Raw G-Band Chromosome Image Segmentation Using U-Net Based Neural Network

Emrecan Altinsoy[1(✉)], Can Yilmaz[1], Juan Wen[2], Lingqian Wu[2], Jie Yang[1], and Yuemin Zhu[3]

[1] School of Electronic, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
{emrecanaltinsoy,montagnard,jieyang}@sjtu.edu.cn
[2] Center for Medical Genetics, School of Life Sciences, Central South University, Changsha 410078, China
{juanwen,wulingqian}@sklmg.edu.cn
[3] Creatis, INSA Lyon, Villeurbanne, France
zhu@creatis.insa-lyon.fr

Abstract. Chromosome analysis plays an important role in investigating one's genetic disorders and abnormalities. Many works are done on automating this operation for decades. Segmentation of chromosomes is the first step of this process, and it is essential for the next step which is classification. However, it is not an easy task due to a very noisy background, the presence of other cells and the variation of chromosome structures. In this paper, we propose a raw G-band chromosome image segmentation method using U-net based convolutional neural network. To this end, we constructed a raw G-band chromosome dataset which consists of 40 images. In order to prevent over-fitting, we implemented augmentations on the training and the validation set images. The trained model achieved 96.97% dice score. The experimental results showed that, the convolutional neural network can provide satisfying results, especially with highly noisy images.

Keywords: Raw G-band chromosome image · Segmentation · U-net · Convolutional neural network · Deep learning

1 Introduction

Chromosomes are considered the main genetic information carriers. Every healthy individual has 46 chromosomes in total, 22 pairs of non-sex and two sex chromosomes. Analyzing the quantity and the structure of these chromosomes helps us to investigate one's genetic disorders (e.g. Down syndrome and

Turner syndrome) and genetic abnormalities [3]. It is difficult to differentiate the chromosomes under a light microscope. Therefore, scientists developed different staining methods (bandings), which are used to color different parts of the chromosomes, in order to display their structural details. These bandings make the chromosome identification easier and more reliable. The most known ones are G-banding, Q-banding, R-banding, and C-banding. G-band images are obtained by using Giemsa dye, which reacts differently for each nucleic acid base pairs. It leads to a set of bright or dark gray bands throughout the chromosomes. Today, G-banding is the most used method for karyotyping.

Thanks to the improvements in computer science in the last 30 years, the use of computers on chromosome analysis is now achievable. Automated chromosome analysis has been studied by many researchers. Preprocessing and segmentation of chromosomes from the background pixels are the first step for an automated solution and necessary for the next steps, which are the classification of chromosomes and karyotyping. However, it is a challenging task for several reasons: a very noisy background, the presence of other cells and the variation of chromosome structures.

Otsu thresholding method [11] was used by Ji [5]. However, using a global threshold value caused losses on the bright chromosome parts and the chromosome satellites. To cope with this issue, Ji [6] and Stanley [15] proposed a local re-thresholding method. This method consists of two steps. Firstly, Otsu thresholding is applied on the whole image. Secondly, the objects are separated from the segmented image using connected components and then applied a second Otsu thresholding on each object. Grisan et al. [4] presented locally adaptive thresholding for the segmentation of Q-band chromosome images. The images are divided into small squares, 100 pixels by 100 pixels, and Otsu thresholding was applied on each square. Then, the squares are resized to the original image size using bilinear interpolation to obtain a pixel-wise map. Sugapriyaa et al. [16] adopted this method and applied it on G-band metaphase images. Lerner [9], Cao et al. [1] and Soumya [14] employed clustering based segmentation methods for the chromosome and background separation, such as K-means clustering and fuzzy C-means clustering algorithm. Poletti et al. [12] implemented and compared the performance of different thresholding methods on Q-band chromosome images in their review paper. As observed in the article, adaptive and region based thresholding methods provided better results compared to global thresholding methods. Yilmaz et al. [17] cleared the background pixels of G-band chromosome images by implementing a thresholding method based on the peak value of the image histogram. Afterward, the interphase cells were removed using their characteristics and Gaussian filter was applied to obtain clear chromosome borders. Unlike the others, in [17] raw images are used. Most of the works carried in this field need human interaction to achieve good results. In addition, the used images are preprocessed: the chromosomes are already separated from the background pixels and the noise is reduced or removed.

In this paper, we introduced a U-net [13] based neural network for segmentation of raw G-band chromosome images. First, we created the dataset with the

images taken from Renji Hospital. For the purpose of improving the segmentation results and to prevent over-fitting, we implemented augmentations on the images in the dataset. For the evaluation of the model, Jaccard distance loss and dice coefficient loss are employed as metrics, since these metrics provide a better evaluation for the instance segmentation processes. Finally, we compared the segmentation results with the local adaptive thresholding results.

2 Method

2.1 Dataset and Data Augmentation

For the experiments, the dataset is created with the images taken from Renji Hospital. It consists of 40 raw G-band chromosome images, 25 images for the training set, 5 images for the validation set and 10 images for the test set. The images are all in the same resolution, 1200×1600 pixels. However, it is reduced to 480×640 pixels to be able to train the model with the available GPU memory. The training set and validation set images are labeled manually. Raw G-band chromosome image and its mask are shown in Fig. 1.
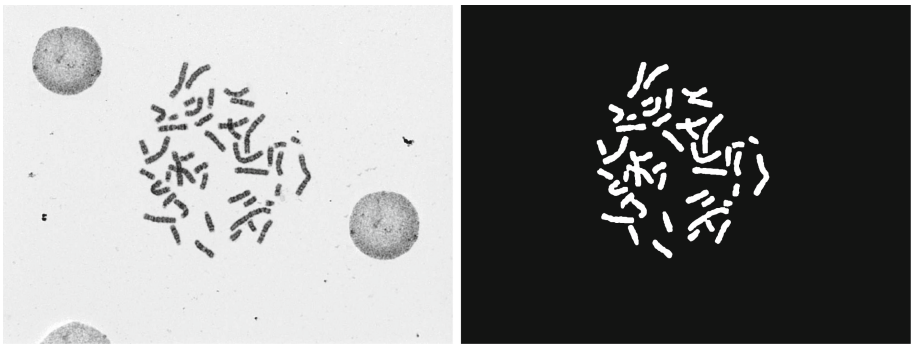

Fig. 1. Raw G-band chromosome image and its mask.

Since the images in the dataset are not various and not numerous enough to obtain satisfactory results and to prevent over-fitting, augmentations are used to increase the number of images in the training set and in the validation set up to 3500 and 700, respectively. The summary of the applied augmentation methods are given in Table 1.

2.2 U-Net Based Neural Network

Segmentation of medical images has constraints due to its characteristics, such as detailed patterns, unclear object boundaries, among others. Long et al. [10] proposed to use skip connections which combine two convolutional layers (encoding

Table 1. Summary of applied augmentations

Method	Range
Rotation	$\pm 45°$
Horizontal flip	50% chance
Vertical flip	50% chance
Width shift	5% of the total width
Height shift	5% of the total height
Shear intensity	10° in the counter-clockwise direction
Zoom	$\pm 10\%$

and decoding layers) to generate better segmentation results. Drozdzal et al. indicated that skip connections can be used on biomedical images [2]. Ronneberger et al. [13] also adopted this idea and proposed the original U-net architecture for biomedical image segmentation.
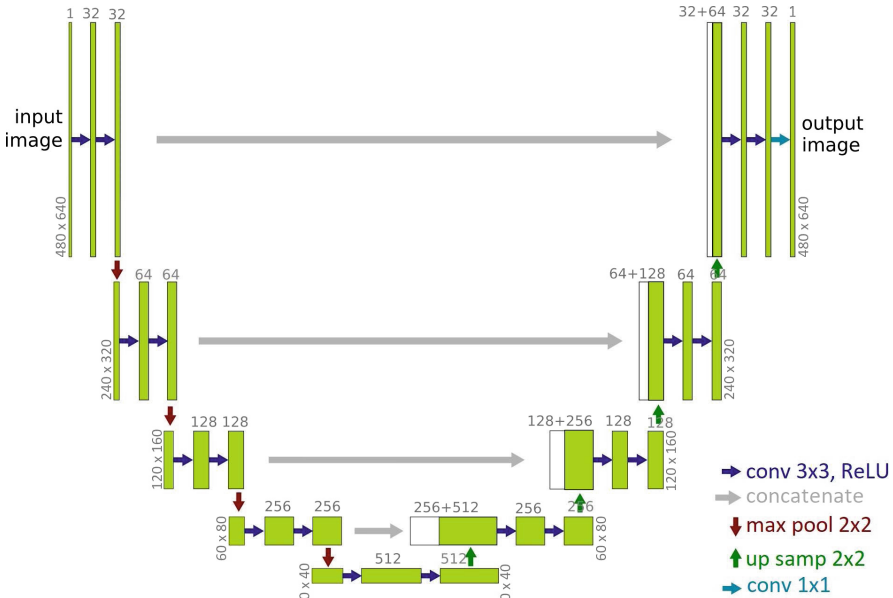
Fig. 2. U-net architecture

Original U-net model [13] is symmetrical fully convolutional neural network and it consists of two parts which are down-sampling (left side) and up-sampling (right side). In total the network has 9 convolutional blocks and each block consists of two convolutional layers with 3×3 kernel size. On the down-sampling

part, each convolutional block is followed by a max pooling layer with 2×2 pool size. After every max pooling layer, the size of the feature maps is divided by two and the number of feature maps is duplicated by two. On the up-sampling part, before every convolutional block, there is an up-sampling layer with the size of 2×2. The output of the up-sampling layers concatenates with the corresponding feature maps from the down-sampling part. Unlike the down-sampling part, after every convolutional block, the size of the feature maps is duplicated by two and the number of feature maps divided by two. All convolutional layers use ReLU as an activation function, except the last one which uses sigmoid function.

In this study, the U-net architecture (Fig. 2) is used with several changes:

– The number of the feature maps are half of the original U-net architecture to be able to train the model with the available GPU memory.
– Input and Output image sizes are changed to 480×640.
– Since the number of future maps is large compared to the number of images in the training dataset, to avoid over-fitting dropout layer is added after 4th (before the max pooling layer) and 5th convolutional blocks.
– Adam optimizer [8] is used instead of stochastic gradient descent optimizer.
– Jaccard distance loss [7] and dice similarity coefficient are used for the evaluation of the model.

2.3   Evaluation Metrics

Jaccard Index. In biomedical image segmentation, binary cross-entropy is not a good indicator for the evaluation. Instead, the Jaccard index, also known as intersection over union score (IoU), is widely used. It measures the similarity between ground truth and predicted masks. It is defined as the intersection of the masks divided by the union of the masks and it returns a value between 0 and 1.

$$Jacc = \frac{|M_{pred} \cap M_{truth}|}{|M_{pred} \cup M_{truth}|} = \frac{|M_{pred} \cap M_{truth}|}{|M_{pred}| + |M_{truth}| - |M_{pred} \cap M_{truth}|}$$

$$= \frac{\text{True Positive}}{\text{Union}} \tag{1}$$

where M_{pred} is the vector of predicted mask and M_{truth} is the vector of ground truth mask.

As mentioned before, for the evaluation of the model, Jaccard distance loss [7] is used. Jaccard distance measures the dissimilarity between ground truth and the predicted masks, and it can be derived as

$$1 - Jacc = \frac{|M_{pred} \setminus M_{truth}| + |M_{truth} \setminus M_{pred}|}{|M_{pred} \cup M_{truth}|}$$

$$= \frac{\text{False Positive} + \text{False Negative}}{\text{Union}} \tag{2}$$

Dice Similarity Coefficient. Dice similarity coefficient (DSC) is the other metric widely employed in instance segmentation problems. It calculates the spatial overlap between the predicted mask and the ground truth mask. It is defined as two times the intersection of the ground truth and the predicted masks divided by the sum of the masks (Eq. 3) and it returns a value between 0 and 1.

$$DSC = \frac{2 \times |M_{pred} \cap M_{truth}|}{|M_{pred}| + |M_{truth}|}$$

$$= \frac{2 \times \text{True Positive}}{\text{False Positive} + \text{False Negative} + (2 \times \text{True Positive})} \tag{3}$$

3   Experiments

3.1   Training

The model is implemented with Keras. As mentioned earlier, Adam optimizer [8] is adopted for the model weights estimation. The beginning learning rate is set to 0.0001, batch size is set to 2 and the number of epochs is set to 100. Keras callback functions are used during the training. The learning rate is multiplied by 0.2 when the loss metric stopped improving for two epochs. On the 22nd epoch, the training was stopped, because the loss metric had not been improved for the last 5 epochs. At the end of the training, the lowest Jaccard distance loss recorded on the 17th epoch as 0.0589. The dice similarity coefficient was recorded as 96.97%. Learning curves of Jaccard distance loss and dice similarity coefficient are shown in Fig. 3.
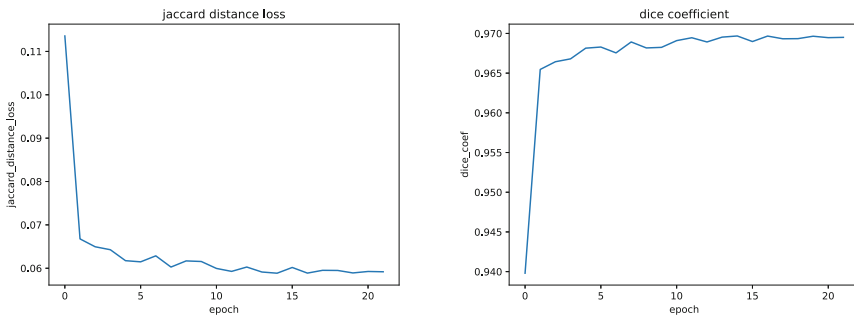

Fig. 3. Jaccard distance loss and dice similarity coefficient learning curves

3.2 Results

The trained model is tested on several images. It takes around 0.25 s to segment one image. Segmentation results are depicted in Fig. 4. As it can be observed, the proposed method also clears the interphase cells with the background pixels. Besides, the chromosome boundaries are almost completely preserved.
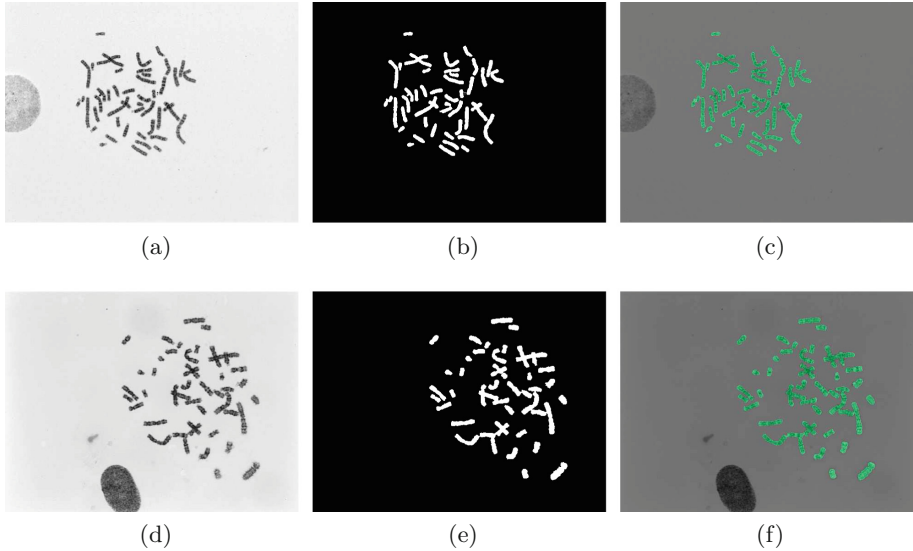

Fig. 4. Segmentation results. (a) (d) Raw G-band chromosome images, (b) (e) The predicted masks, (c) (f) Overlayed predicted masks

The proposed method is compared with the local adaptive thresholding method. The comparison results of the two methods are shown in Figs. 5 and 6. When the G-band chromosome image is not very noisy, both methods provide satisfying results. However, the proposed method produces clearer chromosome borders than the local adaptive thresholding method. When the input image is highly noisy, the local adaptive thresholding method provides very poor results. On the contrary, the proposed method segments the chromosomes with very small errors. The segmentation errors are marked with red ellipses in Fig. 6b and e.
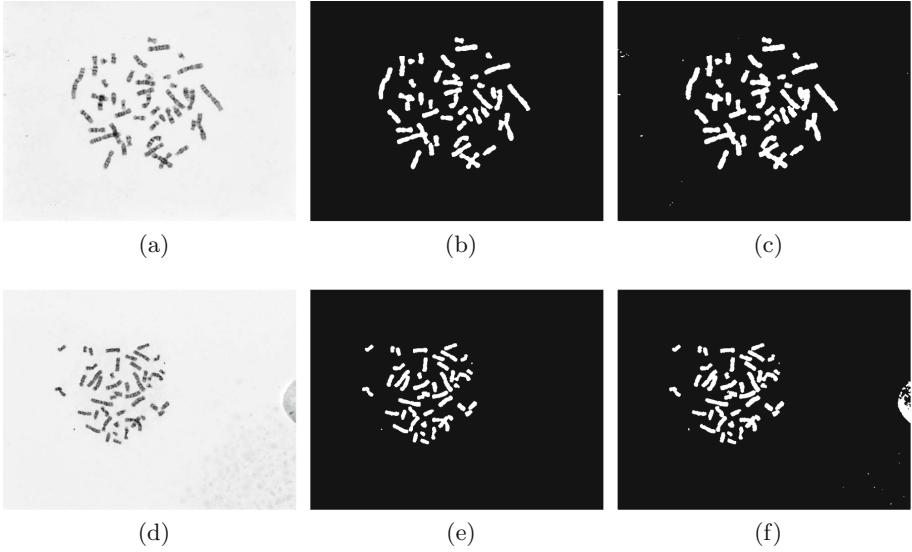
Fig. 5. Comparison of the proposed method and locally adaptive thresholding method. (a) (d) Original images, (b) (e) Predictions of the proposed method, (c) (f) Local adaptive thresholding results
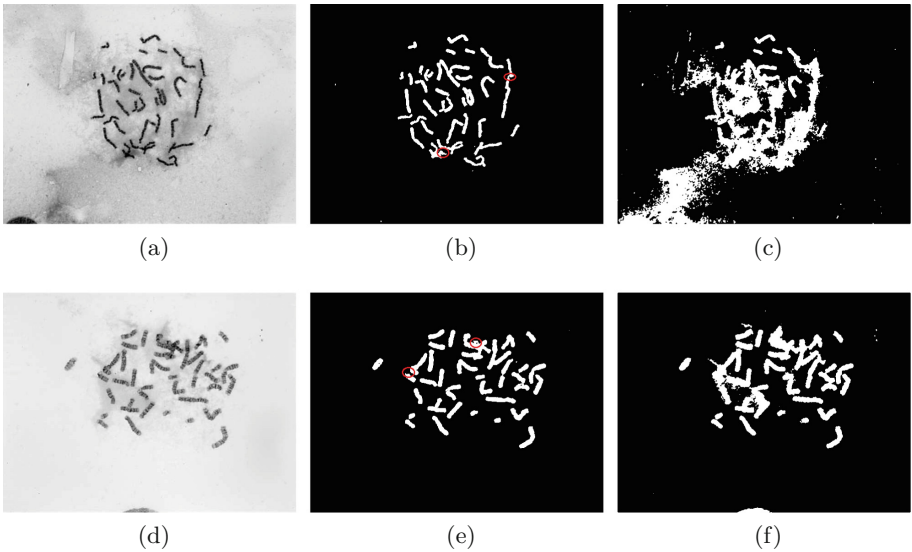

Fig. 6. Comparison of the proposed method and locally adaptive thresholding method. (a) (d) Original highly noisy images, (b) (e) Predictions of the proposed method, (c) (f) Local adaptive thresholding results

4 Conclusion

In this paper, we presented a segmentation method for raw G-band chromosome images using U-net based convolutional neural network. The trained model achieved a 96.97% dice score. The experimental results showed that the convolutional neural network provides satisfying results for the segmentation task. Furthermore, the proposed method segmented highly noisy images with small errors compared to local adaptive thresholding method. The residual errors occured due to lack of image variety in the dataset. For this reason, in the future, we plan to increase the quantity and the diversity of the images in the dataset. Additionally, we intend to increase the segmentation accuracy by using multi-class labels, such as chromosomes, interphase cells, and background masks, instead of using single class mask input.

References

1. Cao, H., Wang, Y.P.: Segmentation of M-FISH images for improved classification of chromosomes with an adaptive fuzzy C-means clustering algorithm. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1442–1445. IEEE (2011)
2. Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C.: The importance of skip connections in biomedical image segmentation. In: Carneiro, G., et al. (eds.) LABELS/DLMIA-2016. LNCS, vol. 10008, pp. 179–187. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46976-8_19
3. Graham, J., Piper, J.: Automatic karyotype analysis. In: Gosden, J.R. (ed.) Chromosome Analysis Protocols. MIMB, pp. 141–185. Springer, Cham (1994). https://doi.org/10.1385/0-89603-289-2:141
4. Grisan, E., Poletti, E., Ruggeri, A.: Automatic segmentation and disentangling of chromosomes in Q-band prometaphase images. IEEE Trans. Inf. Technol. Biomed. **13**(4), 575–581 (2009)
5. Ji, L.: Intelligent splitting in the chromosome domain. Pattern Recognit. **22**(5), 519–532 (1989). https://doi.org/10.1016/0031-3203(89)90021-6
6. Ji, L.: Fully automatic chromosome segmentation. Cytom.: J. Int. Soc. Anal. Cytol. **17**(3), 196–208 (1994)
7. Kayalibay, B., Jensen, G., van der Smagt, P.: CNN-based segmentation of medical imaging data. arXiv preprint arXiv:1701.03056 (2017)
8. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
9. Lerner, B.: Toward a completely automatic neural-network-based human chromosome analysis. IEEE Trans. Syst. Man Cybern. Part B (Cybern.) **28**(4), 544–552 (1998)
10. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
11. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979). https://doi.org/10.1109/TSMC.1979.4310076

12. Poletti, E., Zappelli, F., Ruggeri, A., Grisan, E.: A review of thresholding strategies applied to human chromosome segmentation. Comput. Methods Programs Biomed. **108**(2), 679–688 (2012)

13. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

14. Soumya, D., Arya, V.: Chromosome segmentation using k-means clustering. Int. J. Sci. Eng. Res. **4**(9), 937–940 (2013)

15. Stanley, R.J., Keller, J.M., Gader, P., Caldwell, C.W.: Data-driven homologue matching for chromosome identification. IEEE Trans. Med. Imaging **17**(3), 451–462 (1998)

16. Sugapriyaa, T., Kaviyapriya, P., Gomathi, P.: Segmentation and extraction of chromosomes from G-band metaphase images. Indian J. Sci. Technol. **11**(18) (2018)

17. Yilmaz, I.C., Yang, J., Altinsoy, E., Zhou, L.: An improved segmentation for raw G-band chromosome images. In: The 2018 5th International Conference on Systems and Informatics, pp. 944–950. IEEE (2018)