

Object Detection as Campylobacter Bacteria and Phagocytotic Activity of Leukocytes in Gram Stained Smears Images

Kyohei Yoshihara and Kouich Hirata

Kyushu Institute of Technology, Kawazu 680-4, Iizuka 820-8502, Japan

Keywords: Object Detection, Campylobacter Bacteria, Phagocytotic Activity of Leukocytes, Gram Stained Smears Images, Faster R-CNN, RetinaNet, YOLOv5.

Abstract: In this paper, we apply object detection to Gram stained smear images, where objects are Campylobacter bacteria and phagocytotic activity of leukocytes. Then, we adopt three CNN-based object detectors of Faster R-CNN, RetinaNet and YOLOv5. The outline of the detection is first to annotate the regions of objects as Campylobacter bacteria and phagocytotic activity of leukocytes in training images, and then to detect the regions of objects in the remained test images by using the detectors. Finally, we give experimental results of detecting Campylobacter bacteria and phagocytotic activity of leukocytes in Gram stained smear images by using the detectors.

1 INTRODUCTION

The *Gram stain* (Bartholomew and Mittwer, 1952) is the method for microbial smears test in microscope test per $1,000\times$ field, introduced by Hans Christian Gram (1853–1938) at 1884. For the Gram stain, based on the stained colors as purple/violet or red/pink, the stained shapes as sphere-shape, rod-shape, singles, pairs, chains, clusters, and so on, we detect bacteria occurring in the smears for the samples of blood, sputum, feces, pus and urine.

After Gram staining, we call the bacteria colored by purple or violet *Gram positive* and those by red or pink *Gram negative*. Also we call the bacteria stained as sphere-shape *cocci* and those as rod-shape *basilli*. Hence, we can classify bacteria into the four kinds as *Gram positive cocci*, *Gram positive bacilli*, *Gram negative cocci* and *Gram negative bacilli*¹.

Since the Gram stain is applicable inexpensively and fast returns the results (within 30 min.), it is important for the initial medical care of infectious diseases. On the other hand, Gram staining is possible to stain not only bacteria but also non-bacteria substances such as leukocytes, dusts, oil and crystals. Also there exist many kinds of phlogogenic bacteria for infectious diseases.

In the microscope test, Gram stained smears images are checked manually and visually but not auto-

matically in general. The reason is that we can detect bacteria exactly by applying the culture test and the identification test after the microscope test. On the other hand, since anaerobic bacteria are never lived in the culture test, they cannot be detected by the identification test. Then, the detected bacteria through the culture test and the identification test are the part of bacteria in smears.

Also since the culture test and the identification test spend one day, we cannot apply them to the initial medical care of infectious diseases. Furthermore, whereas expert skills are necessary to detect bacteria manually and visually from Gram stained smears images, such technicians with expert skills are not enough to apply the initial medical care in Japan. Hence, the automatic detection of bacteria from Gram stained smears images is required.

In this paper, we focus on *Campylobacter bacteria* and *phagocytotic activity of leukocytes*. The *Campylobacter bacteria* are Gram negative basilli and phlogogenic bacteria causing so called *Campylobacter enteritis*. Also the phagocytotic activity of leukocytes works as the natural immunity. Then, the detection of them is an important tasks for the microscope test of the Gram stained smears images. However, it is also well-known to be difficult in the microscope test.

The difficulty of the detection of *Campylobacter bacteria* is that we cannot distinguish the *Campylobacter bacteria* from dusts, since the *Campylobacter bacteria* are as small as dusts and their shapes are also

¹Sometimes we call basilli *rods* (Smith et al., 2018).

similar as dusts. Figure 1 illustrates the image containing Campylobacter bacteria, which occur in just two red circles, and a Campylobacter bacterium in Gram stained smears images.

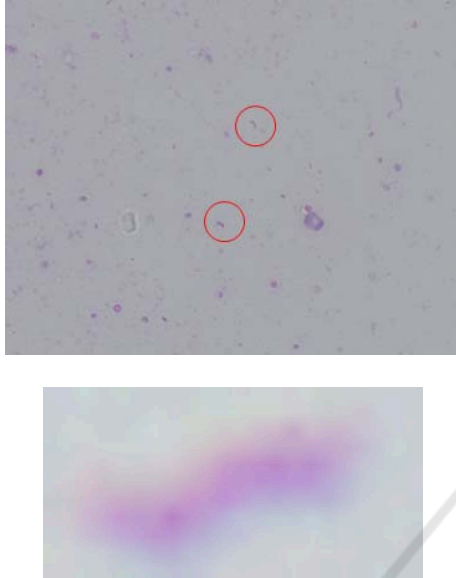
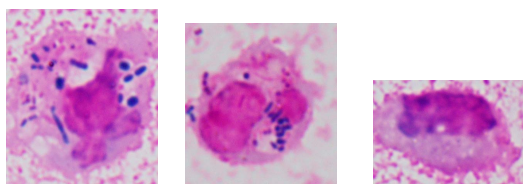


Figure 1: The Gram stained smears image containing Campylobacter bacteria (upper) and a Campylobacter bacterium in Gram stained smears images (lower).

The difficulty of the detection of phagocytotic activity of leukocytes is that not only phagocytotic images that a leukocyte enclose bacteria as avoiding to its nucleus but also *quasi-phagocytotic images* that a leukocyte and bacteria are just overlapping are observed in Gram stained smears images. Figure 2 illustrates a phagocytotic image, a quasi-phagocytotic image and a non-phagocytotic image in Gram stained smears images.



phagocytotic quasi-phagocytotic non-phagocytotic

Figure 2: A phagocytotic image, a quasi-phagocytotic image and a non-phagocytotic image in Gram stained smears images.

Recently, Yoshihara and Hirata (Yoshihara and Hirata, 2021) have classified Campylobacter bacteria and phagocytotic activity of leukocytes, respectively, by using VGG16 (Simonyan and Zisserman, 2015) and its improvement with high accuracy. In their work, for Campylobacter bacteria, after manually extracting the regions in which both a Campy-

lobacter bacterium occurs and no Campylobacter bacterium occurs, they have classified them. Also, for phagocytotic activity of leukocytes, after manually extracting the regions as phagocytotic images, quasi-phagocytotic and non-phagocytotic images, they have classified them. Hence, the purpose of their work is to classify Campylobacter images and phagocytotic images which have been manually extracted, *without object detection*.

On the other hand, in this paper, we detect them *with object detection*. For Campylobacter bacteria, after annotating the regions of Campylobacter bacteria in training images, we detect the regions of Campylobacter bacteria in the remained test images. For phagocytotic activity of leukocytes, after annotating phagocytotic, quasi-phagocytotic and non-phagocytotic regions in training images, we detect the regions of phagocytotic activity of leukocytes in the remained test images.

In this paper, we adopt three CNN-based object detectors of Faster R-CNN (Ren et al., 2015), RetinaNet (Lin et al., 2017) and YOLOv5 (Jocher, 2020). By using these detectors, we give experimental results of detecting Campylobacter bacteria and phagocytotic activity of leukocytes in Gram stained smears images.

1.1 Related Works

As the works dealing with Gram stained smears images, Carvajal *et al.* (Carvajal et al., 2014) have developed the system to learn the candidate areas from fixed-size (51×38 pixels) images applicable to the microscope test with high magnification. Hashimoto *et al.* (Hashimoto et al., 2020) have developed the system to detect Geckler classification defined by the number of buccal squamous epithelial cells and leukocytes for the Gram stained smears images per $100\times$ field for the sample of sputum, in order to guarantee how the Gram stained smears image per $1,000\times$ field is quality for the microscope testing.

Lejon and Andersson (Lejon and Andersson, 2016) have developed the system to classify the bacteria occurring in the areas for the sample of blood by using the template matching. Smith *et al.* (Smith et al., 2018) have classified Gram negative bacilli, Gram positive cossi in clusters and Gram positive cossi in pairs or chains from the Gram stained smear images for the sample of blood by using CNN, after extracting fixed size (146×146 pixels) images. Iida *et al.* (Iida et al., 2020) have developed the system to classify four kinds of Gram positive cocci, Gram positive bacilli, Gram negative cocci and Gram negative bacilli from the Gram stained smear images for not only blood but also other samples by using CNN.

2 OBJECT DETECTION

The purpose of this paper is to detect *Campylobacter* bacteria and phagocytotic activity of leukocytes in Gram stained smears images. The outline of the detection is first to annotate the regions of objects as *Campylobacter* bacteria and phagocytotic activity of leukocytes in training images, and then to detect the regions of objects in the remained test images by using the detectors.

2.1 Detectors

In this paper, we adopt the following three detectors. Here, we implement them through PyTorch, which is an open source machine learning library for Python.

2.1.1 Faster R-CNN

Faster R-CNN (Ren et al., 2015) is a two-stage detector and consists of RPN (region proposal network) and RoI (region of interest) pooling layer. Then, it classifies objects by the RoI pooling layer after proposing regions by the RPN.

To implement Faster R-CNN, we use Detectron2 (Wu et al., 2019) as an object detection library, and tune up the model of FRN+ResNeXt-101-32x8d in PyTorch.

2.1.2 RetinaNet

RetinaNet (Lin et al., 2017) is an one-stage detector and consists of a feature pyramid network backbone on the top of a feedforward ResNet architecture. ResNet consists of subnetworks for classifying anchor boxes and those for regressing from anchor boxes to ground truth object boxes. This network design is intentionally simple, which is a reason why one-stage detectors are faster than two-stage detectors.

To implement RetinaNet, we use Detectron2 (Wu et al., 2019) as an object detection library, and tune up the model of ResNet101 in PyTorch.

2.1.3 YOLOv5

YOLOv5 (Jocher, 2020), where YOLO is an acronym “You only look once” and v5 means “version 5,” is an one-stage detector integrating of the entire object detection and classification process in a single network. The network of YOLO has 24 convolutional layers followed by 2 fully connected layers. Then, YOLO pertains the convolutional layers on the ImageNet classification task at half the resolution and then double the resolution for detection.

To implement YOLOv5, we use a default PyTorch library as an object detection library and tune up the model of YOLOv5x in PyTorch.

2.2 Setting

In this paper, our computer environment is under Google Colab that OS is Ubuntu 18.04.5 LTS, CPU is Intel(R) Xeon(R) CPU @ 2.20GHz, RAM is 25GB and GPU is Tesla P100-PCIE.

Also Table 1 illustrates the values of hyperparameters such as *epoch*, *batch_size* and *lr* (learning rate).

Table 1: The values of hyperparameters that *epoch*, *batch_size* and *lr*.

detector	<i>epoch</i>	<i>batch_size</i>	<i>lr</i>
Faster R-CNN	200	16	0.001
RetinaNet	200	16	0.001
YOLOv5	400	16	0.01

To determine the values of hyperparameters, we tune up manually them that the values of mAP or other AP's are large.

2.3 Annotation

In this paper, we use 67 Gram stained smear images for the sample of feces to detect *Campylobacter* bacteria and 101 Gram stained smear images for the sample of sputum to detect phagocytotic activity of leukocytes.

Then, with helping the opinions of the medical technologist, we annotate the regions of object as *Campylobacter* bacteria and phagocytotic activity of leukocytes in training images. For *Campylobacter* bacteria, we annotate the region where an *Campylobacter* bacteria occurs. On the other hand, for phagocytotic activity of leukocytes, we annotate the regions where leukocytes have phagocytotic activity (phagocytotic), have the look of phagocytotic activity but not (quasi-phagocytotic) and have no phagocytotic activity (non-phagocytotic).

Figure 3 illustrates the images of annotating *Campylobacter* bacteria and leukocytes.

In the detection of *Campylobacter* bacteria (*resp.*, phagocytotic activity of leukocytes), we use about 85% (*resp.*, 90%) of images as training images including validation images and the remained about 15% (*resp.*, 10%) images as test images. Since the number of training images is too small, we increase them at triple by applying data augmentation. After annotating, we resize 640×640 pixels for all the images.

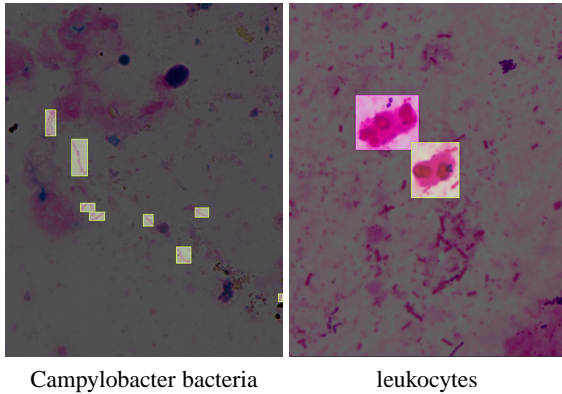


Figure 3: Annotations for Campylobacter bacteria (left) and leukocytes (right).

2.4 Evaluation

In order to evaluate the results of the detection, we adopt the standard measures (Everingham et al., 2010) for object detection. First, we introduce the following *intersection over union* (IoU) between the area P of the predicted box and the area T of the ground truth box:

$$\text{IoU} = \frac{P \cap T}{P \cup T}.$$

For a given threshold δ (%), let TP be the number of the predicted boxes such that $\text{IoU} \geq \delta$, FP the number of the predicted boxes such that $\text{IoU} < \delta$ and FN the number of the ground truth boxes such that $\text{IoU} < \delta$. Then, the standard measures of *precision* and *recall* are defined as follows.

$$\text{precision} = \frac{TP}{TP + FP}, \quad \text{recall} = \frac{TP}{TP + FN}.$$

Also, an *average precision for δ* ($AP\delta$) is defined as the average detection precision under different recalls. We use AP when $\delta = 50$ and $\delta = 75$, that is, $AP50$ and $AP75$. Furthermore, we adopt a (COCO) *mean AP* (mAP) that is an average of APs when varying δ is from 50 to 95 with a step of 5.

3 EXPERIMENTAL RESULTS

In this section, we give the experimental results to detect Campylobacter bacteria and phagocytotic activity of leukocytes.

3.1 Detection of Campylobacter Bacteria

For the detection of Campylobacter bacteria, we use 57 training images including 10 validation images and

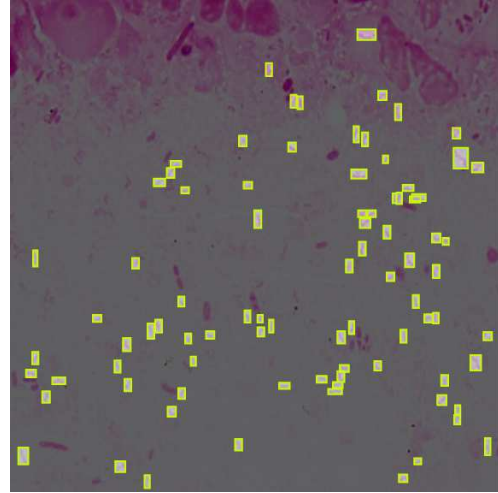


Figure 4: The Gram stained smear image such that Campylobacter bacteria are annotated correctly.

10 test images for total 67 images. Then, Table 2 illustrates the values of mAP , $AP50$ and $AP75$ to detect Campylobacter bacteria by using three detectors.

Table 2: The values of mAP , $AP50$ and $AP75$ to detect Campylobacter bacteria (%).

detector	mAP	$AP50$	$AP75$
Faster R-CNN	5.7	17.9	1.6
RetinaNet	4.1	10.5	1.6
YOLOv5	14.6	43.6	4.1

Table 2 shows that YOLOv5 has larger values of mAP , $AP50$ and $AP75$ than Faster R-CNN and RetinaNet. On the other hand, for all of the Faster R-CNN, RetinaNet and YOLOv5, the values of $AP75$ is much smaller than those of mAP and $AP50$.

By the definition of mAP , the threshold δ such that the value of $AP\delta$ is equal to that of mAP is in the range of $[50, 75]$. Then, for the detected regions of Campylobacter bacteria, the overlap between the predicted box and the ground truth box is not large as IoU .

Next, we represent the results of detecting Campylobacter bacteria in Gram stained smear images. Figure 4 illustrates the Gram stained smear image such that Campylobacter bacteria are annotated correctly.

Then, Figure 5 illustrates the result of detecting Campylobacter bacteria by Faster R-CNN, RetinaNet and YOLOv5 from the Gram stained smear images in Figure 4.

Figure 5 shows that both Faster R-CNN and YOLOv5 detect many regions occurring Campylobacter bacteria, whereas RetinaNet fails to detect. By comparing the results of Faster R-CNN with those of YOLOv5, YOLOv5 detects smaller Cam-



Figure 5: The result of detecting Campylobacter bacteria by Faster R-CNN, RetinaNet and YOLOv5.

phylobacter bacteria which Faster R-CNN cannot detect. Also, Faster R-CNN has the case detecting that non-Campylobacter bacteria are Campylobacter bacteria. Hence, YOLOv5 is the most appropriate detector. On the other hand, the value of AP50 for YOLOv5 is 43.6%, which is the reason that many entwined Campylobacter bacteria exist as Figure 4.

Table 3 illustrates the average running time of detectors for detecting objects of Campylobacter bacteria in one image.

Table 3: The average running time (msec) for detecting objects of Campylobacter bacteria in one image.

detector	time
Faster R-CNN	164.0
RetinaNet	71.8
YOLOv5	36.8

Table 3 shows that YOLOv5 is the fastest in the three detectors, about the half of the average running time of RetinaNet and about the quarter of that of Faster R-CNN.

3.2 Detection of Phagocytotic Activity of Leukocytes

For the detection of phagocytotic activity of leukocytes, we use 91 training images including 10 validation images and 10 test images for total 101 images. Then, Table 4 illustrates the values of mAP, AP50 and AP75 to detect phagocytotic activity of leukocytes by using three detectors.

Table 4: The values of mAP, AP50 and AP75 to detect phagocytotic activity of leukocytes (%).

detector	mAP	AP50	AP75
Faster R-CNN	38.0	62.8	41.2
RetinaNet	18.4	30.0	18.8
YOLOv5	45.5	70.4	53.1

Table 4 shows that YOLOv5 has the largest values of mAP, AP50 and AP75. In contrast to Table 2 in Section 3.1, the values of AP75 is larger than those of mAP in Table 4, the threshold δ such that the value of $AP\delta$ is equal to that of mAP is in the range of [75, 95]. Then, for the detected regions of phagocytotic activity of leukocytes, the overlap between the predicted box and the ground truth box is large as IoU.

Hence, from the viewpoint of the average precision, detecting objects of phagocytotic activity of leukocytes is more successful than detecting objects of Campylobacter bacteria.

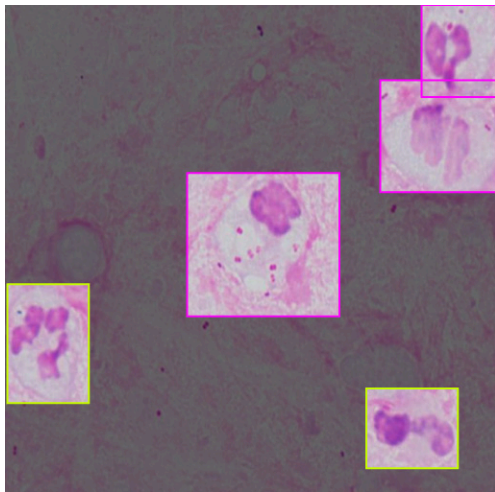


Figure 6: The Gram stained smear image such that phagocytotic activity are annotated correctly, where phagocytotic and non-phagocytotic regions are enclosed by red and yellow colors, respectively.

Table 5: The values of AP50 for phagocytotic (pha), non-phagocytotic (non) and quasi-phagocytotic (quasi) images (%).

detector	pha	non	quasi
Faster R-CNN	83.7	57.3	47.5
RetinaNet	32.1	28.1	29.7
YOLOv5	92.7	59.5	59.1

Next, we represent the results of detecting phagocytotic activity of leukocytes in Gram stained smear images. Figure 6 illustrates the Gram stained smear image such that phagocytotic activity are annotated correctly, where the phagocytotic and the non-phagocytotic regions are enclosed by red and yellow colors, respectively.

Then, Figure 7 illustrates the result of detecting phagocytotic activity of leukocytes by Faster R-CNN, RetinaNet and YOLOv5 from the Gram stained smear images in Figure 6. Here, the phagocytotic images are labeled by “true,” the quasi-phagocytotic images by “false” and the non-phagocytotic images by “no.”

Figure 7 shows that YOLOv5 is the most appropriate detector for phagocytotic activity of leukocytes, which detects almost leukocytes with correct classes. Faster R-CNN detects almost leukocytes but leukocytes with incorrect classes and non-leukocytes substances as leukocytes. On the other hand, RetinaNet is insufficient to detect leukocytes.

Table 5 illustrates the values of AP50 for phagocytotic, non-phagocytotic and quasi-phagocytotic images.

Table 5 shows that the value of AP50 for phagocytotic images by YOLOv5 is much larger than those by

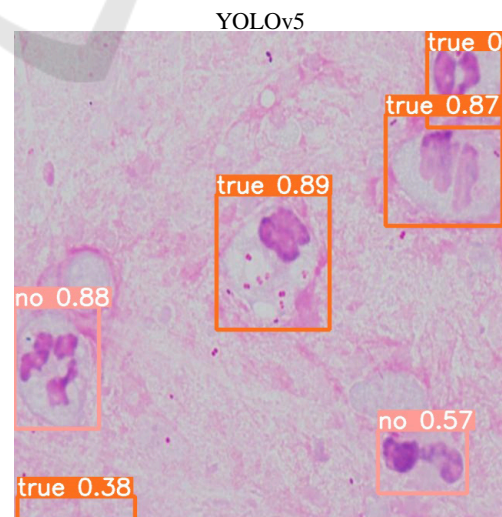
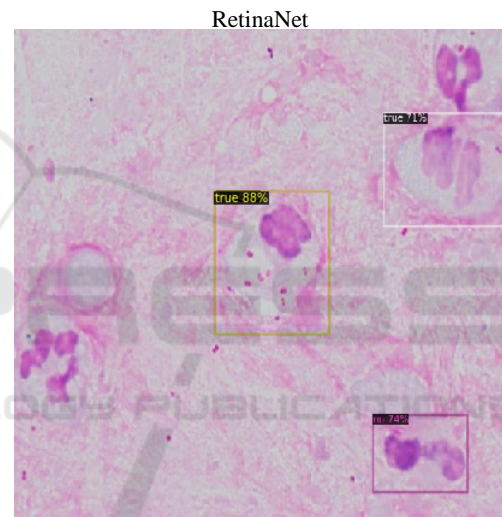
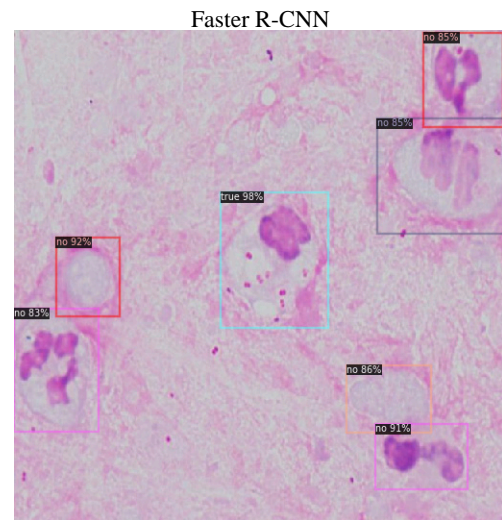


Figure 7: The result of detecting phagocytotic activity of leukocytes by Faster R-CNN, RetinaNet and YOLOv5.

Faster R-CNN and RetinaNet. Since the purpose of this paper is to detect phagocytotic images correctly, YOLOv5 is the most appropriate detector of the purpose. Also, the values of AP50 for non-phagocytotic images by all the detectors are very small. The reason is that quasi-phagocytotic images is misclassified to non-phagocytotic images. Nevertheless, if we regard both quasi- and non-phagocytotic images as non-phagocytotic images, this misclassification can be ignored to detect phagocytotic images correctly. As a result, YOLOv5 succeeds to detect phagocytotic images correctly.

Table 6 illustrates the average running time of detectors for detecting objects phagocytotic activity of Leukocytes in one image.

Table 6: The average running time (msec) for detecting objects of phagocytotic activity of leukocytes in one image.

detector	time
Faster R-CNN	146.0
RetinaNet	69.6
YOLOv5	34.9

Table 6 shows that YOLOv5 is the fastest in the three detectors, about the half of the average running time of RetinaNet and about the quarter of that of Faster R-CNN. By incorporating Table 3 in Section 3.1 with Table 6, we can conclude that YOLOv5 is the fastest in the three detectors for not only Campylobacter bacteria but also phagocytotic activity of Leukocytes.

Furthermore, the running time of detecting objects of Campylobacter bacteria is slightly larger than that of phagocytotic activity of leukocytes but they are almost equal. Hence, the running time of the object detection in this paper is independent from the object as the target.

4 CONCLUSION AND FUTURE WORKS

In this paper, we have detected Campylobacter bacteria and phagocytotic activity of leukocytes in Gram stained smear images by using the detectors of Faster R-CNN, RetinaNet and YOLOv5. Then, RetinaNet have failed to detect them, and YOLOv5 is more appropriate to detect them than Faster R-CNN.

In particular, for phagocytotic activity of leukocytes, YOLOv5 have succeeded to detect almost leukocytes with correct classes. On the other hand, YOLOv5 have succeeded to detect Campylobacter bacteria in many cases, whereas the cases that

Campylobacter bacteria have not detected exist. The reason is that YOLOv5 does not work well for the images having many small objects.

Then, it is a future work to improve YOLOv5 to work well for such images. In particular, we apply YOLOv5 after decreasing the number of objects by dividing an image and enlarging the objects.

Whereas YOLOv5 succeed to detect phagocytotic images at 90% under AP50 in Table 5 as stated in Section 3.2, there exist some images that non-phagocytotic images are detected as phagocytotic. Figure 8 illustrates such images.

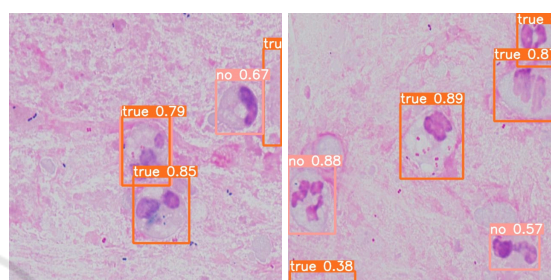


Figure 8: The images that non-phagocytotic images are detected as phagocytotic.

The upper-right region labeled by “no” in the left image in Figure 8 and the lower-left and lower-right regions labeled by “no” in the right image in Figure 8 are not leukocytes. Even if *recall* is more important than *precision* in the medical data, it is a future work to solve this misclassification by improving annotations in test images.

Since the number of training images in this paper is too small to succeed object detection, it is necessary to collect the large number of training images by the medical technologist. Also since we use Gram stained smears images photographed under the same environment, it is necessary to collect the images under the several environment, by using different equipment and different brightness. These are future works.

REFERENCES

- Bartholomew, J. and Mittwer, T. (1952). The Gram stain. *Bacteriol. Rev.*, 16:1–29.
- Carvajal, J., Smith, D., Zhao, K., Wiliem, A., Finucane, P., Hobson, P., Jennings, A., McDougall, R., and Lovell, B. (2014). An early experience toward developing computer aided diagnosis for Gram-stained smear images. In *Proc. CVPR'14*, pages 62–28.
- Everingham, M., Gool, L. V., Williams, C. K. I., Winn, J., and Zisserman, A. (2010). The PASCAL visual object classes (VOC) challenge. *Internat. J. Comput. Vision*, 88:303–338.

- Hashimoto, K., Iida, R., Hirata, K., Matsuoka, K., and Yokoyama, S. (2020). Detecting Geckler classification from Gram stained smears images for sputum. In *Proc. ICPRAM'20*, pages 469–476.
- Iida, R., Hashimoto, K., Hirata, K., Matsuoka, K., and Yokoyama, S. (2020). Detection system of Gram types for bacteria from Gram stained smears images. In *Proc. ICPRAM'20*, pages 477–484.
- Jocher, G. (2020). YOLOv5. <https://github.com/ultralytics/yolov5>.
- Lejon, S. and Andersson, E. (2016). *Semi-automatic segmentation, detection and classification of Gram stained bacteria in blood sample*. Master Thesis, Lund University.
- Lin, T., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. In *Proc. ICCV'17*, pages 2999–3007.
- Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Proc. NIPS'15*, pages 91–99.
- Simonyan, K. and Zisserman, A. (2015). Very deep convolution networks for large-scale image recognition. In *Proc. ICLR'15*.
- Smith, K., Kang, A., and Kirby, J. (2018). Automated interpretation of blood culture Gram stains by use of a deep convolutional neural network. *J. Clin. Microbio.*, 56:e01521–17.
- Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., and Girshick, R. (2019). Detectron2. <https://github.com/facebookresearch/detectron2/>.
- Yoshihara, K. and Hirata, K. (2021). Detecting Campylobacter bacteria and phagocytotic activity of leukocytes from Gram stained smear images. In *Proc. ESKM'21*, pages 10–15.