

## Plasma Cell Detection in Multiple Myeloma Cases using Mask Region Based Convolutional Neural Network Method (Mask R-CNN)

Milyun Ni'ma Shoumi<sup>1</sup>, Radian Malek Rayrendra<sup>2</sup>, Dwi Puspitasari<sup>3</sup>, Pramana Yoga Saputra<sup>4</sup>

<sup>1,2,3,4</sup>Department of Information Technology, State Polytechnic of Malang

email: [milyun.nima.shoumi@polinema.ac.id](mailto:milyun.nima.shoumi@polinema.ac.id)<sup>1</sup>, [radianmalik9@gmail.com](mailto:radianmalik9@gmail.com)<sup>2</sup>, [dwi.puspitasari@polinema.ac.id](mailto:dwi.puspitasari@polinema.ac.id)<sup>3</sup>, [pramanay@polinema.ac.id](mailto:pramanay@polinema.ac.id)<sup>4</sup>

### Abstract

Multiple myeloma cancer is the third major of hematologic malignancy after lymphoma and leukemia, which is about 1% of 13% of hematologic malignancies. Unlike other cancers, myeloma does not form a tumor or lump, but rather causes an accumulation of abnormal plasma cells in the bone marrow which is more than 10% and paraprotein in the body. One of the first steps in diagnosing Multiple Myeloma cancer is by detecting plasma cells in a bone marrow sample taken from the patient's body. Blood samples are taken on several preparations, and the number of plasma cells will be counted from the entire sample. If the number of plasma cells is more than 30% of all cells that have nuclei, then the patient is diagnosed with Multiple Myeloma cancer. The process of detecting plasma cells and calculating the entire sample takes quite a long time and can lead to misdiagnosis due to inaccuracy in the calculation process and the fatigue factor of the medical personnel who check it. In this study, a model was developed to detect Plasma Cells in Multiple Myeloma Cases Using the Mask Region Based Convolutional Neural Network (Mask R-CNN) method, which is expected to speed up the diagnosis process. The use of the Mask Region Based Convolutional Neural Network (Mask R-CNN) method is implemented using the SegPC-2021-dataset for the model training process, and data from the Kepanjen general hospital for the testing process. Using this dataset, the mAp value is 75.94%, the mean precision is 73.93%, and the mean recall is 53.9%.

**Keywords :** multiple myeloma, plasma cell detection, Mask R-CNN, deep learning, object detection

---

**Received:** 15-10-2022 | **Revised:** 04-03-2023 | **Accepted:** 31-03-2023  
DOI: <https://doi.org/10.23887/janapati.v12i1.53119>

---

### INTRODUCTION

Cancer is a condition in which the body has cells that grow out of control and spread to the surrounding healthy tissue [1]. One type of cancer that is often found is Multiple Myeloma (MM). Multiple myeloma cancer is the third major hematologic malignancy after lymphoma and leukemia, which is about 1% of 13% of hematological malignancies. In western countries, this case is found as many as 5.6 per 100,000 population per year. Often found in the elderly with an average age of 70 years [2].

This type of cancer can be caused by abnormal growth of myeloma cells (plasma cells) in the bone marrow, where the number of plasma cells reaches more than 10% [2] and plasma cells are found circulating in the blood. Unlike other cancers, myeloma does not form a tumor or lump, but rather causes an accumulation of abnormal plasma cells in the bone marrow and paraproteins in the body. Myeloma cells can be differentiated from normal plasma cells based on

histological and morphological features, which are generally the methods for diagnosing Multiple Myeloma cancer [1].

One of the first steps in diagnosing Multiple Myeloma cancer is by detecting plasma cells in a bone marrow sample taken from the patient's body. Blood samples are taken on several preparations, and the number of plasma cells will be counted from the entire sample. If the number of plasma cells is more than 30% of all cells that have nuclei, then the patient is diagnosed with Multiple Myeloma cancer. The process of detecting plasma cells and calculating the entire sample takes quite a long time and can lead to misdiagnosis due to inaccuracy in the calculation process and the fatigue factor of the medical personnel who check it. Another problem that arises in detecting plasma cells, including plasma cells that appear in a group of cells, or even isolated from other cells [3].

Plasma cell detection in cases of Multiple Myeloma cancer can be overcome by Deep

Learning. Deep Learning is a subfield of Machine Learning that is represented in several feature hierarchies, where lower feature levels can help higher feature levels [4]. The Deep Learning structure extends the neural network or Neural Network with more hidden layers between the input layer network and the output layer network. There are various types of Deep Learning architectures, one of which is the Mask Region Based Convolutional Neural Network (Mask R-CNN) which is most used in recent years. Mask R-CNN innovatively combines machine learning approaches with deep learning for the segmentation process, namely in the fields of object detection and semantic segmentation [5].

The use of Deep Learning methods has increased the possibility of providing solutions in the medical world, one of which is by detecting plasma cells in cases of Multiple Myeloma cancer [6]. Research [1] explores deep learning methods for detection and segmentation functions in Multiple Myeloma. This study uses the Attention Depplabv3+ regional multi-scale model for myeloma plasma cell segmentation and the U-net model for nucleus instance segmentation. This study also proposes to implement the aggregation function to increase the prediction score.

Research [7] aims to detect and diagnose pancreatic cancer which is quite difficult to do traditionally because the pancreas is a retroperitoneal organ that has a complicated anatomical position. This study accurately applied the Faster R-CNN deep neural network with clinical significance for preoperative pancreatic cancer. Implementation of pancreatic cancer diagnosis using Faster R-CNN can help make the diagnosis more accurate.

In this research, a plasma cell detection system was developed in cases of Multiple Myeloma cancer using a deep learning method, namely Mask R-CNN. This research is an initial study with the limitation of only detecting plasma cells in images. The developed system is expected to help solve the problem of plasma cell detection in cases of Multiple Myeloma cancer in terms of time efficiency required for detection by considering the high workload of a doctor.

## METHOD

This study has several stages, there are: data collection, the process of establishing a plasma cell detection model using the Mask R-CNN method which consists of a model training process, a testing process. The last stage is the calculation of the accuracy of plasma cell detection results.

### A. Data Collection

The dataset used in this study is microscopic images taken from spinal cord aspiration slides of patients diagnosed with Multiple Myeloma (MM). Slides were stained using Jenner-Giemsa stain. The dataset consists of a total of 298 images, which are broken down into three directories namely: data containing training sets, validation sets, and final test sets. This dataset has been used as part of the IEEE ISBI 2021 Challenge SegPC-2021 [8][9][10]. As for data testing, this study used data from Kepanjen General Hospital, Malang, with a total of 17 images. In this normalized image, there is an annotation of the presence of plasma cells. The data available in the dataset is only in the form of the original image and the mask which is an annotation of the location of the plasma cells, and there is no information whether the image includes multiple myeloma cancer.

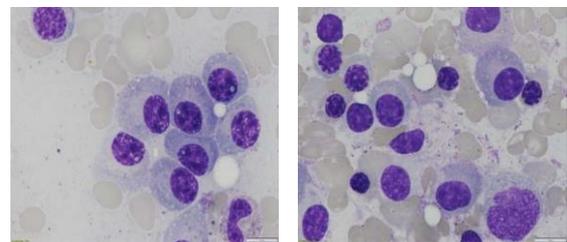


Figure 1. Training Data Samples

The sample in Figure 1 shows the plasma cells used as training data. While the sample in Figure 2 shows an image of plasma cells that will be used as testing data.

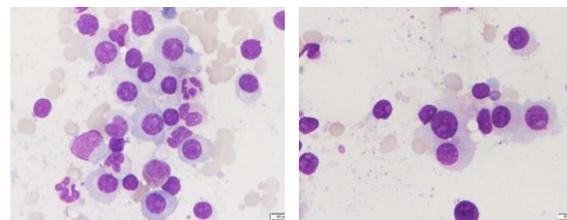


Figure 2. Testing Data Samples

### B. Plasma Cell Detection Method

To be able to detect plasma cells in cases of Multiple Myeloma cancer, several processes are needed as shown in Figure 3. In the Flow Chart Figure 3, the Histogram Equalization process will carry out the input in the form of image data so that the distribution of the gray degree values in the input image is even. The results of the Histogram Equalization will enter the Feature Extraction stage which produces a Feature Map. The Feature Map is the input for the Region Proposal Network process which produces a Region of Interest (ROI) proposal. The two

inputs, which is the Feature Map and the ROI proposal enter the ROI Classification and Bounding Box Regression process, where the ROI Align process which branches into the Segmentation Mask process will produce masks, and the FC Layer which produces the classification and Bounding Box coordinates.

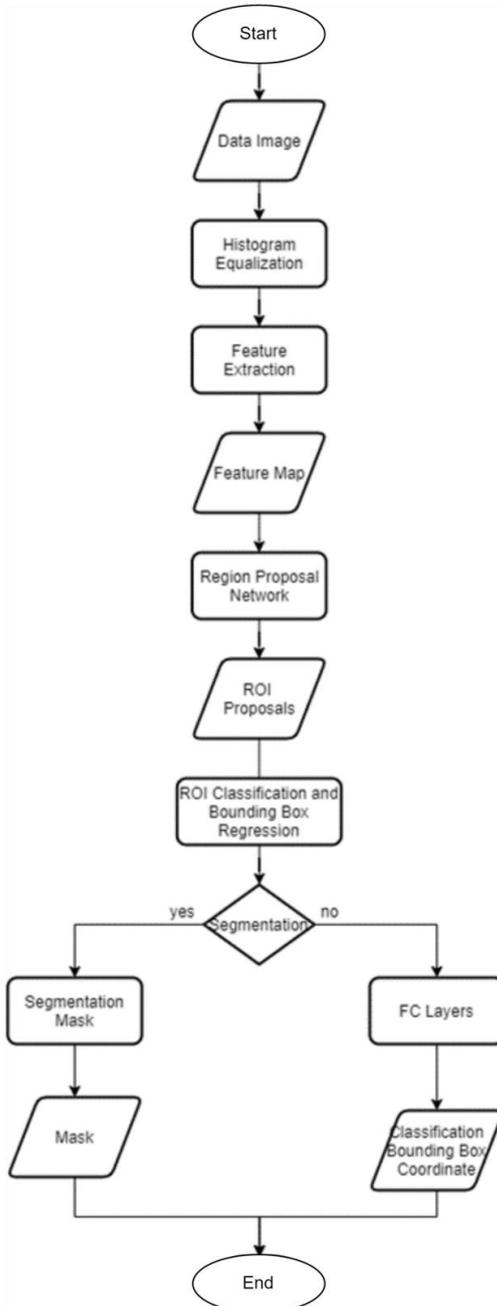


Figure 3. The flow of the plasma cell detection model

### 1. Histogram Equalization

In this process, the input data in the form of an image changes the sharpness values of the image, so that the distribution is uniform, or in

other words, the degree of gray (grayscale) of the number of image pixels is almost the same. The histogram equalization process can be seen in Figure 4.

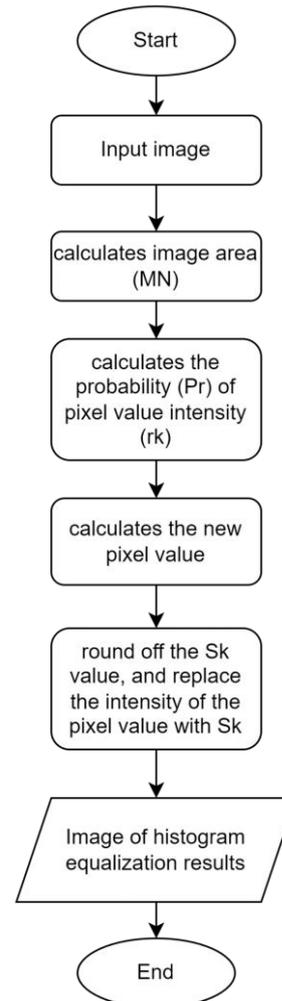


Figure 4. Histogram Equalization Process

From Figure 4, it can be seen that the Histogram Equalization process consists of:

- Enter image data.
- Calculates the area ( $M \times N$ ) of image data (Image Length ( $M$ ) and Image Width ( $N$ ))
- Calculating the probability ( $Pr$ ) of the intensity of the image pixel value ( $rk$ ) by dividing the total value of  $rk$  ( $nk$ ) by the area of the image ( $MN$ )
- Calculating the new pixel value with the formula 1:

$$Sk = (L - 1) \sum_{i=0}^k Pr(r_i) \quad (1)$$

- Round off the new pixel value ( $Sk$ ) and replace the image intensity value ( $rk$ ) with the resulting new pixel value ( $Sk$ ).

f. Generate output images from the histogram equalization process.

## 2. Feature Extraction

Not much different from ConvNet in the Faster R-CNN method, which is tasked with extracting feature maps from images, this research applies the ResNet 101 architecture to extract features from images more reliably. The first step that needs to be done is to retrieve and extract image features using ResNet 101. An illustration of the feature extraction process can be seen in Figure 5.

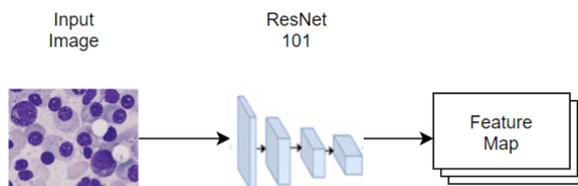


Figure 5. Feature Extraction Process

To extract more reliable features, this research uses ResNet 101 architecture with Feature Pyramid Network (FPN) Backbone. FPN implementation aims to improve object representation better when performing feature extraction, especially on the Region Proposal Network scale.

## 3. Region Proposal Network (RPN)

After going through the feature extraction stage and producing feature maps, the next step is implementing the Region Proposal Network (RPN). The purpose of this stage is to predict whether there is an object in that area or not. It is at this stage that feature maps that store objects are obtained. Feature maps that are forwarded to the RPN network will generate ROI Proposals. The convolutional layer will scan the image using a 3x3 window which generates relevant anchors to represent bounding boxes of different sizes, which will then be distributed throughout the image. There are about 20 thousand anchors of different scales and sizes that will cover the image. Next, a binary classification will be performed to determine whether the anchor contains an object or a background.

## 4. Region of Interest (ROI) Classification

From the ROI proposals and feature maps in the previous stage, it will be forwarded to the ROI Classification stage. This stage classifies ROI in certain classes, such as Plasma Cells or not, which will then increase the size of the bounding box. Bounding Box Regression (BBR) aims to precisely refine the location and size of the bounding box so that it can detect areas of

Myeloma. The bounds on the ROI usually do not match the detailed feature maps, due to the lower sample of the original images. Therefore, it is necessary to change the size of the feature maps by applying the ROI Align layer to extract features with length feature vectors for area candidates. Bounding Box Regression (BBR) will produce bounding boxes according to the Intersection-over-Union (IoU) value. If the anchor has a value of more than 0.5 with ground-truth (GT), then it can be classified as a positive anchor. If the value does not match, it is classified as negative. To calculate IoU, we can use the formula 2:

$$IoU = \frac{Intersection\ Area}{Union\ Area} \quad (2)$$

## 5. Mask Segmentation

This process is carried out through ROI feature maps that are generated by multiple inputs of the same size in ROI Align. After obtaining the ROI value based on the IoU value, a positive ROI value will be identified as input to return the Mask to the existing architecture. This process will return a segmentation mask of 28x28 in each region that has objects which will then be increased during inference to match the dimension box in the ROI bounding box.

## 6. Multi-task Loss

The trained data needs to define multitask loss in each ROI sample.  $L_{mask}$  is defined as the average cross-entropy loss.  $L_{mask}$  allows the network to generate masks for each class. In this case, the authors rely on classification branches to predict the class used in selecting the output mask. Sigmoid at each pixel is the key to providing good segmentation results.

## C. System Test

Testing is done after the implementation phase is complete. The system trial phase will greatly assist research to determine whether the system is running well with the right results. Testing can be done in several ways, there are: (1) Test the running system and make sure there are no errors, (2) Measuring the accuracy value of the pixel-based classification results, whether a cell is a plasma cell or not a plasma cell, using the Confusion Matrix.

Mean Average Precision (mAP) needs to be implemented to produce accuracy values using the Mask RCNN method. If the accuracy results are minimal, then a review is needed starting from coding and classification to produce the best accuracy value. Mean Average Precision is a metric used to measure object accuracy. The

formula used to calculate that is with the following formula 3:

$$mAP = \frac{1}{N} \sum_{i=1}^n AP_i \quad (3)$$

Annotation:

N = Number of AP data

AP = Average Precision

As for getting the AP value, the following formula 4 and 5 is used:

$$precision = \frac{TP}{TP+FP} \quad (4)$$

$$recall = \frac{TP}{TP+FN} \quad (5)$$

Annotation:

TP = True Positive (TP), the condition where the model classifies the data as yes (TRUE) and the actual answer is yes (TRUE).

TN = True Negative (TN), the condition in which the model classifies the data as FALSE, and the actual answer is FALSE

FP = False Positive (FP), a condition where the model classifies the data as yes (TRUE) and the actual answer is no (FALSE)

FN = False Negative (FN), a condition where the model classifies data as no (FALSE) and the actual answer is yes (TRUE).

In the dataset used, for every 1 image, there are several mask images that represent the location of the plasma cells in a microscopic image. This mask image is the ground truth in the dataset. Examples of microscopic images and several mask images can be seen in Figure 6.

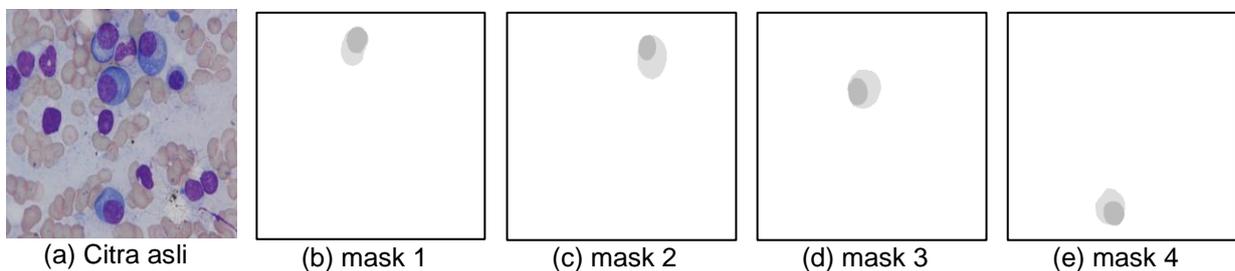


Figure 6. Example of Image Data and Plasma Cell Mask

## RESULT AND DISCUSSION

### A. System Architecture

The architecture of the plasma cell detection system in multiple myeloma cases using the Mask R-CNN method can be seen in Figure 7.

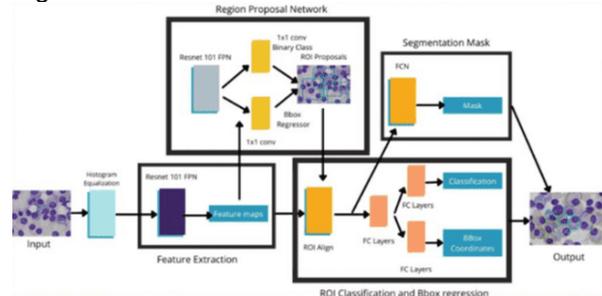
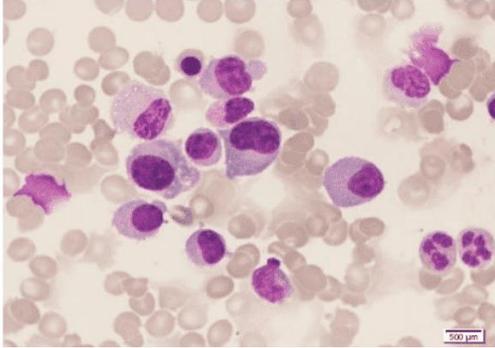
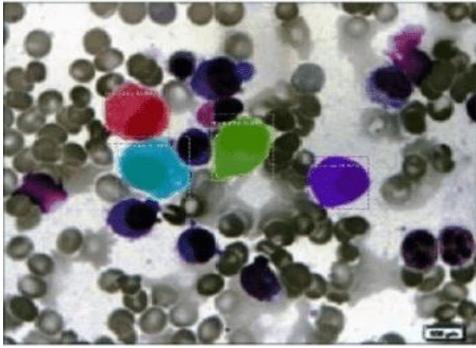
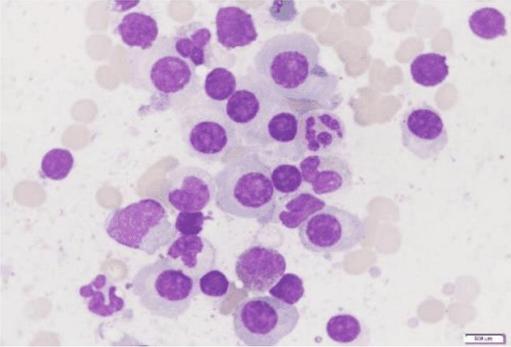
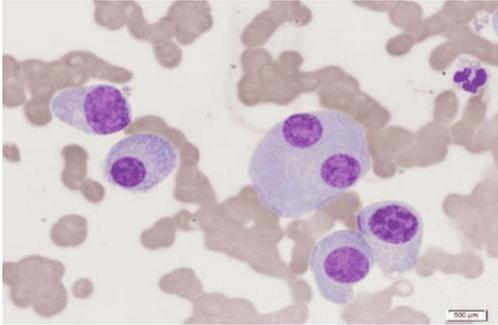
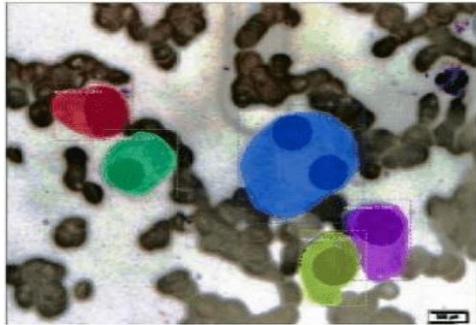


Figure 7. System Architecture

From Figure 7 it can be seen that the image will go through the Histogram Equalization stage to even out the contrast. Then the images will be extracted, thus producing feature maps, and going through the following processes, namely RPN, ROI classification, and segmentation mask as described in the methods section. The back-end network will be built using the Python programming language with the Keras and Tensorflow libraries. Image output will be returned to be displayed on the website page, so that it can be accessed by users. The Mask R-CNN Model developed in Python will be integrated with a website-based application that has a database for data storage.

Table 1. Plasma Cell Detection Results

Input Data	Result
	
	
	

**B. Plasma Cell Detection Results**

The plasma cell detection model in the Multiple Myeloma case was trained using 200 training data and 60 testing data. This data is used to form a model with accuracy calculated using the mAp method of 75%. These results are then implemented by conducting tests using the Kepanjen Hospital dataset. An example of plasma cell detection results can be seen in Table 1.

the accuracy value. Epoch testing is carried out after the Mask R-CNN model is implemented with the COCO Pretrained Model. After the model configuration is complete, the model will be trained according to the specified number of epochs, namely 100 epochs. The configuration for the plasma cell detection model can be seen in Figure 7.

**C. Plasma Cell Detection Results**

In this study, two types of tests were carried out, which is testing the epoch and testing to calculate

```

Configurations:
BACKBONE                resnet101
BACKBONE_STRIDES        [4, 8, 16, 32, 64]
BATCH_SIZE              1
BBOX_STD_DEV            [0.1 0.1 0.2 0.2]
COMPUTE_BACKBONE_SHAPE None
DETECTION_MAX_INSTANCES 100
DETECTION_MIN_CONFIDENCE 0.7
DETECTION_NMS_THRESHOLD 0.3
FPN_CLASSIF_FC_LAYERS_SIZE 1024
GPU_COUNT               1
GRADIENT_CLIP_NORM     5.0
IMAGES_PER_GPU         1
IMAGE_CHANNEL_COUNT     3
IMAGE_MAX_DIM           896
IMAGE_META_SIZE        14
IMAGE_MIN_DIM           896
IMAGE_MIN_SCALE         0
IMAGE_RESIZE_MODE       square
IMAGE_SHAPE             [896 896  3]
LEARNING_MOMENTUM       0.9
LEARNING_RATE           0.001
LOSS_WEIGHTS            {'rpn_class_loss': 1.0, 'rpn_bbox_loss': 1.0,
MASK_POOL_SIZE          14
MASK_SHAPE              [28, 28]
MAX_GT_INSTANCES        100
MEAN_PIXEL              [123.7 116.8 103.9]
MINI_MASK_SHAPE         (56, 56)

```

Figure 7. Configuration of Mask R-CNN Model

Accuracy testing is carried out to find out how much the success rate of the system is in making predictions in conducting classification. Testing is carried out after the training data training process. Tests were carried out with the Mean Average Precision (*mAP*). From the results of testing the *mAP* value of 75% with an accuracy value of 73.93%, which is the result of a True Positive (*TP*) comparison with lots of positive predicted data, these results are supported by the recall value obtained of 53.9% which is the result of a comparison between True Positive (*TP*) with data that is actually positive.

In addition to calculating accuracy, evaluation is also carried out using a loss function. Loss function is a function that is used to find out the difference between the output of an algorithm and the expected output, so that it can be evaluated how well an algorithm is in modeling data and producing the right output. If the detection results contain many errors, then the loss function will produce a high number. Vice versa, the model can be said to be good if the loss function gives a low value. From the results of trials conducted in this study, the loss function diagram can be seen in Figure 8.

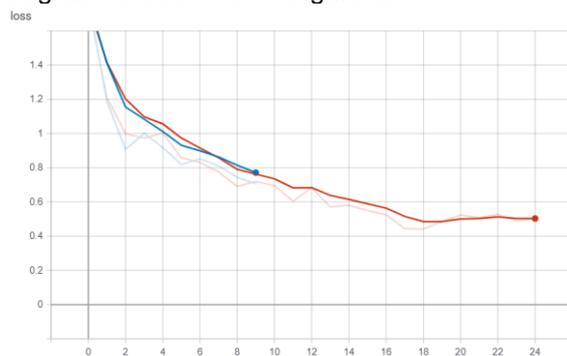


Figure 8. Loss Function Graphic

From Figure 10 it can be seen that the value of the loss function decreases with increasing number of epochs in the training process.

## CONCLUSION

This research has built a website-based Myeloma Detection system with the Laravel framework integrated with the back end using the Flask RESTful API. From the test results, the website-based system can receive input in the form of myeloma images, which will then be automatically stored in the database. Processing The predicted result image will be stored in the database. Plasma detection as an initial stage in diagnosing Multiple Myeloma cancer cases has been built using the Mask Region Based Convolutional Network algorithm using datasets from SegPC-2021-dataset. The RCNN Mask Algorithm is implemented in a website-based system. From the evaluation of the model using the SegPC-2021-dataset, the *mAP* value is 75.94%, the mean precision is 73.93% and a mean recall of 53.9%. This research is an initial study which is only limited to detecting the presence of plasma cells, it is hoped that this research can be developed to produce an output in the form of the number of plasma cells from the entire sample for each subject/patient, so that the percentage of plasma cells in each patient can be known.

## REFERENCES

- [1] B. Afshin, A. Reza, S. Eman, and S. Alaa, "Multi-scale Regional Attention Deeplab3+: Multiple Myeloma Plasma Cells Segmentation in Microscopic Images." PMLR, pp. 47–56, Sep. 16, 2021. Accessed: Aug. 31, 2022. [Online]. Available: <https://proceedings.mlr.press/v156/afshin21a.html>
- [2] D. Hematologi and O. Medik, "Mieloma multipel: aspek patogenesis molekuler," *J. Penyakit Dalam Udayana*, vol. 3, no. 1, pp. 1–7, Jan. 2019, doi: 10.36216/JPD.V3I1.70.
- [3] A. G. Faura, D. Štepec, T. Martinčič, and D. Skočaj, "Segmentation of Multiple Myeloma Plasma Cells in Microscopy Images with Noisy Labels," p. 24, Nov. 2021, doi: 10.48550/arxiv.2111.05125.
- [4] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Futur. Comput. Informatics J.*, vol. 3, no. 1, pp. 68–71, Jun. 2018, doi: 10.1016/J.FCIJ.2017.12.001.
- [5] P. Bharati and A. Pramanik, "Deep Learning Techniques—R-CNN to Mask R-CNN: A Survey," *Adv. Intell. Syst. Comput.*, vol. 999,

- pp. 657–668, 2020, doi: 10.1007/978-981-13-9042-5\_56.
- [6] D. Sagar, “Multiple Myeloma Cancer Cell Instance Segmentation,” Sep. 2021, doi: 10.48550/arxiv.2110.04275.
- [7] J. Liu, M. Mohandes, and M. Deriche, “A multi-classifier image based vacant parking detection system,” *Proc. IEEE Int. Conf. Electron. Circuits, Syst.*, pp. 933–936, 2013, doi: 10.1109/ICECS.2013.6815565.
- [8] A. Gupta, P. Mallick, O. Sharma, R. Gupta, and R. Duggal, “PCSEG: Color model driven probabilistic multiphase level set based tool for plasma cell segmentation in multiple myeloma,” *PLoS One*, vol. 13, no. 12, Dec. 2018, doi: 10.1371/JOURNAL.PONE.0207908.
- [9] A. Gupta *et al.*, “GCTI-SN: Geometry-inspired chemical and tissue invariant stain normalization of microscopic medical images,” *Med. Image Anal.*, vol. 65, Oct. 2020, doi: 10.1016/J.MEDIA.2020.101788.
- [10] S. Gehlot, A. Gupta, and R. Gupta, “EDNFC-Net: Convolutional Neural Network with Nested Feature Concatenation for Nuclei-Instance Segmentation,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2020-May, pp. 1389–1393, May 2020, doi: 10.1109/ICASSP40776.2020.9053633.