Automatic Urine Sediment Detection and Classification Based on YoloV8

Sania Akhtar[1], Muhammad Hanif[1,2(✉)], and Hamidi Malih[1,2]

[1] Faculty of Computer Science and Engineering (FCSE), Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Swabi, Pakistan
sania.akhtar@giki.edu.pk

[2] Department of Electrical and Electronics, Kutahya Dumplupinar University, Kutahya, Turkey
muhammad.hanif@giki.edu.pk, melih.saraoglu@dpu.edu.tr

Abstract. The identification of urine sediment in human urine samples through microscopic images is a critical part of in vitro testing. Currently, automatic urine sediment analyzers are used by doctors to supplement manual examinations. However, the conventional technique of artificial feature extraction used by most analyzers can be labor-intensive and subjectively dependent on the professional's prior knowledge. To overcome these limitations, this work employs YoloV8, a recent version of the Yolo algorithm, to accurately detect and categorize urine particles. In addition, a data-centric strategy has been introduced to address difficulties with missing data, incorrect labeling, and class imbalance. This strategy aims to improve labeling reliability and remove noisy data points. Experimental findings on the dataset show that YOLOv8 has a greater detection accuracy than existing state-of-the-art techniques for detecting eleven different categories of urine sediments. The approach presented in this work outperforms other techniques, yielding a mean average precision (mAP) of 91%. Furthermore, the average detection time of the model is 0.6 microseconds.

Keywords: Urine Sediments · YoloV8 · object detection · Feature Extraction

1 Introduction

Urine sediment analysis is an essential medical diagnostic method used to detect various physical diseases. This low-cost method examines the active components present in urine. It involves qualitative detection of urine sediment cells and quantitative analysis of cell types, numbers, and shapes using microscopy. By studying the characteristics of urine sediment, doctors can determine whether the body is suffering from certain diseases [1,2]. For instance, even though urine with increased erythrocytes appears colorless to the naked eye, hematuria can be observed under a microscope [1]. Overall, urine sediment analysis plays a critical role in identifying potential health problems. When examining urine sediment,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023 Workshops, LNCS 14112, pp. 269–279, 2023.
https://doi.org/10.1007/978-3-031-37129-5_22

elevated levels of leukocytes are often indicative of urinary system infections. If abnormal-shaped erythrocytes are observed under the microscope, the patient may have kidney disease. Additionally, the presence of epithelial cells in urine microscopy suggests that the patient's renal parenchyma has been damaged and requires treatment [1,2]. In short, the results of urine sediment analysis can provide valuable information to identify the underlying health issues affecting the urinary system and kidneys.

Urine microscopic images consist of various tiny and unevenly distributed particles, which makes it challenging to analyze. Therefore, the examination of urine sediment requires skilled laboratory physicians. However, this process is time-consuming and can be affected by the subjective will and working fatigue of the laboratory physicians due to the significant workload. Fortunately, with the advancement of digital image technology, automatic image recognition has become prevalent in the medical field. The conventional approach to image recognition, which involves "target segmentation + feature selection and extraction + classifier," has shown promising results in identifying and categorizing urine sediment images [1–3]

The field of computer vision has made significant progress with the emergence of convolutional neural networks through deep learning. This development has led to the creation of many outstanding object detection algorithms. Currently, object detection frameworks are typically divided into two categories: those based on regional recommendation and those based on dense sampling. In the former approach, the detector initially generates a rough regional proposal for the input image. Afterward, it more accurately classifies and locates the region of interest (ROI) produced [1,3]. Object detectors such as R-CNN [4], Fast R-CNN [5], Faster R-CNN [6], R-FCN [7], and FPN based on the two-stage detection framework have demonstrated exceptional performance. They operate in two stages, starting with the regional recommendation and then performing accurate detection. Due to this approach, they are commonly known as two-stage detectors [7]. In contrast, object detectors based on deep convolutional networks can use dense sampling [8] to directly detect objects by performing position regression and class prediction on all possible grid points on the detected feature layer. Examples of such one-stage detectors are SSD [9] and YOLO [8,10,11], which offer faster detection speed but comparatively lower accuracy. However, many outstanding algorithms and network structure enhancements have been introduced in recent years to improve the accuracy of one-stage detectors, achieving remarkable results.

2 Methodology

This paper describes the use of YOLOv8, an object detection and classification model, for identifying and categorizing different types of cells in urine sediment images. The dataset was labeled using Roboflow and included eleven classes of urine particles such as platelets, red blood cells (RBC), white blood cells (WBC), casts, crystals, epithelial cells (epith), transitional epithelial cells (epithn), erythrocytes (eryth), leukocytes (leuko), fungi (mycete), and background. Deep

learning models were applied to the dataset to analyze the sample images and determine the number and types of cells present in the urine sediment.

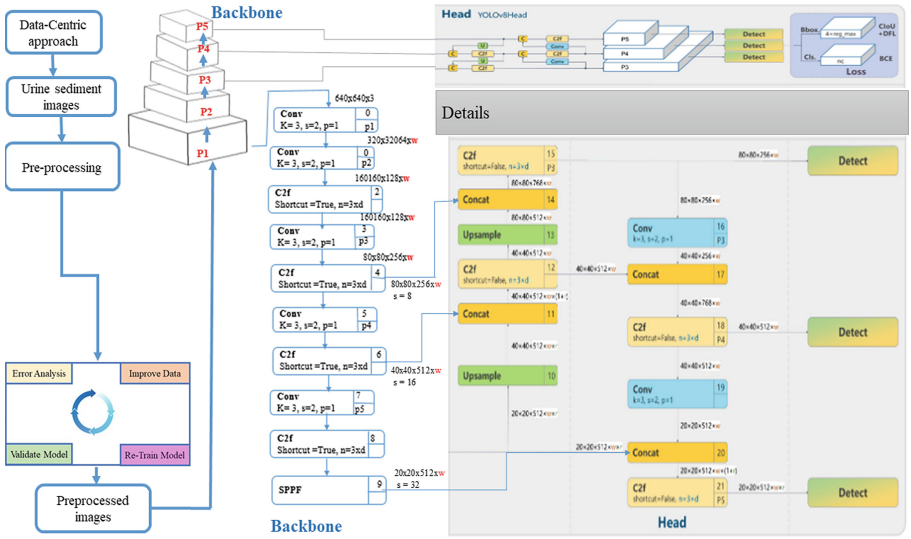

Fig. 1. Workflow of the proposed method.

2.1 Data-Centric Approach

Before training a deep learning model on any dataset, it is crucial to ensure that the data is of high quality and suitable for the intended purpose. In this context, the primary goal is to accurately identify and categorize different types of cells in urine sediment images. Therefore, the first step is to examine the data to detect any missing data, inaccurate labeling, or class imbalance issues.

Missing data refers to any data points that are not available or have not been recorded. Inaccurate labeling refers to any mislabeling of data points or incorrect assignment of class labels. Class imbalance issues occur when there is a significant disparity in the number of data points across different classes, which can lead to bias in the model's training. To address these issues, the data is carefully examined to identify any missing data points or incorrect labels. Any such data points are either removed or corrected to ensure that the dataset is of high quality and reliable. Additionally, the data is checked for any class imbalance issues, and appropriate measures are taken to balance the number of data points across different classes, such as oversampling or undersampling.

By ensuring the data is accurate, reliable, and balanced, the labeling reliability is boosted, and noisy data points are excluded. This process helps to ensure that the deep learning model trained on the dataset can accurately identify and categorize different types of cells in urine sediment images, leading to more accurate results and better insights.

2.2 Working of Yolo

YOLO (You Only Look Once) is an object detection algorithm that performs object detection in a single pass by dividing the input image into a grid and predicting bounding boxes, confidence scores, and class probabilities directly. It utilizes convolutional neural networks (CNNs) to extract features and employs anchor boxes to handle objects of different sizes and aspect ratios. During training, YOLO optimizes its parameters using a loss function that considers localization accuracy and classification performance. In inference, the algorithm processes the image through the network, generates predictions, and applies non-maximum suppression to refine the results.

2.3 Architectural Details of YoloV8

We trained the YOLOv8 model, which is developed by Ultralytics, a latest state-of-the-art YOLO model capable of performing tasks such as object detection, image classification, and instance segmentation. It includes several improvements over the influential YOLOv5 model, including changes in the backbone architecture and a new training methodology that enables faster and more efficient training.

2.3.1 Anchor Free Detection

Unlike models that rely on predefined anchor boxes to predict object location, YOLOv8 is an anchor-free model that directly predicts the center of an object. This approach eliminates the need for anchor boxes and simplifies the object detection process, leading to more accurate predictions. Earlier YOLO models had difficulties with anchor boxes as they might not represent the object box distribution in custom datasets, leading to inaccurate predictions. Anchor-free detection simplifies the object detection process by reducing the number of box predictions, resulting in faster Non-Maximum Suppression (NMS) during post-processing. NMS is a crucial step that filters out false positives and selects the most accurate detections after inference.

2.3.2 New Convolutions

The YOLOv8 model includes several architectural changes from earlier versions, such as replacing the first 6×6 convolution in the stem with a 3×3 convolution, using a new building block, and replacing C3 with C2f. The C2f module concatenates all the outputs from the Bottleneck, while in C3 only the output of

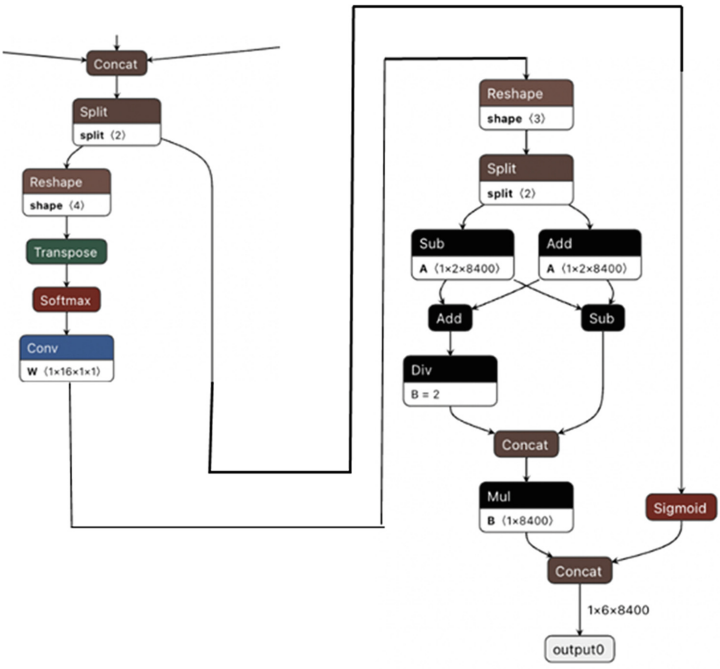

Fig. 2. The detection head for YOLOv8, visualized in netron.app.

the last Bottleneck was used. The module is composed of a Convolution layer, a Batch Normalization layer, and a SiLU activation layer, and the number of features is denoted as "f" while the expansion rate is denoted as "e". A diagram summarizing the module is provided in Fig. 2. The Bottleneck used in YOLOv8 is similar to that of YOLOv5, but with a change in the kernel size of the first convolution from 1×1 to 3×3. This change indicates a shift towards the ResNet block architecture, which was introduced in 2015. In the neck, feature concatenation is done without enforcing the same channel dimensions, which reduces the parameter count and the tensor size.

3 Experiments

3.1 Dataset

The accuracy of network recognition is greatly affected by the quality of datasets. In order to obtain a high-quality dataset for image recognition, several urine samples were collected using micro-photography equipment. From each sample, 650 high-definition images were captured, resulting in a total of 5632 images belonging to 10 different classes and 1 background, as shown in Fig. 4. The number of images per class is 284. The images were manually labeled using Robowflow and verified by the professional medical team and finally, 10 categories of urine

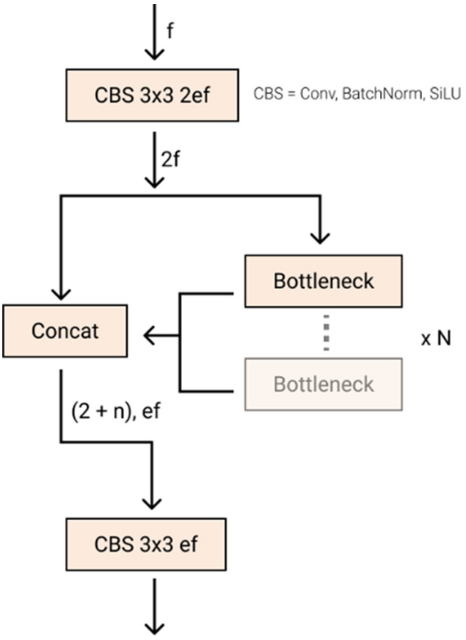

Fig. 3. The new YoloV8 c2f module.

sediment with medical significance were selected. These categories include RBC, WBC, casts, crystals, epithelial cells (epith), transitional epithelial cells (epithn), erythrocytes (eryth), leukocytes (leuko), and mycete.The dataset was divided into three sets - a training set, a validation set, and a test set - with random assignment of images. The training set consisted of 90% of the images, while the validation and testing set each contained 5% of the images.

3.2 Performance Metrics

The study assesses the performance of the model through the utilization of evaluation metrics such as precision, recall, F1 score, confidence interval, and micro-averaging.
Precision and recall are defined as.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

where TP is True Positives and FP is False Positives

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

F 1 score is defined as.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \tag{3}$$

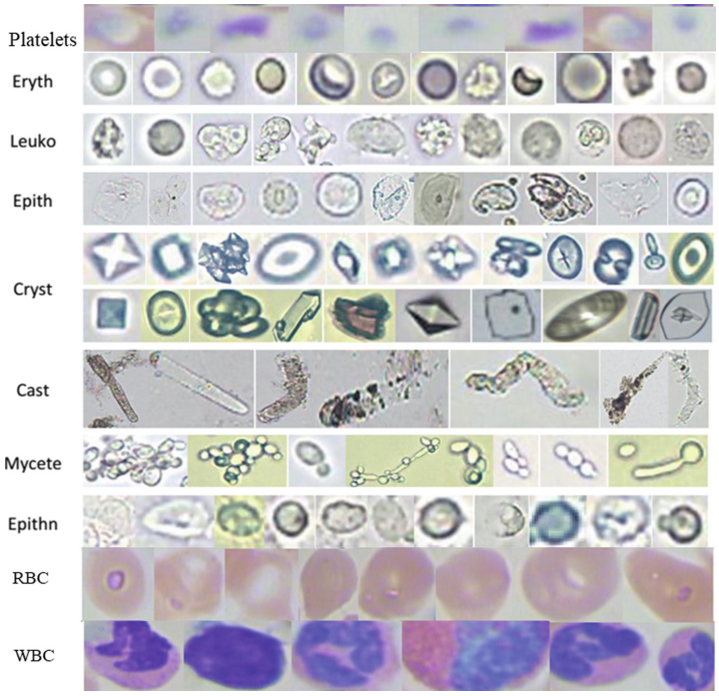

Fig. 4. The sample dataset for urine sediment is presented in the left column with their labels on the right

The F1 score is calculated by taking the harmonic mean of precision and recall. This is an appropriate choice because both precision and recall are rates. By taking the harmonic mean, the F1 score gives equal importance to both precision and recall in its calculation. Therefore, it can be said that the F1 score provides an average of precision and recall with equal weighting. Whereas evaluating multi-category recognition models, micro-averaging is a core indicator. It provides useful insights into the model's performance in multiple categories. To calculate the micro average for the given data, we sum up the number of true positive predictions for all the classes and divide it by the total number of actual positives, instead of predicted positives which are represented by the following equation.

$$PrecisionMicroAvg = \frac{TP1 + TP2 + ... + TPn}{TP1 + TP2 + ... + TPn + FP1 + FP2 + FP3 + ... + FPn} \tag{4}$$

The confidence measure is defined as

$$ConfidenceMeasure = \frac{TruePositive}{TruePositive + FalsePositive} \tag{5}$$

The confidence measure is the ratio of true positive predictions to the sum of true positive predictions and false positive predictions. In other words, it represents the proportion of correctly predicted positive instances out of the total instances predicted as positive by the model. It provides an indication of the model's accuracy in correctly identifying positive instances while taking into account the number of false positive predictions made.

3.3 Experiments and Analysis

The main network was trained and tested using an Intel Core i7-7700K CPU and Tesla T4 GPU with driver version NVIDIA-SMI 525.85.12. The learning rate used during training was 0.001, and the training process was repeated for 55 epochs. The test results were summarized in Fig. 7, which shows the confusion matrix of the main network.It's very clear from the figure that eryth and mycete are correctly detected and classified with the highest confidence of 97%. Whereas RBCs are the particles that are 81% correctly classified with the lowest confidence interval as compared to all other 10 particles. F1 score, precision, and recall are important performance metrics for detection and recognition tasks which are presented in Figs. 6 and 7.

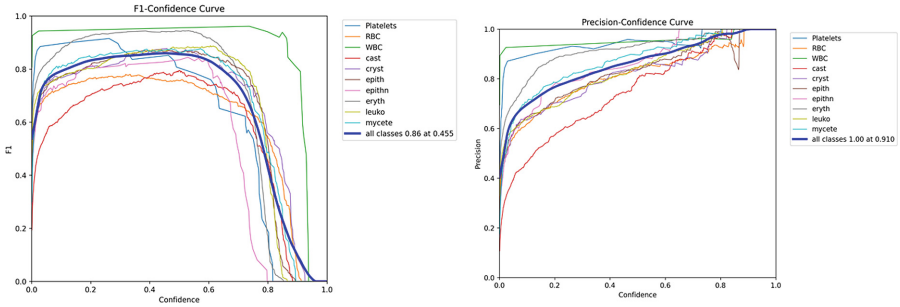

Fig. 5. F1 and Precision- confidence curve for all ten catagories.

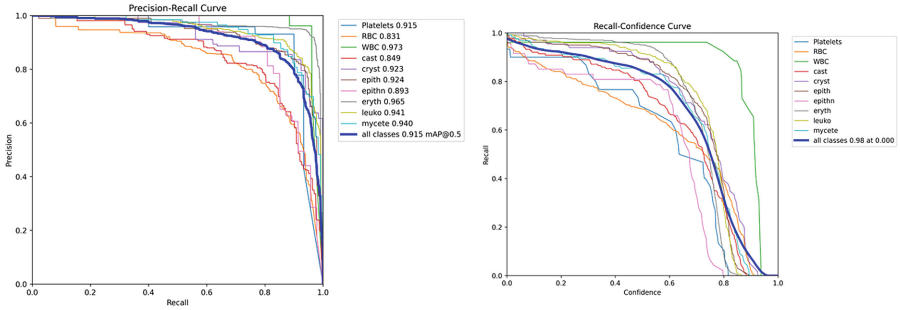

Fig. 6. Precision-Recall and Recall-confidence curve for all ten catagories.

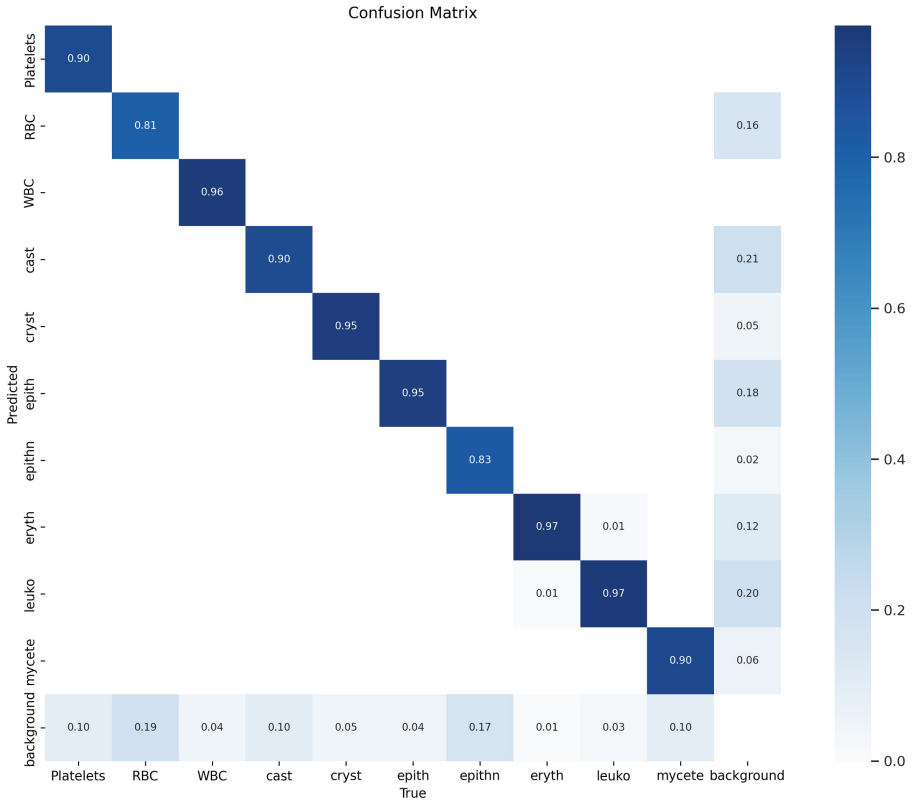

Fig. 7. Detection results on test set of ten catagories of urine sediment.

Table 1. Comparison of accuracy of detection methods from different research articles

N0	Methods	Dataset & categories	mAP	Infernece time
1	Our Method (Yolov8)	10 categories of RBC, WBC, cast, etc.	91%	0.6 ms per image
2 [12]	Faster RCNN+SSD	7 categories of RBC, WBC, Casts, etc.	84.1%	72 ms per image
3 [13]	CNN (DFPN)	7 categories of RBC, WBC, Crystals, etc.	86.9%	Not mentioned
4 [14]	DCNN	2 categories of RBC, WBC, Crystals, etc.	98.94%	2590 ms per image
5 [15]	Improved Yolov3	5 categories of RBC, WBC, Crystals, etc.	90.1%	0.047 s per fram

Figure 8 depicts the performance results obtained from testing a model on a dataset consisting of 10 distinct categories. The average mean precision accuracy achieved by the model was 91%, indicating that the model is capable of accurately classifying objects within the dataset. Furthermore, the inference time, which is the time it takes for the model to make a prediction on a single sample of data, was 0.6 ms, implying that the model is efficient and can make predictions quickly. Overall, the model appears to be reliable and efficient in performing its intended task of object classification.

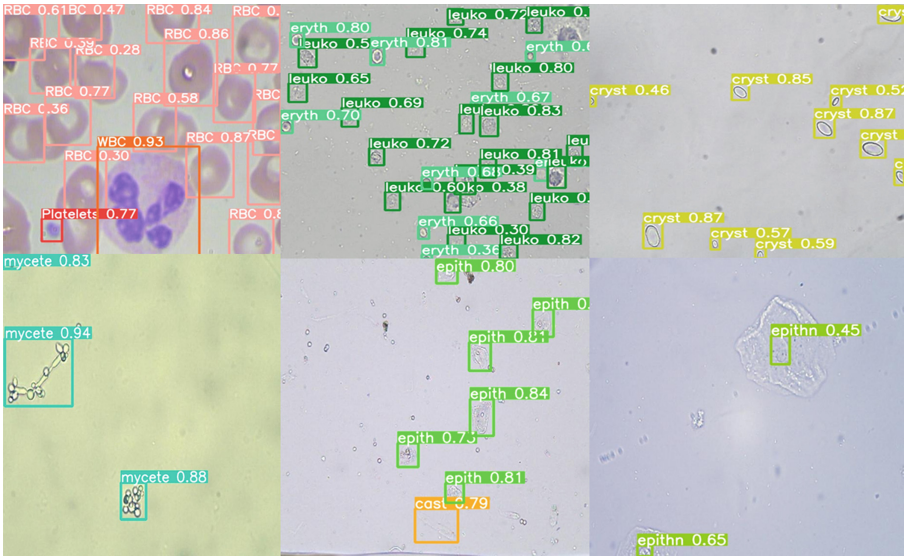

Fig. 8. Prediction results of YoloV8

The experimental results demonstrate that YOLOv8 is the fastest in terms of detection speed. YOLOv8 outperform the Faster R-CNN with VGG16 net in terms of detection speed as they are one-stage algorithms, while Faster R-CNN is a two-stage algorithm. Two-stage algorithms require a rough regional proposal for the input image, followed by more accurate classification and location estimation, which takes more time for detection. Although one-stage algorithms have faster detection speed, they are generally associated with lower accuracy. However, YOLOv8, despite being a one-stage algorithm, achieves high detection accuracy while maintaining fast speed. To further validate the effectiveness of the proposed method and other algorithms, we compare the detection of average class accuracy, which is presented in Table 1. Our method involves increasing the number of parameters and computations for determining the training samples. However, these modifications do not significantly impact the detection speed. Despite the requirement for fast detection, our model achieves a mean average precision (mAP) of 91%, which is superior to other algorithms. The detection time required for our model is 0.6 ms per image, which is slightly longer than improved YOLOv3's. Considering the real-time and accuracy requirements of automatic urinary sediment visible component detection, our model is a viable option.

4 Conclusion

This paper proposes a deep learning-based method for urine sediment image recognition, which is significant for the diagnosis of related diseases in urology and other departments. The traditional recognition methods have limitations,

such as heavy workload and low accuracy. The proposed method uses a data-centric approach to enhance labeling reliability and overcome class imbalance and missing labeling data. The method employs YoloV8, a recent version of the Yolo algorithm, to accurately detect and categorize urine particles. The CNNs are trained on a dataset containing 5632 images, and the test set achieves an average mean precision of 91% with an average recognition time of 0.6 ms. Furthermore, the method has strong expandability and can introduce more categories for recognition in practical applications.

References

1. Goswami, D., Aggrawal, H., Agarwal. V.: Cell detection and classification from urine sediment microscopic images (2020)
2. Aglibot, K.P., Angeles, J.A., Gecana, J.F., Germano, A.B., Macalindong, J.A., Tolentino, R.E.: Urine crystal classification using convolutional neural networks. In: 2022 International Visualization, Informatics and Technology Conference (IVIT), pp. 245–250. IEEE (2022)
3. Ji, Q., Li, X., Zhiyu, Q., Dai, C.: Research on urine sediment images recognition based on deep learning. IEEE Access **7**, 166711–166720 (2019)
4. Uijlings, J.R.R., Van De Sande, K.E.A., Gevers, T., Smeulders, A.W.M.: Selective search for object recognition. Int. J. Comput. Vis. **104**, 154–171 (2013)
5. Girshick. R.: Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448 (2015)
6. Wang, Q., Bi, S., Sun, M., Wang, Y., Wang, D., Yang, S.: Deep learning approach to peripheral leukocyte recognition. PLoS ONE **14**(6), e0218808 (2019)
7. Redmon, I., Farhadi, A.: Yolov3: an incremental improvement. preprint. arXiv preprint arXiv:1804.02767, 4322 (2018)
8. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)
9. Redmon, J., Farhadi, A.: Yolo9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7263–7271 (2017)
10. Li, Q., et al.: Inspection of visible components in urine based on deep learning. Med. Phys. **47**(7), 2937–2949 (2020)
11. Lin, T-y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988 (2017)
12. Liang, Y., Kang, R., Lian, C., Mao, Y.: An end-to-end system for automatic urinary particle recognition with convolutional neural network. J. Med. Syst. **42**, 1–14 (2018)
13. Liang, Y., Tang, Z., Yan, M., Liu, J.: Object detection based on deep learning for urine sediment examination. Biocybernet. Biomed. Eng. **38**(3), 661–670 (2018)
14. Wang, Q., Sun, Q., Wang. Y.: A two-stage urine sediment detection method. In: 2020 International Conference on Image, Video Processing and Artificial Intelligence, vol. 11584, pp. 15–21. SPIE (2020)
15. Dong, S., Zhang, S., Jiao, L., Wang. Q.: Automatic urinary sediments visible component detection based on improved yolo algorithm. In: 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), pp. 485–490. IEEE (2020)