109/ACCESS.2023.3297500.
[23] S. A. S. Mohamed, J. N. Yasin, M. H. Haghbayan, J. Heikkonen, H. Tenhunen, and J. Plosila, "DBA-filter: A dynamic background activity noise filtering algorithm for event cameras," in *Intelligent Computing* (Lecture Notes in Networks and Systems), vol. 283, K. Arai, Eds. Cham, Switzerland: Springer, 2022, doi: 10.1007/978-3-030-80119-9_44.
[24] R. W. Baldwin, M. Almatrafi, V. Asari, and K. Hirakawa, "Event probability mask (EPM) and event denoising convolutional neural network (EDnCNN) for neuromorphic cameras," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 1698–1707, doi: 10.1109/CVPR42600.2020.00177.
[25] L. Jia, G. Wang, and H. Yu, "Bilateral filter algorithm based on noise detection and adaptive variance," *Electron. Des. Eng.*, vol. 30, no. 19, pp. 144–148, 2022, doi: 10.14022/j.issn1674-6236.2022.19.031.
[26] J. Hou, A. Gao, and W. Liu, "Defect image segmentation algorithm based on morphological operations and image fusion," *J. Detection Control*, vol. 43, no. 5, pp. 55–59, 2021.
[27] K. Sudars, I. Namatevs, J. Judvaitis, R. Balašs, A. Nikulins, A. Peter, S. Strautina, E. Kaufmane, and I. Kalnina, "YOLOv5 deep neural network for quince and raspberry detection on RGB images," in *Proc. Workshop Microw. Theory Techn. Wireless Commun. (MTTW)*, Riga, Latvia, Oct. 2022, pp. 19–22, doi: 10.1109/MTTW56973.2022.9942550.

MING LI received the B.Sc. degree from Xi'an Jiaotong University and the Ph.D. degree from Shanghai Jiaotong University. His research interests include industrial machine vision, artificial intelligence, and pattern recognition.

HUIYU WANG received the bachelor's degree in automation from the Xuhai College, China University of Mining and Technology, in 2022. She is currently pursuing the master's degree with Shanghai Dianji University. Her research interests include computer vision and artificial intelligence.

QIANGHUA CHEN received the bachelor's degree in electrical engineering from Shenyang Ligong University, in 2020. He is currently pursuing the master's degree with Shanghai Dianji University. His research interests include computer vision and artificial intelligence.

ZHAOXIANG ZONG received the Ph.D. degree in microelectronics and solid-state electronics from Fudan University, Shanghai, China. She is currently a full-time Faculty Member with Shanghai Dianji University, Shanghai. Her research interests include SOC hardware circuit development based on RISC-V architecture, AI audio and video hardware system development based on the FPGA platform, and chip system-level test equipment development.