IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 12 August 2024, accepted 2 September 2024, date of publication 13 September 2024, date of current version 24 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3459868

METHODS

AFM Imaging Defect Detection and Classification Using Deep Learning

JUNTAO ZHANG[1], JUAN REN[1], AND SHUIQING HU[2]
[1]Department of Mechanical Engineering, Iowa State University, Ames, IA 50011, USA
[2]Bruker Nano Surfaces and Metrology, Santa Barbara, CA 93117, USA

Corresponding author: Juan Ren (juanren@iastate.edu)

This work was supported in part by the National Science Foundation (NSF) under Grant CNS-2409359 and Grant CMMI01751503, and in part by Iowa State University.

ABSTRACT Atomic Force Microscopy (AFM) has been a broadly used platform for high-resolution imaging and mechanical characterization of a wide range of samples. However, this technique can be time-consuming and heavily relies on constant human supervision and human insight for data acquisition and analysis. Recent advancement in artificial intelligence (AI) provides the potential for efficient data analysis for AFM applications. The fusion of AFM with AI for effective image analysis and classification still remains an ongoing research endeavor. In this study, we present a novel AFM image defect detection and classification framework, AFM_YOLO-ResNet, using advanced deep learning (DL) techniques. Central to our approach is a highly integrated DL model that consists of a YOLO image defect detection layer and a ResNet feature extraction and classification layer. The proposed AFM_YOLO-ResNet framework is trained with expert annotated AFM images, and prepared to assess future AFM images from similar samples. Performance of AFM_YOLO-ResNet was validated for AFM image defect detection and classification, and compared with three commonly used transfer learning and computer vision models (Googlenet, Darknet, and YOLOv8). The results with high training and validation accuracies demonstrated that the AFM_YOLO-ResNet framework greatly improves the AFM imaging analysis efficiency.

INDEX TERMS AFM, deep learning, image analysis, identification, classification.

I. INTRODUCTION

Within the family of ''scanning probe microscopy'', atomic force microscopy (AFM) is an excellent platform for high-resolution surface topography imaging and mechanical characterization. It can work with a wide range of materials, such as composites, ceramics, polymers, as well as biological samples [1], [2], [3]. AFM characterization can be carried out in multiple mediums, including ambient air, liquid dispersions, and controlled environments. It is capable of providing 3-D surface/mechanical property profile images [4], [5]. Owing to the superior force, spatial resolution as well as versatility, AFM has been one of the most popular characterization instruments for material studies [6], [7]. However, this technique is not without limitations. AFM characterization is time-consuming, elusive, and prone to errors as it relies on constant human supervision and human insight for image data acquisition and analysis. This is largely attributed to the inherent issues in AFM operations, such as cantilever tip degradation after extended usage, sample damage (especially on soft samples) due to non-optimized force settings, and imaging performed at undesired locations. This requires that experimentalists must manually screen and identify the defects in each AFM image, which often fails to be fast, accurate, and consistent. Hence, the throughput of AFM characterization is greatly limited. Therefore, it is evident that alternative methodologies are needed to alleviate these issues to advance AFM technologies.

Artificial intelligence (AI) and machine learning (ML) tools become the natural choices to reduce the time and labor cost in AFM imaging analysis. The significant advancements in the area of AI and deep learning (DL) in recent years

The associate editor coordinating the review of this manuscript and approving it for publication was Ye Zhou.

have profoundly influenced imaging analysis. Some of these technologies have been developed and implemented for image post-processing and the analysis of various samples, thus refining data quality and improving the overall sample characterization efficiency. Moreover, the materials science community has started integrating AI techniques with AFM to perform high-throughput pattern recognition and advanced data post-processing tasks [8], [9], [10]. Bai and Wu has reported an autodetection method for flexible nanowires using DL technique [8]. Müller et al proposed an ML tool to automate the data quality assessment of AFM force-distance curve measurement of biological tissues [9]. Huang et al. demonstrated an AI-based AFM that is capable of imaging processing in real-time on the fly for ferroelectric materials and electrochemical systems [10]. Additionally, research efforts on combining AI and ML in image-based AFM sample region selection and data modeling have been reported. Rade et al. have developed a DL-based approach to automated sample selection for AFM living cell measurement [5]. Krull et al. proposed an AI framework for AFM good sample region selection and good/bad image sorting [11]. Alldritt et al. developed a DL infrastructure that correlates AFM images with specific molecular configurations [12]. Although these efforts have demonstrated promising results, they are primarily focusing on post-operation data handling for certain limited samples. The broad AFM application still demands better automated approaches that can be adopted for image analysis and identification of a wide variety of samples to substantially improve the throughput.

In this paper, we leverage the DL-based image feature detection and classification tools to propose an AI-driven AFM image defect detection and classification framework. This framework seeks to overcome the inherent limitations of AFM image quality assessment and identification of diverse samples, aiming to substantially cut down on labor costs and time. The AI framework consists of a real-time object detection layer for image defect identification and a convolution neural network (CNN) layer for feature extraction and classification. To ensure defect detection fidelity, this framework has been trained using a dataset of AFM images, each annotated by specialists to highlight different defects. The trained AI framework focuses on the detection and classification of three major AFM image defect profiles: scratches, wobbles, and cracks. Specifically, the proposed framework leverages the ''You Only Look Once''(YOLOv5) [13] as the defect detector and Residual Network (ResNet-34) [14] as an advanced feature delineator and classifier thus, namely *AFM_YOLO-ResNet*. The former is selected for its fast computation, superior detection precision, and lean model dimensions, while the latter is tailor-made for feature curation and classification tasks. Taking AFM images as the input, the framework outputs the detected defect's profile (e.g., the shape of the defect) and its precise coordinates (e.g., location and dimension). These results can serve as the foundation for subsequent sample evaluations and decision-making tool during AFM operation.

For demonstration, AFM_YOLO-ResNet has been trained and tested using AFM images with varying defect types and sizes of two samples: highly ordered pyrolytic graphite (HOPG) and hard disk. The training, validation, and testing accuracies for these samples were 90.13%, 87.68%, and 85.45%, respectively. In addition, AFM_YOLO-ResNet was also compared with existing DL image analysis models, such as Googlenet, Darknet, and the newly announced YOLOv8 using the same dataset. AFM_YOLO-ResNet's accuracies in terms of training, validation, and testing are notably higher than that of the three aforementioned models, highlighting the superiority of this work's approach for defect identification and quality assessment in AFM applications. The key contributions of this work are:

- A successful DL model, AFM_YOLO-ResNet, which integrates YOLOv5 and ResNet-34 for AFM topography image defect detection and classification.
- Custom AFM image preparation algorithms to enhance data quality and consistency.
- Integration of the DL model with data augmentation and transfer learning to enable accurate detection and classification of small objects.
- Thorough model performance evaluation plan to demonstrate the efficacy of the proposed approach and its superiority over existing commonly used object detection and classification models.

With the high accuracies, implementing the proposed AI approach can greatly lower the time and labor cost in AFM quality analysis, and can be potentially extended to other applications where image analysis is critical, such as manufacturing, material engineering, and biomedical engineering.

II. METHODS

A. AFM IMAGE DATA COLLECTION

AFM sample topography images are essential for training and validating this DL model and were acquired during the AFM scanning process of various samples. The sample topography images were acquired from commercial AFM systems (Dimension Icon and BioResolve AFMs, Bruker Nano Inc.).

The images were saved in the .spm format, which is a standard file format for AFM data export. This format retains crucial 3D information about the sample, such as topography, adhesion, and nanomechanical property signals at each lateral imaging location. The choice of .spm format ensures that the data remains compatible with the software and analytical tools commonly used in AFM research.

B. DL MODEL FOR AFM IMAGE DEFECT DETECTION AND CLASSIFICATION

A DL model, AFM_YOLO-ResNet, which could be well-trained to identify a variety of defects in AFM topography images, was developed and explained in the following sections.

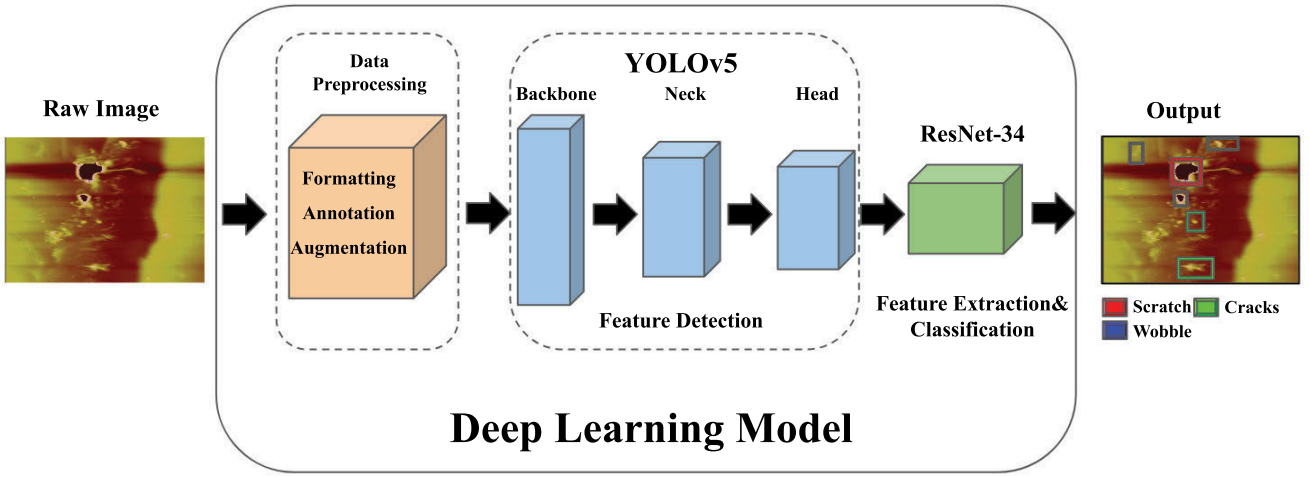

FIGURE 1. Overview of the proposed DL AFM_YOLO-ResNet framework.

1) ARCHITECTURE OF THE AFM_YOLO-ResNet FRAMEWORK

AFM_YOLO-ResNet focuses on detecting the presence of imaging defects (blemishes) and categorizing common defect types, such as cracks, scratches, and wobbles. These defects are often observed in AFM imaging applications.

As depicted in Figure 1, the proposed AFM_YOLO-ResNet model consists of two main components: a defect detector and a feature extractor and classifier.

Despite undergoing several revisions and tests, including the latest version, YOLOv8, YOLOv5 [15] was selected for defect shape detection and localization in this work for its speed, simplicity, and accuracy. Launched in 2020 as an unofficial iteration of the YOLO series, YOLOv5 brings numerous enhancements compared to its predecessors. Although not the most recent version, YOLOv5 significantly improves processing speed, accuracy, and model compactness. Additionally, it offers a more streamlined architecture that reduces parameters and optimizes computing resource utilization. Consequently, it is suitable for object detection tasks involving image feature detection and object localization. Employing a single convolutional neural network [16], YOLOv5 divides input images into grids based on the principle of ''You Only Look Once'' to predict object presence and location within each grid cell. YOLOv5 uses bounding box regression [17] for optimizing object scale variations to enhance its detection capabilities. Each grid cell predicts multiple bounding boxes along with their coordinates, dimensions, and likelihoods of objects within them; conditional class probabilities are also provided for defect detection purposes while bounding boxes serve defect localization needs.

YOLOv5, as illustratively shown in Figure 2, consists of three main components: backbone, neck, and head. It uses CSPDarknet53 [18] as the backbone feature extractor. The main structure of CSPDarknet53 is the stacking of several CBS (Conv+ BatchNorm + SiLU) modules with C3 (Concentrated-Comprehensive Convolution) modules [19], and then connecting to an SPPF (Spatial Pyramid Pooling with Features) module [17]. The CBS module is used to assist the C3 module in feature extraction, while the SPPF module enhances the feature expression ability of the backbone. This setup helps to improve the running speed. An integrated feature fusion network [17], which consists of FPN (Feature pyramid network [20]) and PAN (Path aggregation network [21]), serves as the neck of YOLOv5 for object detection to generate three feature maps P3, P4, and P5 (with the dimensions of 80×80, 40×40 and 20×20) to detect small, medium, and large objects, respectively. FPN performs upsampling of the backbone output to generate feature maps for detecting different scales objects with a top-down feature fusion path. PAN performs downsampling with a bottom-up feature fusion path to further enhance the detection accuracy for different scales objects. The head then executes the confidence score calculation and the bounding box regression of the detected objects [17] using the generated feature maps. The regression process uses the x and y offsets calculated by the model to adjust the center coordinate of the preset prior anchor and the final prediction box size.

Although the superiority of YOLOv5 in object detection has been demonstrated in studies, it is known to have relatively low accuracy in classification, especially for small objects. In this work, an image classification model, ResNet-34 [22], is integrated with YOLOv5 to ensure the classification accuracy. The central innovation of ResNet lies in its utilization of ''residual connections'' [22], [23] that can avoid the gradient vanishing problem, enabling the model to master a residual mapping between the input and output of each layer rather than endeavoring to learn the entire mapping directly. The classical ResNet-34 network

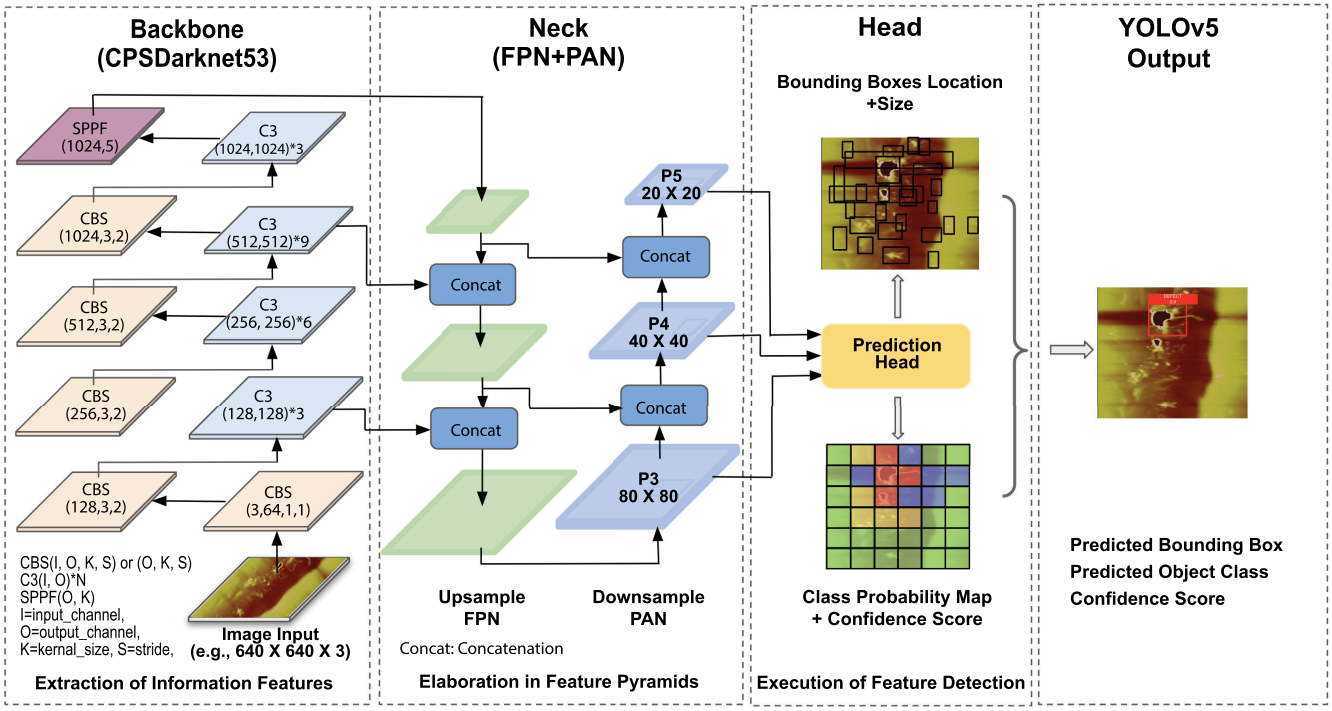

FIGURE 2. Illustrative architecture of YOLOv5 (input image with pixel size of 640 × 640×3 as an example).

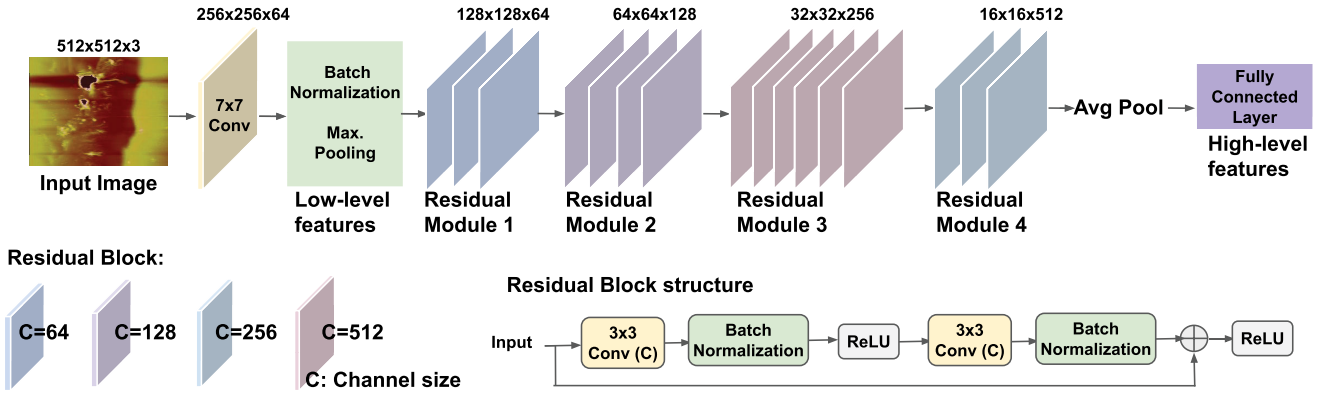

FIGURE 3. Architecture of ResNet-34.

(see Figure 3) contains one convolutional layer, four residual modules, and one fully connected layer [22]. The input image is first processed by the first convolution layer, batch normalization, and maximum pooling to get the low-level features, which are input of the subsequent residual module. A residual module contains repeated residual blocks that have two 3×3 convolutional layers with the same number of output channels. In ResNet-34, the four residual modules have 3, 4, 6, and 3 repeated residual blocks with 64, 128, 256, and 512 channels, respectively. Finally, the fully connected layer outputs the classification results.

To harness the strengths of both YOLOv5 and ResNet-34, the proposed model, AFM_YOLO-ResNet, merges these two sequentially by taking the output of YOLOv5 as the input of the ResNet-34 classifier (see Figure 1). In this configuration, YOLOv5 acts as the "conduit" for defect detection, encompassing both the detection and localization of defects. More precisely, YOLOv5 identifies the defects (blemishes) within the AFM images and pinpoints their locations, which are then forwarded to ResNet-34 for defect extraction and type classification. By integrating these two techniques, the high-level features discerned and categorized by ResNet-34 can be employed to enhance the overall classification performance.

2) DATA PREPROCESSING

a: DATA CONVERSION

In order to speed up the training process of the AFM_YOLO-ResNet, converted JPEG images were used. To streamline the

image processing workflow and enhance efficiency, a custom Python script was developed to seamlessly interface with the AFM Nanoscope Analysis software (Bruker Nano Inc.), allowing for the automatic conversion of all multi-channel .spm files into high-quality JPEG images. This conversion step is critical for data preprocessing, as the majority of deep learning frameworks can directly read and process JPEG images without additional conversions or adaptations.

Employing NanoScope Analysis for batch conversion provided an efficient and reliable way to handle the AFM raw image data. This significantly streamlined the data preparation workflow and ensured the preservation of image quality throughout the conversion process. Compared to the original .spm files, JPEG images are smaller in size, making them more manageable for storage and transmission, yet they still retain ample information for the model to learn from.

b: DATA ANNOTATION

Accurately annotating the defects is labor intensive and essentially important for the success of model learning. For the purpose of defect annotation [24], [25], phase-contrast images were generated from the JPEG files. Then we utilized an open-source tool, OpenLabeling [26], to facilitate manual data annotation. Each defect in the training dataset was manually outlined by designating a bounding box on the images. Through OpenLabeling, defects in each sample image were meticulously marked and categorized. Upon deactivation of the marker, a corresponding text (.txt) file containing the information of the annotated defect data was generated. Specifically, each defect was characterized by three attributes: an index denoting the object class (or defect type), (x, y) coordinates pinpointing the central location of the defect, and the dimensions (both width and height) of the defect. The location and dimensional parameters were normalized with respect to the overall image dimensions.

However, the manual acquisition of these sample images was indeed labor-intensive in terms of time and effort. Notwithstanding our diligent efforts, the resulting annotated dataset remained modest in size.

c: DATA AUGMENTATION

To address the dataset limit issue, we augmented [27] the data to increase the effective size of the dataset and enhance the robustness of machine learning models. The original annotated images undergo various transformations to create the augmented samples. These transformations included mirroring, rotation (both clockwise and counterclockwise by 90Â°), and flipping (both vertically and horizontally).

The augmented samples were then integrated into the training set (see **Algorithm 1**). This not only increased the volume of the training dataset but also introduced a broader range of defect orientations. As a result, the DL model is more resilient and capable of handling diverse defect orientations after training.

This entire process of training data preparation is visualized in Figure 4, which illustrates how data augmentation contributes to the dataset increase.

Algorithm 1 Data Loading and Augmentation

Input:
Dataset Path: *path_to_dataset*

Output:
Augmented Dataset $\mathcal{D}_{aug}$

procedure LoadAndAugmentData(*dataset*)
  Initialize $\mathcal{D}$ as an empty dataset
  Load dataset from *path_to_dataset*
  for each *img_path* in *path_to_dataset* **do**
    Load image load_image(*img_path*)
    Add *img_jpeg* to $\mathcal{D}$
  Initialize $\mathcal{D}_{aug}$ as an empty dataset
  for each *img_jpeg* in $\mathcal{D}$ **do**
    Apply augmentation techniques to *img_jpeg*
    Add augmented images to $\mathcal{D}_{aug}$
  return $\mathcal{D}_{aug}$

C. TRAINING WITH TRANSFER LEARNING

The foundation for training the proposed AFM_YOLO-ResNet DL model lies in the carefully curated training dataset. To ensure comprehensive evaluation and robust training, we employed a meticulous data-splitting approach. Initially, 20% of the dataset was reserved for rigorous testing, while the remaining 80% served as the core for model development.

Within the 80% subset, we further divided it into two distinct sets: 75% of it was designated as the primary training set, and the remaining 25% was allocated as the validation set. This segregation allowed us to harness 60% of the entire dataset for intensive model training while reserving the remaining 20% for continuous validation. This validation subset played a pivotal role in tuning the model's hyperparameters and generalization capabilities throughout the training process.

Advanced transfer learning [28], [29], a process thoroughly detailed in **Algorithm 2**, was implemented to expedite the training process while enhancing overall effectiveness. In stead of training the model from scratch, pretrained weights obtained for traffic lights identification using the Microsoft COCO dataset [30], [31] was used as a start point aiming to reduce the overall training time and resources. Other pretrained weights can be implemented as well to optimize the training efficiency. The network was then further fine-tuned for AFM images using the transfer learning technique. Specifically, the initial training utilized high-resolution images (i.e., those AFM images with a bigger number of scan points at each scan line) from the dataset to ensure the model could discern fine details within the AFM images. Subsequently, the training process was refined

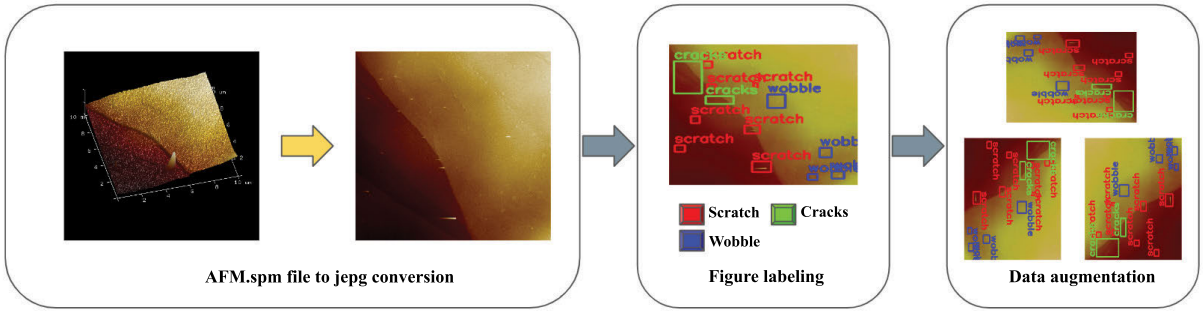

FIGURE 4. Training data preparation.

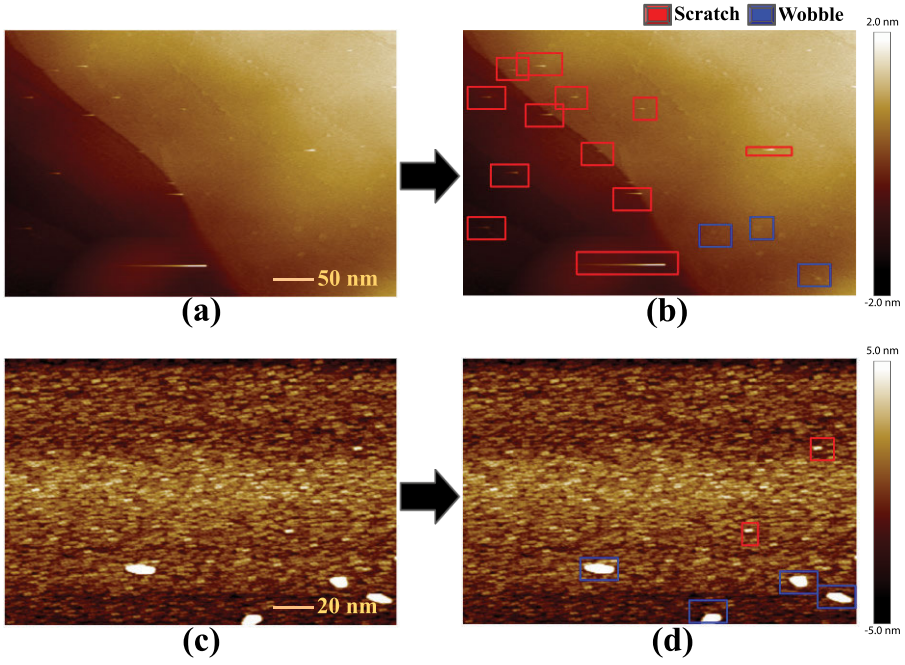

FIGURE 5. Defection detection test results of the HOPG sample (a) and (b), and hard disk surface sample (c) and (d).

using mixed-resolution images, further honing the model's adaptability.

Integration of transfer learning with pretrained model weights significantly contributed to a more robust learning experience, particularly in the recognition and classification of diverse defect types within the intricate AFM images.

III. RESULTS AND DISCUSSION

A. DATA SET

The AFM topographical images of highly ordered pyrolytic graphite (HOPG) and hard disk surfaces formed the basis for the AFM_YOLO-ResNet DL model's training and testing (see Figure 5 for example, AFM images of these two samples). The dataset consists of 200 images for each sample type. These images were divided into three subsets: 60% for training, 20% for validation, and 20% for testing, ensuring no overlap among the subsets to minimize the effect of data discrepancy. Given the relatively small size of the dataset, data augmentation and transfer learning techniques to enhance the model's performance and robustness. These techniques allow the framework to generalize better and achieve high accuracy despite the limited number of images.

B. TRAINING OPTIMIZATION AND MODEL PERFORMANCE

The model's training leveraged the YOLOv5 architecture, complemented by the robust feature extraction capabilities of ResNet-34. Python based PyTorch [32], a popular DL library, was used to develop the proposed AFM_YOLO-ResNet network. The training was conducted up to 1000 epochs, with a focus on optimizing the model training and validation performance. Both Adam [33] and Stochastic Gradient Descent (SGD) [34] optimizers were tested to achieve the optimal performance.

Algorithm 2 Training With Transfer Learning

Input:
Augmented Dataset $\mathcal{D}_{aug}$
Pre-trained YOLOv5 and ResNet-34 models: *model_yolo*, *model_resnet*34
Number of epochs **e**
Learning rate α

Output:
Trained YOLOv5 and ResNet-34 models

procedure TrainWithTransferLearning($\mathcal{D}_{aug}$, *model_YOLO*, *model_resnet*34, **e**, α)
  for $e = 1$ to **e do**
    for each *img_jpeg* in $\mathcal{D}_{aug}$ **do**
      Process the image with YOLOv5:
      `results = model_yolo(img_jpeg)`
      Extract bounding boxes and confidence scores:
      `preds = results.pred[0]`
      for each *box* in *preds* **do**
        Extract ROI:
        `roi = extract_roi(img_jpeg, box)`
        Resize the ROI for ResNet-34:
        `roi_resized = cv2.resize(roi, (224, 224))`
        Perform backpropagation and update models
      Update learning rate α if needed
  return Trained *model_YOLO* and *model_resnet*34

During the process of obtaining the optimal configuration, extensive tests with diverse permutations of network parameters were performed. Adjustments included training batch sizes, epoch numbers, and learning rate variations. The performance of the model under each configuration was rigorously assessed using the mean average precision (mAP) for all defect categories. The outcomes of these evaluations are consolidated in Table 1.

TABLE 1. Performance of the DL model trained with different network configurations. The mAP value over all defect shapes is used for selecting the optimal configuration (highlighted).

Optimizer	Batch	Epochs	Learning Rate	mAP
SGD	16	1000	0.001	62.3
SGD	32	500	0.001	64.4
SGD	32	1000	0.001	66.1
SGD	16	500	0.0001	66.4
SGD	16	300	0.001	69.5
Adam	16	500	0.001	81.0
Adam	32	500	0.0001	83.4
Adam	32	1000	0.001	85.2
Adam	16	1000	0.001	86.2
Adam	32	300	0.001	90.2

From the mAP evaluations, the configuration that delivered the best accuracy is highlighted in Table 1 (the last row), with the mAP score of 90.2% over all defect shapes. It's worth noting that during training, the learning rate initially was set to 0.001 and then changed to 0.0002 when the epoch number was 500. Post the 300-epoch mark, the mAP decreased for both optimizers, signaling overfitting occurred. The final training, validation, and test accuracies are 90.13%, 87.68%, and 85.45%, respectively, as shown in Figure 6 (a).

C. MODEL TEST

AFM_YOLO-ResNet's performance was evaluated using various metrics on the designated test dataset. This evaluation highlighted the model's precision in detecting and classifying common defects. Examples of defect detection outcomes for the two samples are visually depicted in Figure 5, showcasing the precision of defect type, location, and dimensions achieved by this AFM_YOLO-ResNet DL model.

The confusion matrix (CM), shown in Figure 7, was used to assess the classification performance. "True" in CM refers to the actual defect categories, and "Prediction" refers to instances that are classified/predicted by the model. Therefore, the diagonal elements in the CM represent the cases where the model prediction matches the actual "true" result.

Scratches, one of the most common defects in AFM images usually caused by improper probe-sample contact. These were detected with a high precision of 93.62%: 44 out of 47 (along the row) were detected correctly. Meanwhile, the false predictions were low: the false positive detection is 8.51% (4 out of 47), and the false negative rate is 6.38% (3 out of 47). Cracks, another prevalent defect type, were detected with an accuracy rate of 81.13%. The false negative prediction is 5.66% (3 out of 53), and the false positive prediction is 18.87% (10 out 53). Notably, most of the false positive prediction was contributed from background classification. This indicates the model sensitivity needs to be further improved. The prediction performance for wobbles is high as well, with 85.11% (40 out of 47) accuracy, respectively. Similarly seen in the crack detection, misprediction of this defect type as the background is relatively higher compared to other false negative cases.

Overall, the CM demonstrates that the proposed AFM_YOLO-ResNet model was able to detect each type of defect with high accuracy, even considering the false predictions. As most of the false predictions were related to the background, further improvement of the classification performance can be achieved by increasing the model sensitivity.

Furthermore, we have tested the effect of dataset size on the performance of AFM_YOLO-ResNet by checking the testing accuracy for the same full testing subset mentioned earlier while reducing the training and validation dataset size (before augmentation). The results are shown in Figure 8. As it can be seen, with 75% (120 images per sample before augmentation) of the full model development (training and validation) dataset used (same 3:1 training vs. validation split ratio), the training, validation, and testing accuracies

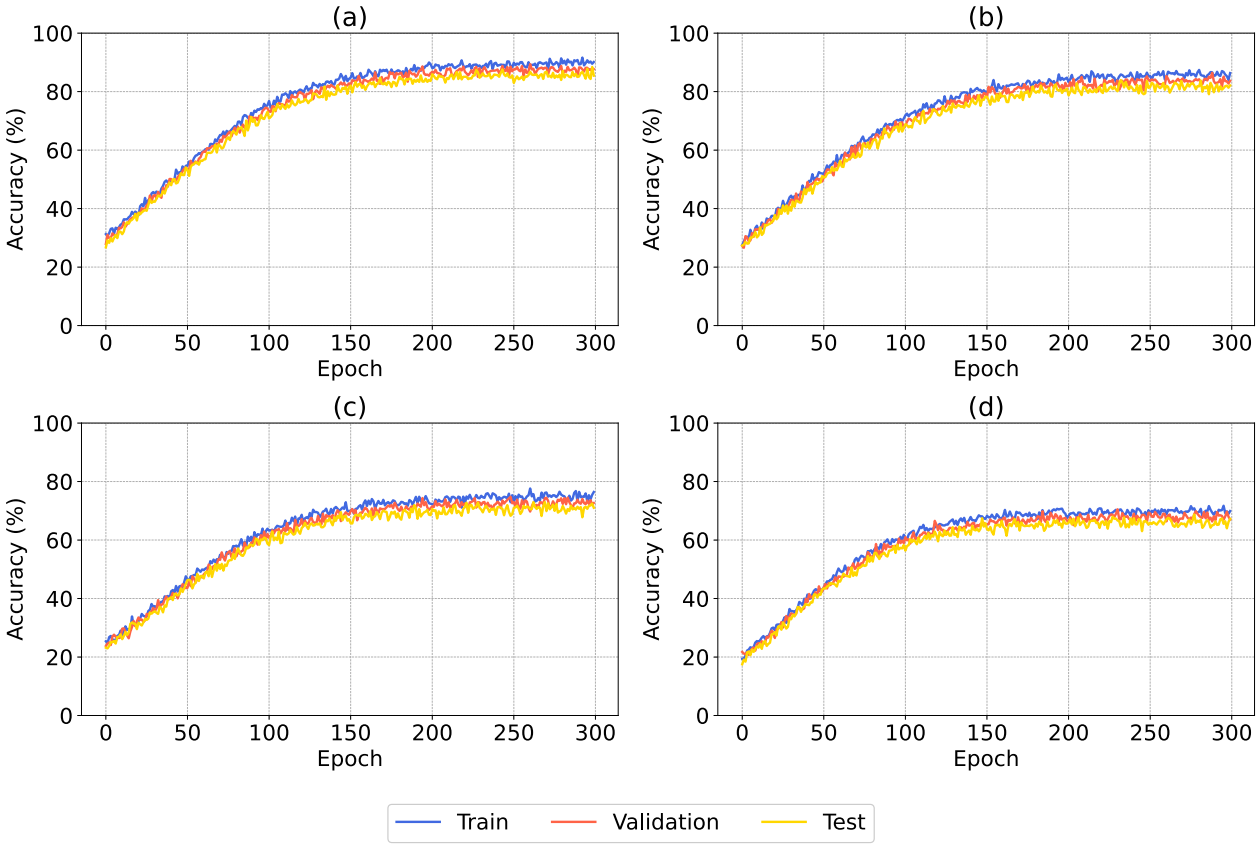

FIGURE 6. Accuracy progression over 300 epochs comparing (a)AFM_YOLO-ResNet with (b)YOLOv8 [35], (c)Darknet [36], and (d)Googlenet [37]. The performance curves highlight the superior accuracy and training efficiency of the AFM_YOLO-ResNet model across the training duration.

are only 75%, 73% and 70%, respectively. This implies the proposed AFM_YOLO-ResNet model requires a dataset close to the size of the full model development set (320 images) for properly fine-tuning: the training, validation, and testing accuracies are all above 85%.

D. CLASSIFICATION RESULT COMPARISON

To further validate the proposed AFM_YOLO-ResNet model, we compared its training and testing performance with that of existing DL image analysis models, such as YOLOv8, Darknet, and Googlenet. The same full dataset was used for training and testing these models. Note that, since the dataset used is relatively small, early stopping was applied during the training process to prevent overfitting, which results in fewer learning iterations. With this consideration, the maximum epoch number used was 300. The performance of all the models is presented in Figure 6.

Despite YOLOv8 being the latest iteration in its series and achieving the highest accuracy among reported techniques, it presented notable challenges. Notably, it demanded significantly longer training time (with the same dataset): 8h 51min for YOLOv8 vs. 2h 43min for the proposed model. This may be caused by the larger size of YOLOv8 model vs YOLOv5, which makes training slower. YOLOv8 also failed to deliver a similar accuracy: YOLOv8 achieved a final training accuracy of 86.34%, testing accuracy of 82.13%, and validation accuracy of 83.71%, compared to 90.13%, 85.45%, and 87.68%, respectively of AFM_YOLO-ResNet model. One reason for the low performance of YOLOv8 is the limited dataset: the limited data in terms of size and versatility (defects only from two samples) is not enough to train the heavy YOLOv8 model. Also, because the task mostly focuses on small object (e.g., defect) detection, while YOLOv8 is known to have limitations on small object detection as it's primarily designed for detecting features across broad scale ranges [38], bigger detection and classification errors have been observed. Moreover, an enhanced classifier (ResNet-34) integrated into the proposed model may also contribute to superior performance, while YOLOv8 doesn't have one. Therefore, the comparison highlighted the superiority of the proposed approach over YOLOv8 for AFM image detection with limited data.

For Darknet, the accuracy achieved was significantly lower: 76.43% for training, 71.03% for testing, and 72.65% for validation. As a popular object detection DL framework, Darknet is commonly used as the backbone of YOLO for feature detection and extraction [39]. Limited by its capability in object classification, it is expected that the performance

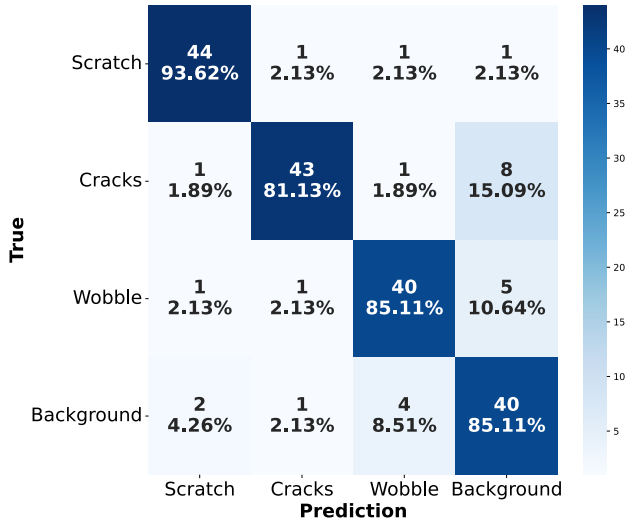

FIGURE 7. Confusion Matrix: True labels are on the vertical axis, while predicted labels are on the horizontal axis. The absolute numbers represent the number of images predicted with each type of defects or background. The percentage fractions were calculated along each row.

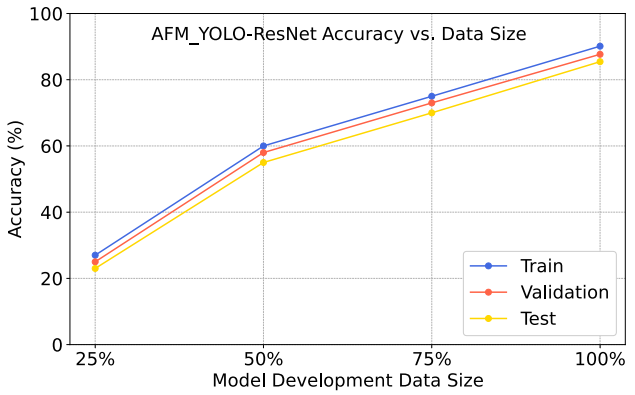

FIGURE 8. AFM_YOLO-ResNet performance vs. model development (training and validation) dataset size. The percentage data size is with respect to that of the full model development subset (160 images each sample). The testing accuracies are for the same aforementioned full testing subset (40 images per sample).

of Darknet for this task cannot compete with that of the proposed approach, which consists of a powerful object detector (YOLOv5) and an enhanced classifier (ResNet-34).

Similarly, the performance was also less satisfying from Googlenet, with a training accuracy of 69.90%, testing accuracy of 67.22%, and validation accuracy of 67.00%. As a generalized transfer learning framework for image classification, Googlenet is known for high computation efficiency but lacks the identification of small objects present in the image and also the localization of the area where the object is present in the image [40]. Compared to the proposed approach, Googlenet falls short of the small object detection capability. Therefore, the accuracies delivered by Googlenet were significantly lower.

Even though these three techniques have been broadly used for feature detection and image classification tasks, the results clearly demonstrated that these generalized models are less optimal for the detection and classification of AFM image defects when compared with the proposed AFM_YOLO-ResNet framework. Therefore, it is indeed useful to develop the proposed hybrid architecture for AFM image assessment.

Although a limited dataset was used for demonstration, this framework can allow users to expand the database for training with newly classified AFM images, thus enhancing precision and adaptability. Such a flexible feature enables the framework to assimilate and learn from an ever-expanding array of sample images, reinforcing its capability to recognize and classify a variety of image defects for various sample types.

With the high accuracy provided by the proposed AFM_YOLO-ResNet for AFM image defect detection and classification, there are several promising directions for future enhancement and application. Future work involves model optimization and architectural enhancements, as well as data expansion. Leveraging newly developed deep learning architectures in combination with transfer learning could lead to better performance and efficiency. Meanwhile, it will be helpful in enhancing the model's real-time processing capabilities to fit the needs for applications requiring immediate responses, such as AFM measurement for dynamic sample evolution monitoring. For data expansion, expanding the training dataset include a more diverse range of defect types and image conditions would improve the model's robustness and generalizability. We will also explore other data augmentation techniques, such as exploring more sophisticated data augmentation techniques like synthetic image generation [41]. These future directions highlight the potential advantages and broader applicability of the AFM_YOLO-ResNet framework, paving the way for innovative applications and further advancements in defect detection and classification.

IV. CONCLUSION

In this paper, a deep learning framework, AFM_YOLO-ResNet, for defect detection and classification of Atomic Force Microscopy topography images was developed. The proposed framework combines the capabilities of YOLOv5 for accurate defect detection with those of ResNet-34 for precise classification. To overcome the limitations of a limited dataset, strategic data augmentation and transfer learning techniques have been integrated. The major significance of the proposed approach are three folds:

- Demonstrated by the model test results, the AFM_YOLO-ResNet framework successfully integrates YOLOv5 and ResNet-34 together with strategic data augmentation and transfer learning to achieve high accuracy in defect detection and classification of AFM images.
- Comparison with three existing image analysis models demonstrated the superior performance of AFM_YOLO-ResNet in terms of efficiency and accuracy.

- The adaptable design of the framework allows for its application in broader image-based quality assessment domains, including manufacturing, material engineering, and biomedical engineering.

These significants highlight the effectiveness and versatility of the AFM_YOLO-ResNet framework in improving defect detection and classification processes across various fields.

DATA AVAILABILITY STATEMENT

The data and the model are available upon request via this GitHub link: https://github.com/Monsterfafa/AFMYOLO-RESNET-project.git

ACKNOWLEDGMENT

The AFM software support from Bruker Nano Inc., is greatly acknowledged.

REFERENCES

[1] M. Aliofkhazraei and N. Ali, *Afm Applications in Micro/Nanostructured Coatings*. Amsterdam, The Netherlands: Elsevier, 2014.
[2] C. Wright, L. Powell, D. Johnson, and N. Hilal, "Microscopy: Atomic force microscopy," in *Encyclopedia of Food Microbiology*. Amsterdam, The Netherlands: Elsevier, 2014, pp. 666–675.
[3] K. Mollaeian, Y. Liu, S. Bi, and J. Ren, "Atomic force microscopy study revealed velocity-dependence and nonlinearity of nanoscale poroelasticity of eukaryotic cells," *J. Mech. Behav. Biomed. Mater.*, vol. 78, pp. 65–73, Feb. 2018.
[4] Z. Zhang, L. Ma, Y. Liu, J. Ren, and H. Hu, "An experimental study of rain erosion effects on a hydro-/ice-phobic coating pertinent to unmanned-arial-system (UAS) inflight icing mitigation," *Cold Regions Sci. Technol.*, vol. 181, Jan. 2021, Art. no. 103196.
[5] J. Rade, J. Zhang, S. Sarkar, A. Krishnamurthy, J. Ren, and A. Sarkar, "Deep learning for live cell shape detection and automated AFM navigation," *Bioengineering*, vol. 9, no. 10, p. 522, Oct. 2022.
[6] Y. F. Dufrene, "Atomic force microscopy, a powerful tool in microbiology," *J. Bacteriology*, vol. 184, no. 19, pp. 5205–5213, Oct. 2002.
[7] A. Tsyrenova, M. Q. Farooq, S. M. Anthony, K. Mollaeian, Y. Li, F. Liu, K. Miller, J. Ren, J. L. Anderson, and S. Jiang, "Unique orientation of the solid–solid interface at the Janus particle boundary induced by ionic liquids," *J. Phys. Chem. Lett.*, vol. 11, no. 22, pp. 9834–9841, 2020.
[8] H. Bai and S. Wu, "Deep-learning-based nanowire detection in AFM images for automated nanomanipulation," *Nanotechnol. Precis. Eng.*, vol. 4, no. 1, Mar. 2021, Art. no. 013002.
[9] P. Müller, S. Abuhattum, S. Möllmert, E. Ulbricht, A. V. Taubenberger, and J. Guck, "Nanite: Using machine learning to assess the quality of atomic force microscopy-enabled nano-indentation data," *BMC Bioinf.*, vol. 20, no. 1, pp. 1–9, Dec. 2019.
[10] B. Huang, Z. Li, and J. Li, "An artificial intelligence atomic force microscope enabled by machine learning," *Nanoscale*, vol. 10, no. 45, pp. 21320–21326, 2018.
[11] A. Krull, P. Hirsch, C. Rother, A. Schiffrin, and C. Krull, "Artificial-intelligence-driven scanning probe microscopy," *Commun. Phys.*, vol. 3, no. 1, p. 54, Mar. 2020.
[12] B. Alldritt, P. Hapala, N. Oinonen, F. Urtev, O. Krejci, F. Federici Canova, J. Kannala, F. Schulz, P. Liljeroth, and A. S. Foster, "Automated structure discovery in atomic force microscopy," *Sci. Adv.*, vol. 6, no. 9, p. 6913, Feb. 2020.
[13] W. Wu, H. Liu, L. Li, Y. Long, X. Wang, Z. Wang, J. Li, and Y. Chang, "Application of local fully convolutional neural network combined with YOLOv5 algorithm in small target detection of remote sensing image," *PLoS One*, vol. 16, no. 10, Oct. 2021, Art. no. e0259283.
[14] M. Gao, D. Qi, H. Mu, and J. Chen, "A transfer residual neural network based on ResNet-34 for detection of wood knot defects," *Forests*, vol. 12, no. 2, p. 212, Feb. 2021.
[15] G. Dai, L. Hu, J. Fan, S. Yan, and R. Li, "A deep learning-based object detection scheme by improving YOLOv5 for sprouted potatoes datasets," *IEEE Access*, vol. 10, pp. 85416–85428, 2022.
[16] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.
[17] H. Liu, F. Sun, J. Gu, and L. Deng, "SF-YOLOv5: A lightweight small object detection algorithm based on improved feature fusion mode," *Sensors*, vol. 22, no. 15, p. 5817, Aug. 2022.
[18] D. Wu, S. Lv, M. Jiang, and H. Song, "Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments," *Comput. Electron. Agricult.*, vol. 178, Nov. 2020, Art. no. 105742.
[19] G. Wen, S. Li, F. Liu, X. Luo, M.-J. Er, M. Mahmud, and T. Wu, "YOLOv5s-CA: A modified YOLOv5s network with coordinate attention for underwater target detection," *Sensors*, vol. 23, no. 7, p. 3367, Mar. 2023.
[20] 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)*, Jul. 2017, pp. 936–944.
[21] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8759–8768.
[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[23] K. He, X. Zhang, and S. Ren, "Deep residual learning," *Image Recognit.*, vol. 7, pp. 1–29, Jun. 2015.
[24] N. F. Greenwald et al., "Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning," *Nature Biotechnol.*, vol. 40, no. 4, pp. 555–565, Apr. 2022.
[25] N. Tajbakhsh, L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, and X. Ding, "Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation," *Med. Image Anal.*, vol. 63, Jul. 2020, Art. no. 101693.
[26] J. Cartucho, R. Ventura, and M. Veloso, "Robust object recognition through symbiotic deep learning in mobile robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 2336–2341.
[27] B. Zoph, E. D. Cubuk, G. Ghiasi, T. Y. Lin, J. Shlens, and Q. V. Le, "Learning data augmentation strategies for object detection," in *Proc. Eur. Conf. Comput. Vis.*, Aug. 2020, pp. 566–583.
[28] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021.
[29] T. Wolf, V. Sanh, J. Chaumond, and C. Delangue, "TransferTransfo: A transfer learning approach for neural network based conversational agents," 2019, *arXiv:1901.08149*.
[30] T. Lin, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.
[31] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, Y. Kwon, K. Michael, J. Fang, Z. Yifu, C. Wong, and D. Montes, "Ultralytics/YOLOv5: V7.0-YOLOv5 sota realtime instance segmentation," *Zenodo*, vol. 1, no. 2, pp. 1–19, 2022.
[32] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, and L. Antiga, "Pytorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–29.
[33] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[34] A. Rakhlin, O. Shamir, and K. Sridharan, "Making gradient descent optimal for strongly convex stochastic optimization," 2011, *arXiv:1109.5647*.
[35] J. Terven and D. Cordova-Esparza, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," 2023, *arXiv:2304.00501*.
[36] P. Biddle, P. England, M. Peinado, and B. Willman, "The darknet and the future of content distribution," in *Proc. ACM Workshop Digit. Rights Manage.*, vol. 6, 2002, p. 54.
[37] R. U. Khan, X. Zhang, and R. Kumar, "Analysis of ResNet and GoogleNet models for malware detection," *J. Comput. Virol. Hacking Techn.*, vol. 15, no. 1, pp. 29–37, Mar. 2019.
[38] G. Wang, Y. Chen, P. An, H. Hong, J. Hu, and T. Huang, "UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios," *Sensors*, vol. 23, no. 16, p. 7190, Aug. 2023.

[39] O. Elharrouss, Y. Akbari, N. Almaadeed, and S. Al-Maadeed, "Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches," 2022, *arXiv:2206.08016*.
[40] P. Tang, H. Wang, and S. Kwong, "G-MS2F: GoogLeNet based multi-stage feature fusion of deep CNN for scene recognition," *Neurocomputing*, vol. 225, pp. 188–197, Feb. 2017.
[41] R. Corvi, D. Cozzolino, G. Zingarini, G. Poggi, K. Nagano, and L. Verdoliva, "On the detection of synthetic images generated by diffusion models," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5.

JUNTAO ZHANG received the bachelor's degree in mechanical engineering from Iowa State University (ISU), Ames, IA, USA, in 2020, where he is currently pursuing the Ph.D. degree. His research interests include modeling, artificial intelligence, and deep learning.

JUAN REN received the B.S. degree in mechanical engineering from Xi'an Jiaotong University, China, in 2009, and the Ph.D. degree in mechanical engineering from Rutgers, The State University of New Jersey, in June 2015. She is currently an Assistant Professor with the Department of Mechanical Engineering, Iowa State University (ISU), where she has been a Faculty Member, since August 2015. She holds the William and Virginia Binger Professorship with the Department of Mechanical Engineering, ISU. Her research interests include learning-based output tracking and control, control tools for high-speed scanning probe microscope imaging, mechanotransduction modeling, and nanomechanical measurement and mapping of soft and live biological materials. She received the NSF CAREER Award, in 2018. She is the Representative of the IEEE Control Systems Society in the IEEE Nanotechnology Council and an Associate Editor of *Mechatronics* (Elsevier).

SHUIQING HU received the Bachelor of Electrical Engineering and Master of Science degrees from Nanjing University, China, in 1998 and 2001, respectively, and the Ph.D. degree in mechanical engineering from Purdue University, in 2007. Since then, he has been a Research Scientist with Bruker Nano Surfaces and Metrology. His research interests include developing new technologies in scanning probe microscopy, including peak force tapping mode, mask cleaning technology, fast scanning, and AFM nano IR.