Automatic Discriminator of Abnormal Chromosomes Using Deep Learning Algorithms

S. S. Jang[a], S. G. Shin[a], M. J. Lee[a], S. S. Han[a], C. H. Choi[a], S. K. Kim[a], W. S. Cho[a], S. H. Kim[a*],
Y. R. Kang[b], W. S. Jo[b], S. K. Jeong[b] and S. J. Oh[b]
[a]*Department of Advanced Nuclear Engineering , POSTECH, 77 Cheongam-Ro, Nam-Gu, Pohang 37673, Korea*
[b]*Dongnam Inst. Of Radiological & Medical Science, 40 Jwadong-Gil, Jangan-Eup, Gijang-Gun, Busan, Korea*
[*]songhyunkim@postech.ac.kr

1. Introduction

Radiation dose estimation is an important procedure for not only radiological operators but also persons who are exposed by a background or artificial radiations such as cosmic rays, specific soils, and medical exposures. One of the renowned methods of biological dose estimation is Dicentric Chromosome Assay (DCA) discriminating the abnormal chromosomes whether cells have aberration chromosomes or not. Although the previous studies suggested the method to automatically estimate the radiation exposure with DCA [1-5], they still have lower accuracy than manual analysis and require human intervention.

We developed an automatic system of discriminating the abnormal chromosomes using an object detection algorithm based on deep learning method. In this research, Feature Pyramid Network (FPN) [6] was adopted and the chromosome data which were provided from Dongnam Institute of Radiological & Medical Sciences (DIRAMS) were used to train the neural network.

2. Method

2.1. Scope of chromosome discriminating algorithms

The automatic system, which is under development in our research group, conducts DCA using FPN algorithm. The architecture of FPN [6-7] is shown in Figure 1. Inception-Resnet v2 [8] was used as a backbone network to extract feature maps.

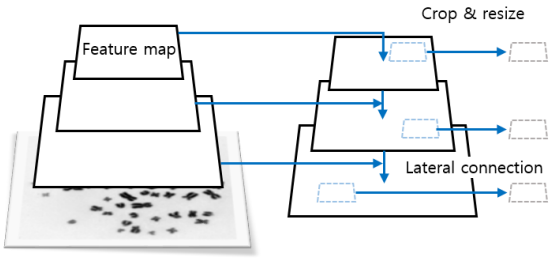

Figure 1. Architecture of faster FPN.

Figure 2. shows various cases discriminated by the network such as normal, dicentric, overlapped chromosomes and fragments. The training data consist of images and annotations with information on locations and sizes of the chromosomes.

To overcome the lack of data and improve the performance of the trained network, the data were augmented by flipping the images horizontally, vertically as shown in Figure 3.

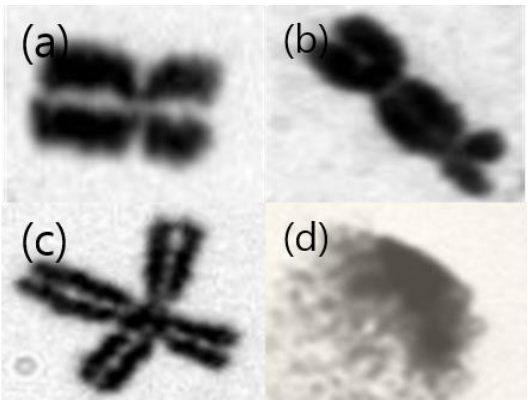

Figure 2. (a) normal, (b) dicentric, (c) overlapped, chromosome images, and (d) fragment images.

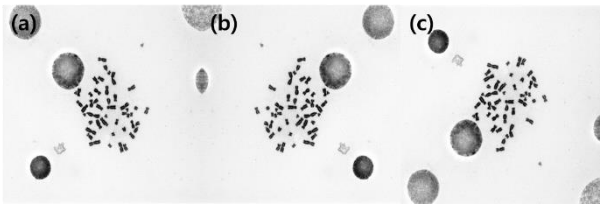

Figure 3. (a) original, (b) horizontal flipped, and (c) vertical flipped images.

2.2. Hyperparameter settings

The training was iterating for 200,000 steps using Adam optimizer [9] with momentum 0.9, input image size 1,280 x 1,024, and batch size 1. We also used learning rate scheduler that is cyclical with minimum learning rate $5x10^{-5}$, maximum learning rate $1x10^{-4}$ and cycle length 50,000 that decays by 0.25 every cycle lengths. The parameters chosen for the machine learning are summarized in Table 1.

Table 1. Hyperparameter information for training the discriminating networks

Image size	1,280 x 1,024
Step	200,000
Batch size	1
Optimizer	Adam
Optimizer momentum	0.9
Learning rate scheduler type	Cyclical
Minimum learning rate	5e-5
Maximum learning rate	1e-4
Cycle length	50,000
Magnitude of decay	0.25

2.3. Performance estimation of the neural networks

For estimating the performance of networks, we have used precision and recall values as criteria. As shown in Table 2, precision and recall are defined as the proportion of the correct detections to all detections and all ground truths respectively. We set the standards for correct detections using Intersection over Union (IoU) value defined as the proportion of intersection area to union area of prediction box and ground truth box.

Table 2. Classification result matrix and definition of precision and recall

Ground Truth	**Prediction result**	
	Positive	**Negative**
Positive	TP	FN
Negative	FP	TN

$$\boldsymbol{Precision} = \frac{TP}{TP + FP} = \frac{TP}{all\ detections}$$

$$\boldsymbol{Recall} = \frac{TP}{TP + FN} = \frac{TP}{all\ ground\ truths}$$

3. Test and results

In the experiments, all networks were trained on single NVIDIA Titan V100 GPU and 2 Inter(R) Xeon(R) Gold 6136 CPUs at 3.000 GHz. 9,904 and 1,219 images were used to train and validate the network, respectively. The model parameters were saved at the step having the lowest validation loss.

During the inference stage, a multi-scale strategy was used to improve the detection accuracy. 1,178 images were used to test the performance of the algorithm, and the results were obtained as shown in Figure 4. The network yields both precision and recall values over 0.9. The results are summarized in Table 3.

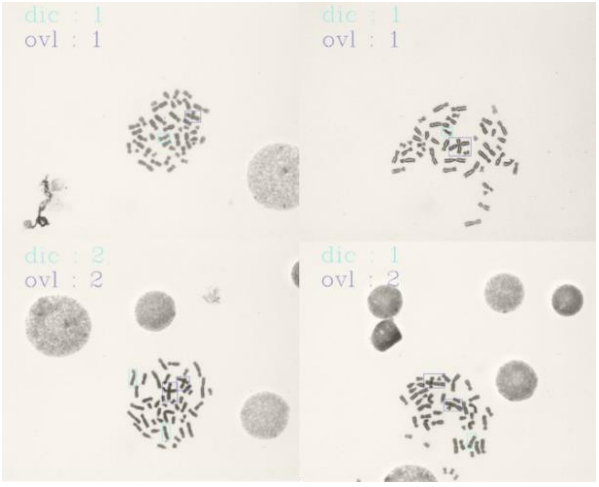

Figure 4. Samples of abnormal chromosomes detections

Table 3. Result of abnormal chromosome detection for test data.

Number of images for performance evaluation	1,178
Precision for test data	0.9035
Recall for test data	0.9056

4. Conclusion

In this study, an automatic system was proposed for discriminating the abnormal chromosomes with the object detection algorithm. The network yields a precision of 90.35% and a recall of 90.56%, respectively. These results showed that the deep learning network has an outstanding performance to distinguish the available data for DCA, and the automatic discriminator can be directly utilized for reducing the estimation time and resources.

Acknowledgment

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT of Korea (MSIT) (NRF-2018M2C7A1A02071506).

REFERENCES

[1] T. L. Ryan et al., “Optimization and validation of automated dicentric chromosome analysis for radiological/nuclear triage applications”, Mutat Res Gen Tox En, 847, (2019)

[2] J. Bi et al., “Rapid and High-Throughput Detection of Peripheral Blood Chromosome Aberrations in Radiation Workers”, International Dose-Response Society, 17, 2 (2019)

[3] A. S. Balajee et al., “Development of electronic training and telescoring tools to increase the surge capacity of dicentric chromosome scorers for radiological/nuclear mass casualty incidents”, Applied Radiation and Isotopes 144, 111-117 (2019)

[4] J. Liu et al., “Accurate cytogenetic biodosimetry through automated dicentric chromosome curation and metaphase cell selection”, F1000Research. 6, 1396 (2017)

[5] X. Shen et al., “A dicentric chromosome identification method based on clustering and watershed algorithm”, Scientific Reports, 9, 2285 (2019)

[6] T. Lin et al., “Feature Pyramid Networks for Object Detection”, arXiv preprint arXiv : 1612.03144v2 (2017)

[7] S. Ren et al., “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks” , arXiv preprint arXiv : 1506.01497v3 (2016)

[8] C. Szegedy et al., “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning” , arXiv preprint arXiv : 1602.07261v2 (2016)

[9] D.P.Kingma, “ADAM : A Method for Stochastic Optimization", arXiv preprint arXiv : 1412.6980 (2014)