

# **IEICE** **TRANSACTIONS**

## **on Fundamentals of Electronics, Communications and Computer Sciences**

**DOI:10.1587/transfun.2024EAL2053**

**Publicized:2025/01/09**

**This advance publication article will be replaced by  
the finalized version after proofreading.**

**A PUBLICATION OF THE ENGINEERING SCIENCES SOCIETY**



**The Institute of Electronics, Information and Communication Engineers  
Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3 chome, Minato-ku, TOKYO, 105-0011 JAPAN**

# Automatic classification of human chromosome shapes using convolutional neural network models

Shinya Matsumoto<sup>†</sup>, Nonmember, Daiki Ikemoto<sup>†</sup>, Nonmember, Takuya Abe<sup>††a</sup>, Nonmember, Kan Okubo<sup>†b</sup>, Regular member and Kiyoshi Nishikawa<sup>†c</sup>, Senior member

**SUMMARY** Metaphase chromosome classifications based on the positional relationship between sister chromatids are used to evaluate the function of the cohesin complex, which tethers sister chromatids until cell division. Currently, classification is manually performed by researchers, which is time consuming and biased. This study aims to automate the analysis using multiple convolutional neural network (CNN)-trained models. By improving our prototype model with a 73.1% concordance rate, one of the proposed new models achieved a maximum concordance rate of 93.33% after applying a fine-tuning method and ensemble learning method. The results suggest that CNN-based models can automatically classify chromosome shapes.

**key words:** Convolutional neural network (CNN), cohesin, image classification, sister chromatid.

## 1. Introduction

Mitotic chromosomes comprise a pair of sister chromatids that form an X-like structure. Until the onset of chromosome segregation, sister chromatids are tethered by a protein complex called cohesin. Deficiency of the cohesin function causes premature sister chromatid separation and is related to the development of cancer and genetic diseases [1], [2], [3].

Direct microscopic observation of a Giemsa-stained metaphase chromosome is a simple and common method to measure the cohesin function deficiency [4], [5]. Currently, researchers manually evaluate and classify chromosomes based on the positional relationship between sister chromatids, which is time-consuming, and the result potentially includes subjective judgment by researchers.

In the field of chromosome analysis using deep learning, almost all of the previous research has focused on karyotype analysis [6], [7], [8], which involves examining and arranging an individual's set of chromosomes in a predefined order. In other words, no previous study has

considered detecting chromosomal abnormalities caused by a deficiency of the cohesin function. Therefore, we aimed to develop convolutional neural network [9] (CNN)-based models to automatically classify the shape of chromosomes, based on the positional relationship between sister chromatids. In our previous study, we trained transfer learning models based on SqueezeNet [10] or ResNet-18 [11] using more than 600 labeled chromosome images classified based on the positional relationship of sister chromatids [12]. Consequently, the SqueezeNet-based model achieved a concordance rate of 73.1% with labeled answers.

In this study, we developed a method to increase the classification accuracy and reduce labor. The main features include: 1) increasing the number of chromosome images and accuracy of data labeling, 2) introduction of automated chromosome clipping, 3) fine-tuning of models, and 4) introduction of ensemble learning. The model achieved a maximum concordance rate of 93.33% with labeled answers. Therefore, we infer that the new method is practically feasible, providing a useful tool for the automatic classification of chromosome shapes.

## 2. Dataset Preparation

### 2.1 Data preparation and labeling

First, we prepared metaphase chromosome spreads from the TK6 human B lymphoblastoid cell line, as previously described [12]. Single-chromosome images were cropped from 500 metaphase cells in the Red, Green, Blue (RGB) color format using OpenCV's `findContours()`. Specifically, `findContours()` detects chromosome outlines and rectangularly crops single chromosomes based on the coordinate information. An example of this result is illustrated in Fig. 1.

In total, 9767 cropped single-chromosome images were labeled by collaborators. Seven participants classified the cropped chromosome images into three classes based on the following classification criteria. If the majority of the responses matched, the images were used as label data. Well-cohered tight chromatids were classified as class A, chromatids in which the arms were separated were classified as class B, and chromatids in which sister chromatids were

<sup>†</sup>The authors are with the Department of Electrical Engineering and Computer Science, Faculty of Systems Design, Tokyo Metropolitan University, Hino-shi, 191-0065 Japan.

<sup>††</sup>The author is with the Division of Biochemistry, Faculty of Pharmaceutical Sciences, Tohoku Medical and Pharmaceutical University, 4-4-1 Komatsushima, Aobaku, Sendai, Miyagi 981-8558, Japan.

a) E-mail: [abe.takuya@tohoku-mpu.ac.jp](mailto:abe.takuya@tohoku-mpu.ac.jp)

b) E-mail: [kanne@tmu.ac.jp](mailto:kanne@tmu.ac.jp)

c) E-mail: [kiyoshi@tmu.ac.jp](mailto:kiyoshi@tmu.ac.jp)

also separated at the centromere were classified as class C [13], [14]. Example images and illustrated images of each class are shown in Fig. 2 and Fig. 3, respectively.

Based on these results, 1146 class A, 7386 class B, and 867 class C chromosome images were obtained from 500 cells. Subsequently, the 867 images of each class (the number of class A and class B chromosome images was reduced to adjust the number of class C chromosome images for unbiased training) were divided into 667 images for training data, 100 images for verification data, and 100 images for test data to fix the number of training data for each class. The same data were always used for the test data, but for the training and validation data, three patterns of datasets were prepared for cross-validation.

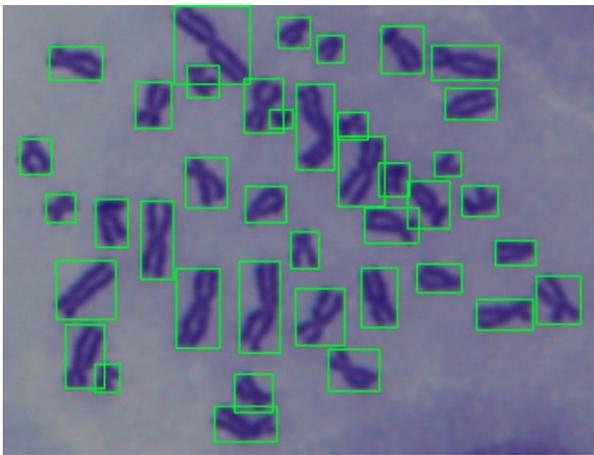


Fig. 1 Detected sister chromatids.

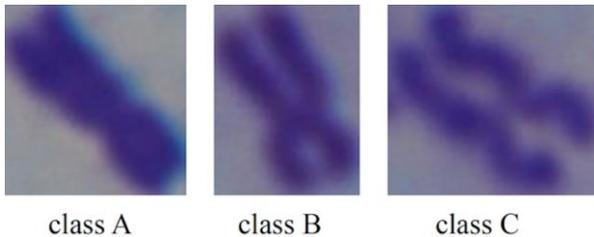


Fig. 2 Example images for each class.

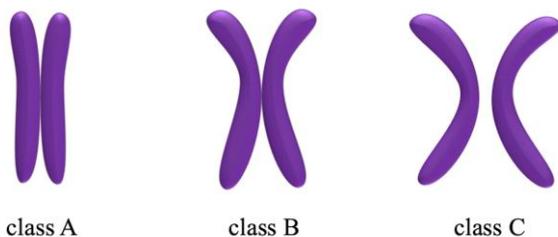


Fig. 3 Illustrated images of sister chromatids of each class.

## 2.2 Preprocessing of input data

Preprocessing was performed on the pixel values to enhance the shape of the chromosome images. First, the mean and standard deviation of the pixel values were set to 90 and 16, respectively, and the pixel value  $src(x, y)$  at position  $(x, y)$  of the image was modified with contrast  $\alpha = 3.0$ , brightness  $\beta = 30$ , and gamma  $\gamma = 3.0$ , based on (1) and (2).

$$dst(x, y) = \alpha src(x, y) + \beta \quad (1)$$

$$y = \left(\frac{x}{255}\right)^\gamma \times 255 \quad (2)$$

Subsequently, OpenCV's `findContours()` was employed to detect contours and remove chromatids or other objects in the images that were not the target of classification. The criteria for removing the objects were that the contours were at the edges of the image and the areas of the contours were less than 8% of the cropped image area. The removed areas in the image were replaced with the pixel values  $(R, G, B) = (255, 255, 255)$ . An example of preprocessing input data is shown in Fig. 4.

Finally, the images were resized to  $224 \times 224$  or  $299 \times 299$  pixels based on the standard input image size in each pretrained model.

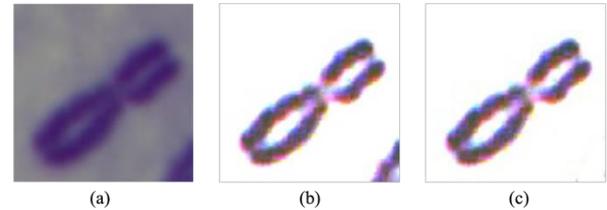


Fig. 4 An example of preprocessing input data. (a) Cropped image. (b) Image with enhanced chromosome shape. (c) Image after removing the partially captured chromosome according to the criteria.

## 3. Image Classification Model

### 3.1 Transfer learning model

The number of chromosome images (667 for each class) did not appear to be sufficient to start machine learning from the beginning. Therefore, we used transfer learning to improve the classification accuracy, even with a small amount of image data, by using a pretrained model as a feature extractor [15]. The fully connected layers (fc) prior to the network output were trained using the chromosome dataset, and the other layers of the CNNs were used as fixed feature extractors.

Fig. 5 shows the total and per-class concordance rates between the labeled data and predicted answers from the SqueezeNet pretrained models [16]. The concordance rates improved with the number of training images reached a peak at 300 images, and did not further improve as the number of

images increased. Thus, we considered 667 images per class to be sufficient for transfer learning.

In our previous study, we compared SqueezeNet [10] and ResNet-18 [11] and showed that the SqueezeNet-based model is more suitable for chromosome classification [12]. In this study, as shown in Fig. 6, we examined 12 pretrained models: AlexNet [17], efficientnet\_b7 [18], DenseNet161 [19], vgg19 [20], inception\_v3 [21], GoogleNet [22], MobileNetV2 [23], ShuffleNetV2 [24], ResNext50 [25], wide\_resnet101\_2 [26], MnasNet [27], and SqueezeNet [10]. All these models are provided by Torchvision. The maximum concordance rate was 81.31% in the SqueezeNet-based model, followed by the ShuffleNet-based model (78.35%), DenseNet161-based model (78.04%), and MobileNetV2-based model (77.22%), respectively.

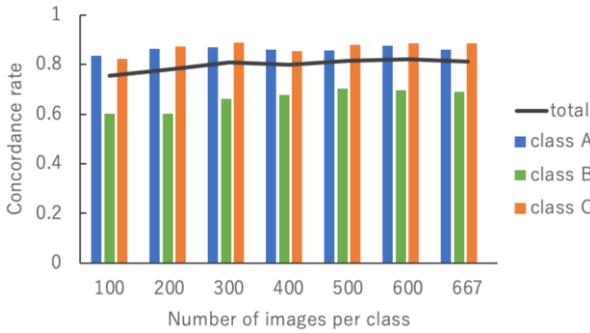


Fig. 5 Concordance rates by the number of images (total and per class).

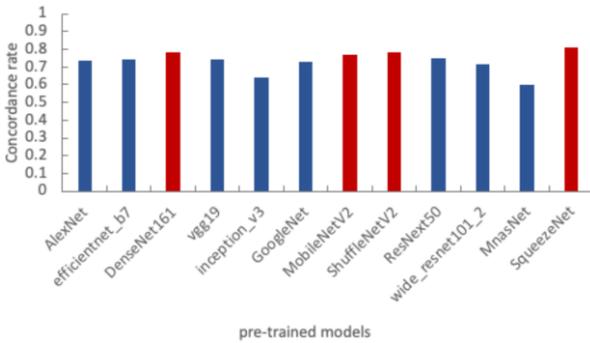


Fig. 6 Concordance rate per pretrained model.

### 3.2 Fine-tuning of models

We fine-tuned the four models that achieved concordance rates higher than 75% in transfer learning. Fine-tuning is a method of retraining the parameters of an arbitrary layer of a trained model to increase its generality. In this study, each of the four pretrained models was tested by increasing the number of layers to be retrained in stages and using the model with the highest concordance rate.

Table 1 lists the highest concordance rate and the layers in which fine tuning was performed for each of the four models. Following fine-tuning, the concordance rates of all

models improved (Table 1). The maximum concordance rate was 89.67% in the DenseNet161-based model, followed by the ShuffleNet-based model (89.58%), MobileNetV2-based model (85.98%), and SqueezeNet-based model (85.06%), respectively.

Table 1 Highest concordance rate and the layers in which fine tuning was performed.

Models	Retrained layers	Concordance rate
DenseNet161	From Dense Block2 to Fully Connected layer	0.8967
MobileNetV2	All layers	0.8598
ShuffleNetV2	From Stage3 to Fully Connected layer	0.8958
SqueezeNet	From Fire8 to Conv10	0.8506

### 3.3 Ensemble learning

Ensemble learning is employed to improve the generalization of models by training multiple models individually and averaging their outputs during inference. It aims to further improve the concordance rate by considering the average score obtained by ensemble learning as the final score for the output of the fine-tuning models developed based on the four trained models in the previous section.

In ensemble learning, we obtained the inference outputs of each model selected for ensemble learning from the four fine-tuning models, obtained their arithmetic mean by class, and used the average score as the final output. Subsequently, the concordance rate with the label data was obtained. We assembled all the possible combinations from four different fine-tuning models.

Table 2 lists the combinations of the base pretrained models and the concordance rates for their ensemble learning. The highest result of 93.33% was obtained when ensembling the outputs of DenseNet161-, MobileNetV2-, and SqueezeNet-based models. The result is a 3.75 percentage point improvement over the DenseNet161-based single model at the time of fine-tuning.

Table 3 lists the confusion matrix when ensembling the outputs of DenseNet161-, MobileNetV2-, and SqueezeNet-based models. The concordance rates for each class were 97%, 85%, and 98%, respectively. Since 10% of class B chromosomes were misclassified as class C, it seems difficult to distinguish between chromosomes that are slightly attached and those that are detached. Considering that the classification boundary could vary even more if analyzed by different human researchers [12], the 85% classification accuracy for class B can be considered sufficient for practical use.

**Table 2** Combination of base pretrained models and concordance rate during ensemble.

DenseNet161	MobileNetV2	ShuffleNetV2	SqueezeNet	Concordance rate
✓	✓	✓	✓	0.9133
✓	✓	✓		0.9267
✓	✓		✓	0.9333
✓		✓	✓	0.9200
	✓	✓	✓	0.9067
✓	✓			0.9200
✓		✓		0.9167
✓			✓	0.9200
	✓	✓		0.9200
	✓		✓	0.9000
		✓	✓	0.9133
✓				0.8967
	✓			0.8598
		✓		0.8958
			✓	0.8506

**Table 3** Confusion matrix when ensembling the outputs of DenseNet161-, MobileNetV2-, and SqueezeNet-based models.

		Predicted label		
		class A	class B	class C
True label	class A	0.97	0.02	0.01
	class B	0.05	0.85	0.10
	class C	0.00	0.02	0.98

**4. Discussion**

Chromosome analyses, including karyotyping and chromosomal aberration detection, are generally manually performed by researchers. However, such manual analyses is laborious, time-consuming, and involve the risks of individual differences and subjective judgement, highlighting the importance of automated analyses. Applying image recognition models to chromosome analyses is an attractive strategy, and many scholars have reported CNN-based chromosome analyses. However, most studies that have applied image recognition models to chromosome analyses have focused on karyotyping, and it is not clear whether image recognition models can be applied to the detection of chromosomal aberrations. In our previous study, we applied CNN-based models to detect sister chromatid cohesion defects [12]. Although our previous model achieved a maximum concordance rate of 73.1 % with the labeled answers, the rate was insufficient for practical use.

Here, we succeeded in increasing the concordance rate to a maximum of 93.33%. In particular, the fine-tuning of the models and ensemble learning greatly contributed to the improvement in the concordance rate. The pretrained models were trained only on natural images and did not include training data for chromosomes. Therefore, we believe that the feature extractor can be retrained to effectively extract features from the chromosomes.

In addition to sister chromatid cohesion defects, many other types of chromosomal aberrations have been reported in human chromosomes, including chromosome breaks,

sister chromatid exchange, chromosome fusion, and translocation. Future studies can examine whether CNN-based models can be applied to detect other chromosomal aberrations.

**5. Conclusion**

In this study, we aimed to automate chromosome analysis using multiple CNN-trained models. The prototype model had a concordance rate of 73.1% with labeled data, and one of our newly developed models achieved a maximum concordance rate of 93.33% upon applying a fine-tuning method and an ensemble learning method. The fine-tuning of the models and ensemble learning significantly contributed to the improvement. These results suggest that CNN-based models have the potential to be used as practical tools to automatically classify chromosome shapes.

**Acknowledgments**

This work was supported by Grants from JSPS KAKENHI (22H05072, 22K12170).

**References**

- [1] S. Remeseiro, A. Cuadrado, and A. Losada, "Cohesin in development and disease," *Development*, vol.140, no. 18, pp.3715–3718, Sep. 2013.
- [2] Y. Ochi, A. Kon, T. Sakata, M. M. Nakagawa, N. Nakazawa, M. Kakuta, K. Kataoka, H. Koseki, M. Nakayama, D. Morishita, T. Tsuruyama, R. Saiki, A. Yoda, R. Okuda, T. Yoshizato, K. Yoshida, Y. Shiozawa, Y. Nannya, S. Kotani, Y. Kogure, N. Kakiuchi, T. Nishimura, H. Makishima, L. Malcovati, A. Yokoyama, K. Takeuchi, E. Sugihara, T.-A. Sato, M. Sanada, A. Takaori-Kondo, M. Cazzola, M. Kengaku, S. Miyano, K. Shirahige, H. I. Suzuki, and S. Ogawa, "Combined Cohesin–RUNX1 Deficiency Synergistically Perturbs Chromatin Looping and Causes Myelodysplastic Syndromes," *Cancer Discov*, vol.10, no. 6, pp.836–853, Jun. 2020.
- [3] D. A. Solomon, T. Kim, L. A. Diaz-Martinez, J. Fair, A. G. Elkahlon, B. T. Harris, J. A. Toretsky, S. A. Rosenberg, N. Shukla, M. Ladanyi, Y. Samuels, C. D. James, H. Yu, J.-S. Kim, and T. Waldman, "Mutational Inactivation of *STAG2* Causes Aneuploidy in Human Cancer," *Science (1979)*, vol.333, no. 6045, pp.1039–1043, Aug. 2011.
- [4] T. Abe, R. Kawasumi, H. Arakawa, T. Hori, K. Shirahige, A. Losada, T. Fukagawa, and D. Branzeti, "Chromatin determinants of the inner-centromere rely on replication factors with functions that impart cohesion," *Oncotarget*, vol.7, no. 42, pp.67934–67947, Oct. 2016.
- [5] W. Deng, S. W. Tsao, J. N. Lucas, C. S. Leung, and

- A. L. M. Cheung, "A new method for improving metaphase chromosome spreading," *Cytometry Part A*, vol.51A, no. 1, pp.46–51, Jan. 2003.
- [6] D. Menaka and S. G. Vaidyanathan, "Chromenet: a CNN architecture with comparison of optimizers for classification of human chromosome images," *Multidimens Syst Signal Process*, vol.33, no. 3, pp.747–768, Sep. 2022.
- [7] C. Yang, T. Li, Q. Dong, and Y. Zhao, "Chromosome classification via deep learning and its application to patients with structural abnormalities of chromosomes," *Med Eng Phys*, vol.121, p.104064, Nov. 2023.
- [8] M. D'Angelo and L. Nanni, "Deep-Learning-Based Human Chromosome Classification: Data Augmentation and Ensemble," *Information*, vol.14, no. 7, p.389, Jul. 2023.
- [9] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol Cybern*, vol.36, no. 4, pp.193–202, Apr. 1980.
- [10] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," Feb. 2016.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.770–778, Jun. 2016.
- [12] D. Ikemoto, T. Taniguchi, K. Hirota, K. Nishikawa, K. Okubo, and T. Abe, "Application of neural network-based image analysis to detect sister chromatid cohesion defects," *Sci Rep*, vol.13, no. 1, p.2133, Feb. 2023.
- [13] P. van der Lelij, K. H. Chrzanowska, B. C. Godthelp, M. A. Rooimans, A. B. Oostra, M. Stumm, M. Z. Zdzienicka, H. Joenje, and J. P. de Winter, "Warsaw Breakage Syndrome, a Cohesinopathy Associated with Mutations in the XPD Helicase Family Member DDX11/ChlR1," *The American Journal of Human Genetics*, vol.86, no. 2, pp.262–266, Feb. 2010.
- [14] J. de Lange, A. Faramarz, A. B. Oostra, R. X. de Menezes, I. H. van der Meulen, M. A. Rooimans, D. A. Rockx, R. H. Brakenhoff, V. W. van Beusechem, R. W. King, J. P. de Winter, and R. M. F. Wolthuis, "Defective sister chromatid cohesion is synthetically lethal with impaired APC/C function," *Nat Commun*, vol.6, no. 1, p.8399, Oct. 2015.
- [15] S. Bozinovski, "Reminder of the First Paper on Transfer Learning in Neural Networks, 1976," *Informatica*, vol.44, no. 3, Sep. 2020.
- [16] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, "A survey of modern deep learning based object detection models," *Digit Signal Process*, vol.126, p.103514, Jun. 2022.
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun ACM*, vol.60, no. 6, pp.84–90, May 2017.
- [18] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," May 2019.
- [19] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2261–2269, Jul. 2017.
- [20] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014.
- [21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2818–2826, Jun. 2016.
- [22] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1–9, Jun. 2015.
- [23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.4510–4520, Jun. 2018.
- [24] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design," Jul. 2018.
- [25] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.5987–5995, Jul. 2017.
- [26] S. Zagoruyko and N. Komodakis, "Wide Residual Networks," May 2016.
- [27] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "MnasNet: Platform-Aware Neural Architecture Search for Mobile," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.2815–2823, Jun. 2019.