DeepACE: Automated Chromosome Enumeration in Metaphase Cell Images Using Deep Convolutional Neural Networks

Preprint · February 2019

DOI: 10.13140/RG.2.2.12362.82886

7 authors, including:

Li Xiao
University of California, Irvine
134 PUBLICATIONS **3,521** CITATIONS

SEE PROFILE

DeepACE: Automated Chromosome Enumeration in Metaphase Cell Images Using Deep Convolutional Neural Networks

Li Xiao[1], Chunlong Luo[1], Yufan Luo[1], Tianqi Yu[2], Chan Tian[2], Jie Qiao[2], and Yi Zhao[1]

[1] Key Laboratory of Intelligent Information Processing, Advanced Computer Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
{xiaoli,luochunlong,luoyufan,biozy}@ict.ac.cn

[2] Department of Obstetrics and Gynecology, Third Hospital, Peking University, Beijing, China
1192032990@qq.com,tianchan_cdc@126.com, jie.qiao@263.net

Abstract. Chromosome enumeration is an important but tedious procedure in karyotyping analysis. In this paper, inspired by the successful use of deep convolutional neural networks in natural object detection, we developed a chromosome enumeration algorithm, DeepACE, to automate the enumeration process. To learn unapparent but important information about highly confusing partial chromosomes, we proposed a new Hard Negative Anchors Sampling method to enhance the ability of RPN to extract more prior informations for Fast R-CNN. To alleviate serious occlusion problems, we novelly designed the Truncated Normalized Repulsion Loss to avoid bounding box error when occlusion happens. Our model is trained and tested on a newly collected dataset with thousands of labeled metaphase images from cytogenetic laboratory, and achieves high average precision(AP) value with respect to chromosomes as 98.45 on the testing set. The error rate is 2.5% which also greatly outperforms the previous method.

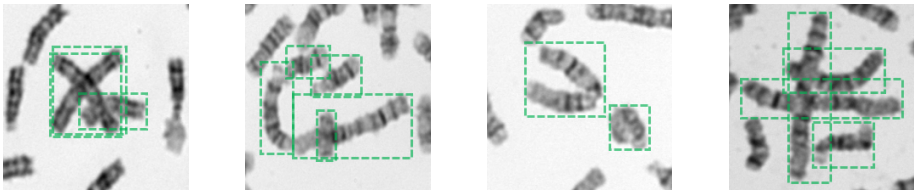
Keywords: Chromosome Enumeration · Object Detection.

1 Introduction

Karyotyping is a cytogenetic experiment method used to help cytologists to analyze the structures and features of chromosomes [12]. In clinical practices, karyotyping generally consists of four stages including chromosome enumeration, segmentation, classification and modification, and finally reporting results. All above processes are based on the metaphase images of cell division generated by microscope camera. Cytologists firstly need to pay attention to the amount of chromosomes for finding out numerical abnormalities of chromosomes that may result in some genetic diseases, such as Down syndrome [11]. Counting chromosomes is performed manually now on at least 20 images per patient and needs 50-100 images more when chromosome mosaicism is explored. Considering

that each human cell normally contains 46 chromosomes, it is tedious and time consuming. Typically, a sophisticated cytologist needs 15 minutes or more to complete chromosome enumeration for one patient, therefore it is an urgent need to develop a computer-aided system for chromosomes enumeration.

However, despite that there are some methods have been developed to solve classification [8, 9, 13, 15, 16, 18] and segmentation [6, 8–10] problems of chromosomes, very few of the researches have tried to develop computer-aided method for chromosome enumeration directly. Gajendran et.al [4] presented a study for chromosome enumeration by combining a variety of pre-processing methods and counting algorithm base on topological analysis, but the error rate was high. The key point of chromosomes enumeration is to find out all the chromosomes on the metaphase image accurately, and the challenges mostly rely on three aspects: firstly, chromosomes in metaphase images usually contain severe occlusion and cross overlapping problem (Fig.1(a)). Secondly, some of the chromosomes present self-similarity in G-band metaphase image (Fig.1(b)): On the one hand, some partial chromosomes are similar to a whole chromosome because they have similar band patterns; On the other hand, two chromosomes are sometimes connected head to head which also brings trouble to identify. But this problem is rarely happened in natural images. Besides, nonrigid chromosomes can be various deformable shapes in images, namely deformation problem (Fig.1(c)). Above three problems are often occurred simultaneously (Fig.1(d)), which usually generate complex chromosome clusters and make it difficult to detect all the chromosome objects accurately.



(a) (b) (c) (d)

Fig. 1. The green boxes are the ground truth bounding boxes of the chromosomes. (a) shows occlusion and cross overlapping problem, two ground truth boxes are very close to each other and hard to identify. (b) shows self-similarity problem, the three connected chromosomes are likely to be classified as one chromosome. (c) shows deformation problem, the deformed chromosomes are likely to be classified as two chromosomes connected to each other. (d) shows complex situation.

In this paper, we proposed a deep learning algorithm to address chromosome enumeration directly on the entire G-band metaphase image. Our model was developed based on the Faster R-CNN [5], one of the state-of-the-art object detection model, which consists of a region proposal network (RPN) to extract features and generate candidate region proposals as well as a Fast R-CNN [5] to finely classify and regress those proposals. We firstly developed a Hard Negative Anchors Sampling method on RPN to learn more information about those

highly confusing partial chromosomes to solve self-similarity problem. Secondly, to alleviate the influence of occlusion and cross overlapping problems within chromosomes, referring to the Repulsion Loss [19], we proposed a Truncated Normalized Repulsion Loss and combine it with the classification and regression loss to jointly optimize the model. Furthermore, to let the Fast R-CNN make full use of spatial information of proposals, we use parallel classification and regression branch to identify chromosomes and regress bounding box of them, in which the fully connected layers are replaced by convolution layers. In the post-processing stage, we use both the traditional NMS and Soft-NMS [2] . Our model is trained and validated on a newly collected dataset with thousands of labeled metaphase images from cytogenetic laboratory, and achieves high average precision value with respect to chromosomes as 98.45 on the testing set. The average error ratio is 2.49% which also greatly outperforms that reported in [4].

2 Method

2.1 Architecture

The architecture of the newly designed network, DeepACE, is shown in Fig.2. The network is developed base on the Faster R-CNN. Firstly a shared feature map is obtained using VGG16 network [17]. Then RPN network is trained with our newly designed Hard Negative Anchors Sampling method and then generates ROIs according to anchor boxes by sliding a small network on the feature map, and the small network consists of classification ($2k$ fc layer) and regression ($4k$ fc layer) branch. The Fast R-CNN uses kernel size of 3 $\times$ 3 with strides equal to 3 and kernel size of 2 $\times$ 2 with strides equal to 1 in the first and second convolution layer of classification branch. And we set the channels to 128, 32 ,8 ,1 with identical kernel size of 3 $\times$ 3 and identical strides of 1 in the four convolution layers of regression branch, aiming at retaining more spatial information. Finally, a newly designed Truncated Normalized Repulsion Loss (TNRL) is computed using the location of predicted boxes and added to the loss function.

Region proposal network The RPN network takes the feature map generated by backbone network as input, then identifies and locates multiple candidate region proposals from anchor boxes. The anchor boxes are set manually as reference boxes. To properly cover size distribution of real objects, we firstly compute the height and width distribution of the ground truth bounding box among all chromosomes in the training and validation set, and then set anchors of 4 scales $(32^2, 64^2, 96^2, 128^2)$ and 9 aspect ratios $(\frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5)$, which cover most of the chromosome sizes. Unlike natural objects, chromosomes have self-similarity problem which may result in that a huge proportion of ROIs containing only part of a chromosome which may puzzle the Fast R-CNN. To solve this problem, we investigated a new sampling method which collect these hard samples emphatically when training the RPN. The details are shown in Section 2.2.

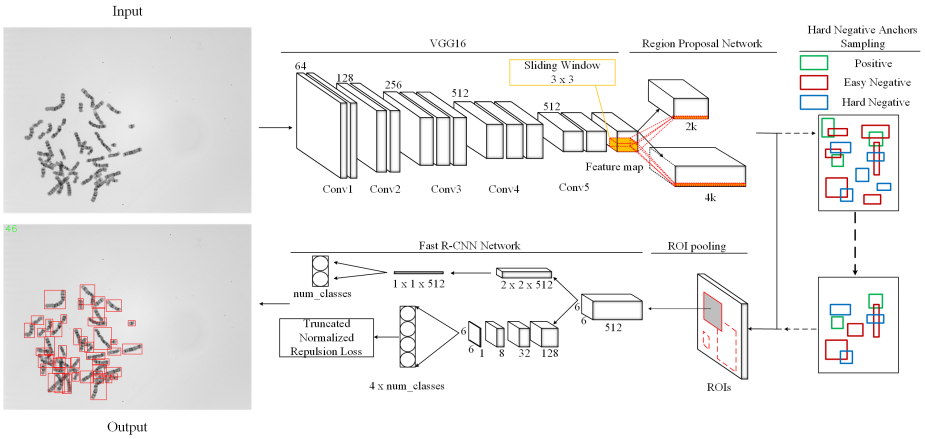


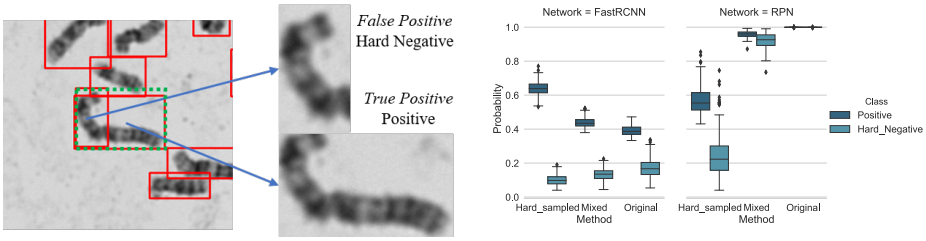
Fig. 2. General architecture of our DeepACE. The cuboids in the figure represent feature maps with different sizes and channels. The raw karyotyping image is sent into the network. The chromosomes are labeled with rectangle boxes and the total number of chromosomes are computed in the output image.

Fast R-CNN network We changed the structure of the Fast R-CNN network to increase detection accuracy. Originally, Fast R-CNN is powerful enough to complete tasks of finer subclass classification of natural objects and bounding box regression with simple fully connected layers. However, fully connected layers in Fast R-CNN may lose important spatial information and shared parameters may create conflicts between classification and regression tasks. Therefore, to extract spatial structure information in order to enhance characters of each channel in the classification branch, we firstly use convolution layers with identical channels and only fuse those channels at the end .The regression branch needs to assemble high-level semantic components to outline a chromosome for more precise location, we replaced the fully connected layer by several convolution layers with identical feature map sizes.

Loss Function and Post-processing Object occlusion and deformation are popular topics in the area of computer vision. Some of the chromosomes in karyotyping images are clustered together and bring more complicate and severe occlusion problem, which makes the counting very difficult. We introduced repulsion-based loss to improve the performance. Compared with pedestrians detection problem where the repulsion loss method is firstly introduced [19], overlapping of chromosomes in the karyotyping image is more frequent and severe. We novelly developed the Truncated Normalized Repulsion Loss method which focus more on the drifting part of the object, the details are described in Section 2.3. Besides, in the post-processing stage, when ground truths shelter each other, a ground truth may not be detected since related bounding boxes are suppressed by the predicted bounding box of another ground truth when applying traditional NMS. Therefore we introduced the Soft-NMS method in our model to avoid such abrupt deletion.

2.2 Hard Negative Anchors Sampling

In the original Faster R-CNN, the two stages, RPN and Fast R-CNN, coordinate with each other through shared feature maps. The RPN roughly filters out the backgrounds and regresses candidate region proposals (RoIs), and then Fast R-CNN fine classify and regress them. However, the self-similarity problem of chromosome images and limitation of Fast R-CNN make it difficult to train the original network and achieve a satisfied result. Specifically, unlike natural objects, as shown in Fig.3(a), chromosomes usually have various length and blurry banding patterns, and the banding patterns are similar in vision between different chromosomes, which make challenges for the network to discriminate partial and whole chromosomes, namely self-similarity problem. Here we call these partial chromosomes as hard negative proposals. Some process of the Fast R-CNN also have the risk of losing important information. For example, cropping features according to coordinates of candidate proposals may lose some surrounding information around proposals and the RoI Pooling layer will also irreversibly remove some information by the downsampling process. As shown in Fig.3(b), most of the positives and hard negative proposals are identified as background (probability < 0.5) by original Fast R-CNN, which means the Fast R-CNN is not powerful enough to discriminate positives from hard negative proposals. Therefore, considering that the RPN does not have such cropping and down sampling process to lose information, we try to learn the information from hard negative objects through the RPN to support Fast R-CNN.



(a) (b)

Fig. 3. (a) shows an example of similarity between whole chromosomes and partial chromosomes, green and red boxes represent ground truths and predicted bounding boxes, respectively.(b) shows foreground probability distribution of positives and hard negatives proposals/anchors predicted by the Fast R-CNN(left) and the RPN(right), respectively, using different methods.

Originally, The RPN is only trained by foreground (positive, IoU ≥ 0.7) and background (negative, IoU < 0.3) anchors. However, those anchors that have the IoU in the interval $[0.3, 0.7)$ are ignored, which exactly cover most of the partial chromosomes. In our task, for clarity, we regard these anchors as hard negative anchors and the original negative anchors with an IoU < 0.3 are named as easy negative anchors. A direct but unsuccessful trail is to mix up hard negative anchors with easy negative anchors in the sampling procedure

of RPN. Nevertheless, RPN still suffers from severe inter-class imbalance and intra-class imbalance. On the one hand, the inter-class imbalance is happened between positive and negative anchors ($\approx 1:2000$). On the other hand, it is discovered that intra-class imbalance happened where the number of hard negative anchors is considerably small comparing to the huge amount of easy negative anchors ($\approx 1:25$). In this situation, easy negative anchors dominate the negative anchors and few of hard negative anchors are trained by RPN if we only uniformly sample easy and hard negative anchors together. As shown in Fig.3(b), the mixed method still makes the RPN predict similar foreground probability of positives and hard negatives, which results in only small improvement of Fast R-CNN.

A new Hard Negative Anchors Sampling method is then proposed, which is inspired by stratified sampling. We divided all anchors into positive, hard negative and easy negative according to IoU overlap with ground truth box. To balance the amount of the three classes of anchors in our experiments, we used mini-batches of size R=512 for training RPN and took 25% of the anchors from positive anchors. The half of remaining are uniformly sampled from hard negatives, and the rest are sampled from easy negatives. Finally, positive anchors are labeled with a foreground object class, both hard and easy negative anchors are labeled as background, the loss function is the same as the original RPN. In this way, the function of RPN is extended to help Fast R-CNN to better classify hard negative proposals. During training process, feature maps generated by VGG16 backbone network are enhanced by hard negative anchor information, and Fast R-CNN is improved by these features that contain strong prior knowledge of hard negatives. As shown in Fig.3(b), positives and hard negatives are correctly identified as foreground by the RPN using the new sampling (Hard_sampled) method, which leads to a good performance of the Fast R-CNN (most of the positives are identified as chromosome with probability > 0.5 and most of the hard negative are identified as background with probability < 0.5).

2.3 Truncated Normalized repulsion loss

Occlusion is a factor that harms the performance of object detectors heavily. Especially, chromosomes often gather together with various body shapes and bring severe intra-class occlusion and cross overlapping simultaneously, in which case network is unable to localize the chromosomes correctly. To solve the problem, we proposed a new loss function aiming at alleviating the influence caused by severe intra-class occlusion in karyotyping images, The new loss function is shown as below:

$$L_{total} = L_{attraction} + \alpha \cdot L_{repulsion} \tag{1}$$

The function consists of two terms: the attraction term $L_{attraction}$ and the repulsion term $L_{repulsion}$. The first term is the same as original loss of Faster R-CNN and the second one is proposed to prevent predicted boxes from shifting to adjacent objects when occlusion of ground truths occurs.

Assume $\mathbb{P}_+$ is the set of positive proposals produced by RPN network and $P \in \mathbb{P}_+$, B^P is the predicted box of P, G^P is the designated target of P defined as the ground truth that has the highest IOU with P, R^P is the repulsion ground truth object of P defined as the ground truth that has the highest IOU with P except G^P. The repulsion loss term is firstly proposed in Wang et.al [19] as:

$$L_{repulsion} = \frac{\sum_{P \in \mathbb{P}_+} Smooth_{ln}(IOG(B^P, R^P))}{|\mathbb{P}_+|} \tag{2}$$

Where $IOG(A, B) \triangleq \frac{area(A \cap B)}{area(B)}$, and $smooth_{ln}$ is a smooth function.

However, as shown in Fig.4(a), if the overlap of two ground truth bounding boxes G^P and R^P is severe, the huge intersection part will dominate the repulsion loss even when B^P matches G^P well. In this situation the repulsion loss will not be sensitive to shifting. Also the repulsion loss will always be equal to 1 if the ground truth G^P contains R^P as shown in Fig.4(a). These phenomena limit the ability of the repulsion loss function to accurately localize each chromosome. Furthermore, Faster R-CNN may be hard to converge when G^P and R^P are overlapped heavily because of the large loss value. Therefore we proposed a novel repulsion loss function, called Truncated Normalized Repulsion Loss (TNRL) as:

$$L_{TNrepulsion} = \frac{\sum_{P \in \mathbb{P}_+} Smooth_{ln}(max(\frac{IOG(B^P, R^P) - IOG(G^P, R^P)}{1 - IOG(G^P, R^P)}, 0))}{|\mathbb{P}_+|} \tag{3}$$

Where,

$$\frac{IOG(B^P, R^P) - IOG(G^P, R^P)}{1 - IOG(G^P, R^P)} = \frac{area(B^P \cap R^P) - area(G^P \cap R^P)}{area(R^P) - area(G^P \cap R^P)} \tag{4}$$

The comparison of TNRL and original repulsion loss is depicted in the right equations of Fig.4(a). Similar to the original repulsion loss, the novel loss function can only be decreased by decreasing $area(B^P \cap R^P)$. But the loss depends on the shifting of the prediction box only and is not affected by the overlap of G^P and R^P, which means the loss value changes more directly according to the severity of shifting. And also, it has an upper bound equal to one if B^P is coincident with R^P; and a lower bound equal to zero when it's coincident with G^P, which means the range of the loss value is greater than the original repulsion loss depending on the severity of the prediction error, especially in the severe occlusion situation as shown in Fig.4(c[3]). Furthermore, when G^P contains R^P, the new loss is equal to zero rather than confused by a large value when applying the original repulsion loss. Therefore, our new loss function is a more accurate representation of the shifting errors (as shown in Fig.4(b)). The experiments in Section 3.3 proved that the novel loss function regresses proposals more precisely.

[3] For simplify, the diagram is depicted when two ground truths are equal in size and the predicted box is shifted from one to another straightly.

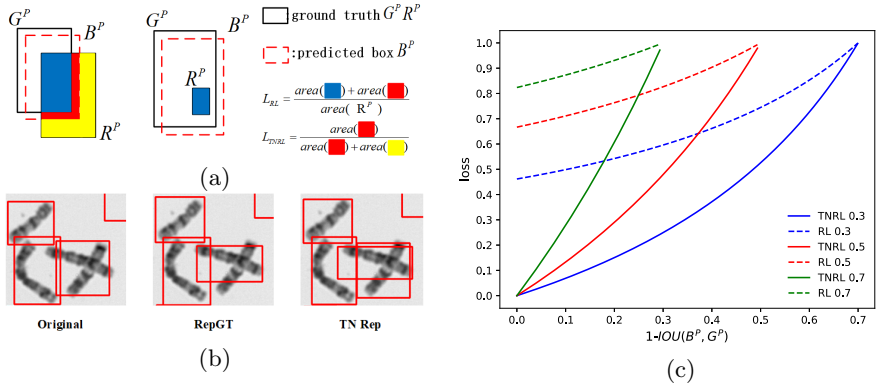


Fig. 4. (a) demonstrates the difference between TNRep Loss and RepGT Loss. (b) is the predicted results of same instance using different loss function when severe occlusion happens. (c) is the comparison of TNRep Loss and RepGT loss with different shifting error, at different occlusion situations($IOU(G^P, R^P)$ values are set as 0.3,0.5,0.7).

3 Experiments Results

3.1 Data Set and Evaluation Metrics

To validate the proposed method on the entire metaphase image, we collected 1375 metaphase images with resolution of 1600×1200 containing 63026 objects from Peking University Third Hospital. Each whole metaphase image was labeled by a cytologist with a rectangle bounding box associate with each chromosome, and then verified by another cytologist. All images were randomly separated into a training set of 1000 images, a validation set of 175 images and a testing set of 200 images.

Since chromosome enumeration problem is closely related to clinic, besides the traditional object detection criterions such as average precision (AP), we introduced the Whole Correct Ratio (WCR) and Average Error Ratio (AER) as the metrics to evaluate the ensemble performance, which cytologists may more care about. In these experiments, we called images that all chromosomes are correctly detected as all right images, and WCR is defined as the percentage of all right images in the whole testing set. Furthermore, given a matching threshold value, we regarded a predicted bounding box as a false positive if it does not have an IoU greater than the threshold with any ground truth or it has the max IoU with a ground truth that has already been detected. A ground truth that is not detected by any bounding box is regarded as a false negative. The AER is defined as the value computed by the sum of the false positives and false negatives divided by the number of ground truths.

3.2 Implementation Details

We conducted augmentations for training set, and only vertical and horizontal flippings were used to augment data to reduce overfitting. During the preprocessing step, we assigned 255 to pixels whose values are larger than 200. In the

training process, we initialized the backbone network with pretrained VGG16 and the rest layers were randomly initialized as in [14]. In the Truncated Normalized Repulsion Loss, the hyper parameter was set as 1.0 and the weight of repulsion term α was set as 10. In the testing process, we used Soft-NMS after RPN and both NMS and Soft-NMS after Fast R-CNN. The thresholds of both traditional NMS and soft-NMS were set as 0.5. All above hyper parameters were determined via grid search on the validation set. Network was implemented on TensorFlow framework [1]. We trained the network for $100k$ iterations with the initialized learning rate set to 0.001 and decayed by a factor of 10 at $40k$ and $80k$ iterations. We used Stochastic Gradient Descent (SGD) to optimize our network on a Nvidia Titan Xp GPU with momentum= 0.9.

3.3 Results

Detection Results and Comparison. We firstly verified the effectiveness of our proposed method by comparing with the chromosome enumeration method proposed in [4], which was based on digital image analysis and evaluated by Cluster-Based Error criterion. The criterion only concerns the error of the chromosome number caused by cutting entire chromosome incorrectly or connecting individual chromosomes incorrectly. So, the criterion is much looser than AER criterion in this paper. Nevertheless, as shown in Table 1, our method still greatly outperformed the previous method [4] and achieved an AER of 2.49%.

Table 1. The comparison of chromosomes counting methods.

Method	Average Error Ratio
Gajendran et al., ICIP'04 [4]	6.4% (average Cluster-Based Error)
DeepACE + NMS + Soft-NMS	2.49%

Ablation Study. To further investigate the contribution of each key component in our framework, extensive ablation experiments were performed on the testing set. The true positive IoU matching threshold is set as 0.5 which is used in the popular PASCAL VOC [3] competition. As shown in Table 2, we trained DeepACE using different components and the first row represents baseline of our method. We find that HNAS performs a huge improvement on all the metics. Especially, it improves the performance of WCR by about 40%. This improvement demonstrates that HNAS indeed enhances hard negative features to support the Fast R-CNN. On the other hand, though occlusion problem is not the main problem at low matching threshold, TNRL can also improve ensemble performance by decreasing AER and increasing AP. Combined together, our proposed DeepACE achieves 98.45 of AP, 48% of WCR and 2.49% of AER, which greatly outperforms the basic network. The comparison of predicted results on the example cases is illustrated in Fig.5. Both the classification probabilities and bounding box locations can be significantly improved by applying the HNAS and TNRL.

Table 2. Ablation study about each component on the testing set. HNAS represents Hard Negative Anchors Sampling; TNRL represents Truncated Normalized Repulsion Loss.

	HNAS	TNRL	AP	WCR	AER
DeepACE			95.12	7%	8.15%
		√	95.25	9.5%	8.05%
	√		98.43	**49%**	2.5%
	√	√	**98.45**	48%	**2.49%**

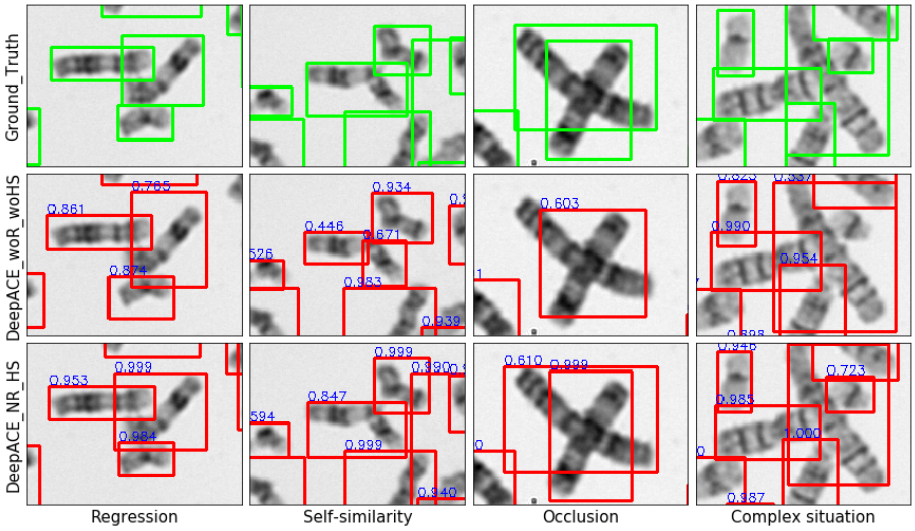


Fig. 5. Grid of examples illustrates the performance of the classification probability and predicted bounding box when applying HNAS and TNRL under different situations. (top row) Example cases labeled with ground truth (Ground_Truth). (middle/bottom row) predicted results of DeepACE without applying HNAS and TNRL (DeepACE_woR_woHS) and DeepACE with HNAS and TNRL (DeepACE_NR_HS). Different columns represent example cases under different situations, including general examples (Regression), self-similarity, occlusion and complex situations. All the example cases show that HNAS and TNRL can significantly improve the performance and make the predicted results closer to the ground truth.

Repulsion Loss and Network. To further comprehensively evaluate the effectiveness of the Truncated Normalized Repulsion Loss and DeepACE structure, we refer to COCO [7] evaluation criterion to design a series of experiments for measuring the performance of different types of Repulsion Loss in different models at eight different matching thresholds. Low matching thresholds represent that more predicted bounding box will be identified as true positive and higher thresholds means more prudent acceptance for predicted objects. Furthermore, we used Faster R-CNN as baseline model for the comparison. As shown in Table 3, by comparing different loss function and different detection model, we observed two interesting phenomenons. Firstly, both RL and TNRL can improve the performance in most of the situations, which demonstrates the effectiveness of the repulsion-based loss functions. Further more, the mean of results shows that our proposed TNRL achieves general improvement on both WCR and AER, which proves that TNRL outperforms L and RL. Secondly, DeepACE reports better results than basic Faster R-CNN. Especially for WCR, DeepACE constantly achieves better performances than basic Faster R-CNN on all the matching thresholds ranging from 0.5 to 0.85. For AER, performance of the DeepACE is boosted by a remarkably large margin under conditions of high matching thresholds. These results demonstrate that the specific design of DeepACE in Section 2.1 can keep more spatial informations than Faster R-CNN and thus make the predicted bounding box more accurate. In short, both TNRL and DeepACE can improve the performance of chromosome enumeration.

Table 3. Whole Correct Ratio and Average Error Ratio on different true positive IoU thresholds. L represents original loss; RL represents original Repulsion Loss; TNRL represents our proposed Truncated Normalized Repulsion Loss.

	DeepACE+HNAS						Faster R-CNN+HNAS					
	WCR			AER			WCR			AER		
	L	RL	TRNL	L	RL	TRNL	L	RL	TRNL	L	RL	TRNL
0.5	**49**	48	48	2.5	**2.44**	2.49	46	46.5	**47**	2.48	**2.42**	2.57
0.55	47.5	44.5	**47.5**	**2.78**	2.79	2.81	44	45	**47**	2.83	**2.68**	2.79
0.6	44	41	**45.5**	3.26	**3.2**	3.25	38	41	**45**	3.33	3.17	**3.03**
0.65	35.5	33.5	**40.5**	4.29	4.16	**4.04**	34.5	37.5	**38.5**	4.29	4.13	**3.9**
0.7	23.5	23	**27.5**	6.84	6.67	**5.96**	19	**27**	26.5	6.87	6.18	**5.88**
0.75	9	11.5	**13.5**	11.95	10.96	**10.09**	6.5	**11**	9.5	11.81	**10.89**	11.09
0.8	-	-	-	24.2	22.9	**20.39**	-	-	-	23.65	**22.78**	23.02
0.85	-	-	-	51.45	49.47	**45.66**	-	-	-	**50.42**	51.39	51.12
Mean	34.8	33.6	**37.08**	13.41	12.82	**11.84**	31.3	34.7	**35.58**	13.21	12.96	**12.93**

4 Conclusion

In this paper, we developed a chromosome enumeration algorithm, DeepACE, to automate chromosome enumeration process. To learn unapparent but important information about hard negative objects, we proposed a new Hard Negative Anchors Sampling method to enhance the ability of RPN to extract more feature

informations. To alleviate serious occlusion problem, we novelly designed the Truncated Normalized Repulsion Loss to avoid bounding box error when occlusion happens. To localize the bounding box more precisely, we replaced the fully connected layers of Fast R-CNN with convolution layers to keep more spatial informations. Experiments were performed on the collected clinical metaphase images and demonstrated that DeepACE with HNAS and TNRL can achieve solid and consistent improvements at all metrics. We believe our model is a powerful clinical tool to help cytologists make their counting process more fast and efficient.

References

1. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al.: Tensorflow: a system for large-scale machine learning. In: OSDI. vol. 16, pp. 265–283 (2016)
2. Bodla, N., Singh, B., Chellappa, R., Davis, L.S.: Soft-nms—improving object detection with one line of code. In: Computer Vision (ICCV), 2017 IEEE International Conference on. pp. 5562–5570. IEEE (2017)
3. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International Journal of Computer Vision **88**(2), 303–338 (Jun 2010)
4. Gajendran, V., Rodríguez, J.J.: Chromosome counting via digital image analysis. In: Image Processing, 2004. ICIP'04. 2004 International Conference on. vol. 5, pp. 2929–2932. IEEE (2004)
5. Girshick, R.: Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision. pp. 1440–1448 (2015)
6. Hu, R.L., Karnowski, J., Fadely, R., Pommier, J.P.: Image segmentation to distinguish between overlapping human chromosomes. arXiv preprint arXiv:1712.07639 (2017)
7. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: European conference on computer vision. pp. 740–755. Springer (2014)
8. Munot, M.V.: Development of computerized systems for automated chromosome analysis: Current status and future prospects. International Journal of Advanced Research in Computer Science **9**(1) (2018)
9. Nair, R.M., Remya, R., Sabeena, K.: Karyotyping techniques of chromosomes: a survey. Int J Comput Trends Technol **22**(1) (2015)
10. Pardo, E., Morgado, J.M.T., Malpica, N.: Semantic segmentation of mfish images using convolutional networks. Cytometry Part A (2018)
11. Patterson, D.: Molecular genetic analysis of down syndrome. Human genetics **126**(1), 195–214 (2009)
12. Piper, J.: Automated cytogenetics in the study of mutagenesis and cancer. In: Advances in Mutagenesis Research, pp. 127–153. Springer (1990)
13. Qin, Y., Song, N., Zheng, H., Huang, X., Yang, J., Zhu, Y.M., Yang, G.Z.: Varifocalnet: A chromosome classification approach using deep convolutional networks. arXiv preprint arXiv:1810.05943 (2018)
14. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. pp. 91–99 (2015)

15. Sharma, M., Saha, O., Sriraman, A., Hebbalaguppe, R., Vig, L., Karande, S.: Crowdsourcing for chromosome segmentation and deep classification. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 786–793 (July 2017)
16. Sharma, M., Swati, Vig, L.: Automatic chromosome classification using deep attention based sequence learning of chromosome bands. In: 2018 International Joint Conference on Neural Networks (IJCNN). pp. 1–8 (July 2018)
17. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
18. Swati, Gupta, G., Yadav, M., Sharma, M., Vig, L.: Siamese networks for chromosome classification. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW). pp. 72–81 (Oct 2017)
19. Wang, X., Xiao, T., Jiang, Y., Shao, S., Sun, J., Shen, C.: Repulsion loss: Detecting pedestrians in a crowd. arXiv preprint arXiv:1711.07752 (2017)