This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics

CIR-Net: Automatic Classification of Human Chromosome based on Inception-ResNet Architecture

Chengchuang Lin, Gansen Zhao*, Zhirong Yang*, Aihua Yin, Xinming Wang, Li Guo, Hanbiao Chen, Zhaohui Ma, Lei Zhao, Haoyu Luo, Tianxing Wang, Bichao Ding, Xiongwen Pang, Qiren Chen

Abstract—**Background:**In medicine, karyotyping chromosomes is important for medical diagnostics, drug development, and biomedical research. Unfortunately, chromosome karyotyping is usually done by skilled cytologists manually, which requires experience, domain expertise, and considerable manual efforts. Therefore, automating the karyotyping process is a significant and meaningful task.
Method: This paper focuses on chromosome classification because it is critical for chromosome karyotyping.In recent years, deep learning-based methods are the most promising methods for solving the tasks of chromosome classification. Although the deep learning-based *Inception* architecture has yielded state-of-the-art performance in the 2015 ILSVRC challenge, it has not been used in chromosome classification tasks so far. Therefore, we develop an automatic chromosome classification approach named *CIR-Net* based on *Inception-ResNet* which is an optimized version of *Inception*. However, the classification performance of origin *Inception-ResNet* on the insufficient chromosome dataset still has a lot of capacity for improvement. Further, we propose a simple but effective augmentation method called *CDA* for improving the performance of *CIR-Net*.
Results: The experimental results show that our proposed method achieves 95.98% classification accuracy on the clinical G-band chromosome dataset whose training dataset is insufficient.Moreover, the proposed augmentation method *CDA* improves more than 8.5% (from 87.46% to 95.98%) classification accuracy comparing to other methods.In this paper, the experimental results demonstrate that our proposed method is recent the most effective solution for solving clinical chromosome classification problems in chromosome auto-karyotyping on the condition of the insufficient training dataset. Code and Dataset are available at https://github.com/CloudDataLab/CIR-Net.

Index Terms—Chromosome Classification,Inception-ResNet,Biomedical Image Analysis,Chromosome Image Augmentation.

1 Introduction

HUMAN chromosomes contain human genetic information, which are commonly used for analyzing human genetic diseases. In general, there are 23 pairs of chromosomes in a healthy human body, including 22 pairs autosomes and a pair of sex chromosomes (X and Y chromosome in male cells and double X in female cells) [1]. Karyotype analysis, illustrated by Figure 1, is a fundamental approach for clinical cytogeneticists to diagnose human chromosomes genetic diseases, which is generated by arranging these chromosomes after extracting them from the metaphase chromosome images [2]. For cytogeneticists, karyotyping is laborious work, many researchers have dedicated to auto-karyotyping using computation techniques [3], [4], [5], [6], [7], [8] for years. This paper focuses on chromosome classification because it is a very significant but a labor-intensive stage in auto-karyotyping.

Although many researches [1], [8], [9], [10] have made some contribution for chromosome classification tasks, classifying chromosomes accurately and robustly in the clinical application on the condition of the insufficient labeled dataset is still a challenging task for the following reasons:

- Rich deformations of chromosome shape. Chromosomes in stained cell microphotographs have non-rigid intrinsic nature, so it is very common that chromosomes of the same type have completely different shapes and orientations.
- Difficulty to collect a large amount of labeled data. As chromosome images are highly correlative to the individual privacy of patients, it is very difficult for researchers to correct enough data from medical institutions to train their classifiers.

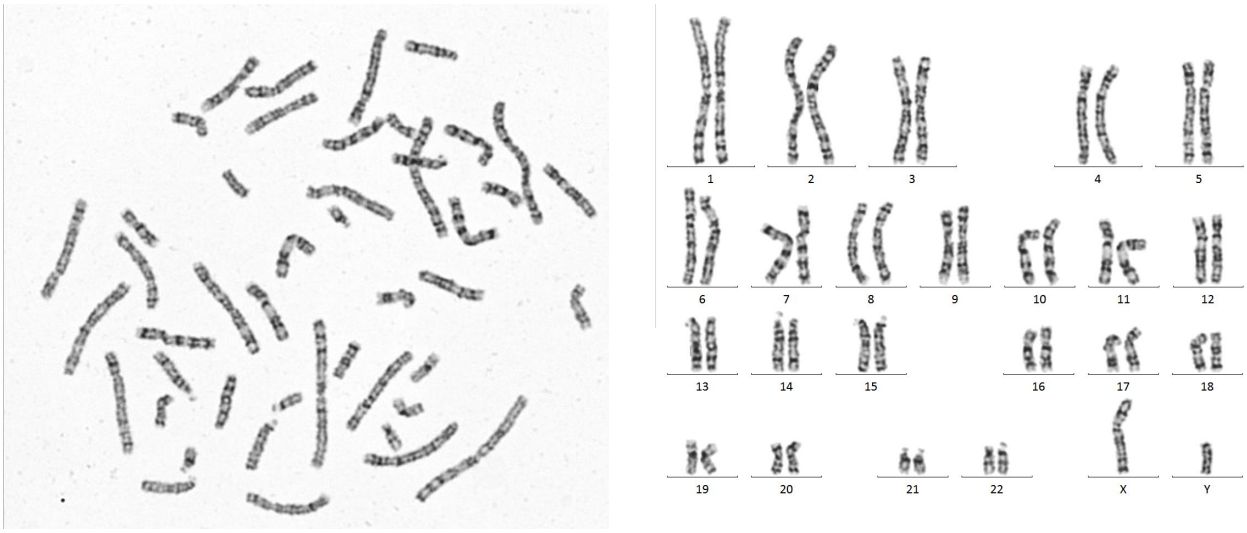
To solve the above problems, we propose an end-to-end classification approach named *CIR-Net* based on Inception-ResNet [11] architecture and design a data augmentation algorithm *CDA* using affine transformation for improving the performance of the classifier.

- *C. Lin, G. Zhao, L. Zhao, H. Luo, T. Wang, B. Ding, X. Pang, and Q. Chen are with School of Computer Science, South China Normal University, Guangzhou, China.*
- *Z. Yang is both with Norwegian University of Science and Technology, Norway and Aalto University, Finland.*
- *X. Wang is with Lakala Group, Building D1, Zhongguancun Yihao, Beiqing Road, Beijing, China. E-mail: wangxinming@lakala.com.*
- *A. Yin, L. Guo, and H. Chen are with Medical Genetic Centre and Maternal and Children Metabolic-Genetic Key Laboratory, Guangdong Women and Children Hospital, Guangzhou, China.*
- *Z. Ma is both with School of Computer Science, South China Normal University and School of Information Science and Technology, Guangdong University of Foreign Studies, Guangzhou, China.*
- *∗ G. Zhao and Z. Yang are corresponding authors. Email: gzhao@m.scnu.edu.cn, zhirong.yang@ntnu.no.*

Manuscript received August 20, 2019; revised November 25, 2019.

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics



(a) (b)

Fig. 1. (a)stained cell microphotograph G-band chromosome image, (b)chromosome karyotype

Additionally, to improve the clinical applicability, we design an image adaptive interface(*IAI*) module for converting an arbitrary size of the input image into the target input size of Inception-ResNet without any preprocessing.

At last, we quantitatively analyze the performance of our method compared with others' researches evaluating by general metrics ($precision$, $recall$, $f1$). Moreover, we qualitatively analyze discriminating the capacity of our model using t-SNE [12].

According to mentioned explorations, this paper makes the following contributions:

- We propose *CIR-Net* for automatically classifying chromosomes in an end-to-end pipeline.
- We explore an effective image augmentation algorithm *CDA*. This algorithm has two benefits: it not only can enlarge the training dataset and improve the accuracy and robustness of classifier but also can eliminate the directional features of chromosomes, which allow us to classify chromosomes without special preprocess (e.g., rotating, straightening) in the test or clinical application stage.
- We design the *IAI* module to receive an arbitrary input size image without preprocessing, which is greatly improving the application capability of the system while others' solutions can not do that.

The rest of this paper is organized as follows: Section 2 will review previous works on chromosome auto-karyotyping and classification problems. Section 3 is about to describe the chromosome classification problem in mathematics, chromosome images augmentation and deep neural network structure. In Section 4, we will give experimental performance results of the *CIR-Net* compared to other methods. We will make a discussion and conclusion in Section 5.

2 Related Work

Chromosome karyotyping is a crucial task for genetic disease detection, which is also a hot spot in recent years.Traditional methods for chromosome classification are generally performed manually [13], which is time-consuming and inefficient. With the development of computational methods, the technologies for automatic chromosome karyotyping system have come into being, mainly including two parts: chromosome segmentation and classification. This paper focused on the classification problem under the assumption that chromosomes are well segmented.

Earlier chromosome classification methods include artificial neural network methods [14], [15] and probabilistic artificial neural network methods [16], [17], [18]. The former methods are principally based on MLP (Multi-Layer Perceptron), which have complex feature selection processes. The latter methods have lower accuracy while they offer shorter training time compared to the training stage of former methods due to backpropagation.

Recently, several deep learning-based approaches have been employed in chromosome classification, such as Siamese Networks [1], Attention Based Sequence Learning [8], vanilla Convolutional Neural Network (Vanilla-CNN) [9] and Varifocal-Net [10].

Jindal et al. [1] proposed a chromosome classification method using deep learning technique based on Siamese Networks [19]. In Jindal's method, they firstly straighten all chromosomes using SMAC (Straightening via Medial Axis and Crowdsourcing) and SPV (Straightening via Projection Vectors) methods, and then they design and pre-train Siamese network using Base-CNN on paired chromosomes for getting parameters W in Base-CNN. Lastly, they design the final classifier using Base-CNN with pre-training parameters W from Siamese Network and adding the MLP to it. The experimental results yield 84.6% classification accuracy in their private G-band dataset.

Monika et al. [8] proposed an automatic chromosome classification method using a deep attention mechanism for learning chromosome band features. They firstly propose Residual Convolutional Recurrent Attention Neural Network (Res-CRANN) which exploits the property of chromosome band sequences. After that, they feed these

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics

chromosome sequences into the Recurrent Neural Network (RNN). Subsequently, an attention mechanism is applied to the top of RNN. At last, the attention module outputs the sequences which are further classified into 24 labels. Monika et al. evaluate their method on the public available Q-band chromosome dataset which yields 91.94% classification accuracy.

Wenbo et al. [9] proposed a Vanilla-CNN method for chromosome classification using deep learning. The authors collect karyotypes from a local company and extract chromosomes from each karyotype. And then, they resize each chromosome image into 142 by 282 pixels with the vertical direction. At last, they feed all chromosome images into Vanilla-CNN (two convolution blocks, one flatten layer, and one dense Layer). In this literature, the authors claimed that this method has yielded 92.5% classification accuracy on the private G-band chromosome dataset which contains 10304 chromosome images.

Yulei et al. [10] proposed Varifocal-Net for chromosome classification using deep learning. This approach consists of one global-scale network (G-Net) and one local-scale network (L-Net). Firstly, the authors extract global features and detect finer local regions via G-Net. Subsequently, they zoom into local parts and extracts local features using a varifocal mechanism. At last, residual and multi-task learning strategies are utilized to promote high-level feature extraction. Evaluation results from 1909 karyotyping cases(87814 G-band chromosome images) showed that Varifocal-Net achieved the highest accuracy of 99.2%.

Although the above recent methods seem to have solved the chromosome classification problem, there are still some limitations in these methods. Firstly, previous methods usually use straightening as preprocessing operation, which means that in the clinical application we need to straighten the target chromosomes in strict accordance with authors do. Secondly, these methods [8], [9], [10] need a large amount of labeled chromosome dataset to guarantee their good performance. Lastly but not least, these methods sound good enough for solving the classification tasks. However, some of them [8], [10] are too complex for clinical researchers to reproduce caused by too much detail information is missing in their papers.

Motivating by these problems in clinical application, we propose an end-to-end classifier fine-tuned from Inception-ResNet. Additionally, to improve the robustness of the classifier, we provide an effective algorithm termed *CDA* for image augmentation. Moreover, we design an *IAI* module for accepting an arbitrary shape of chromosome images without extra preprocessing in clinical application.

3 Proposed Method

The proposed method includes four parts, the first one is chromosome classification problem definition and formal description while the second part details data augmentation. The third part is about the Image Adaptive Interface (*IAI*). The last part is the description of *CIR-Net* architecture.

3.1 Problem Description

As chromosome image is the grayscale image which composes H rows and W columns where H and W mean the spatial height and spatial width. So we can use two-dimensional tensor x to denote a chromosome image, use $x[r][c]$ to denote the value of pixel in r-th row, c-th column of x. After that we can use $X = \{x_1, x_2, \ldots, x_n\}$ to denote a chromosome image set which contains n chromosome images.

Correspondently, we can use a 24-dimensional tensor y to denote the type of chromosome image. More precisely, we use y_i to denote the type of chromosome x_i, use $Y = \{y_1, y_2, \ldots, y_n\}$ to denote the chromosome label set.

The chromosome classification problem can be formalized as finding a good enough classifier Φ to represent the relationship between X and Y. For convenience, we call the output of the classifier $\hat{Y}$ as the predict value set, the actual chromosome label set Y as the ground truth set. The expression of Φ is shown in *Equation 1*.

$$\forall x_i \in X, \exists \hat{y_i} = \Phi(x_i) \in \hat{Y} \tag{1}$$

We define a special function L named loss function to evaluate the degree of the classifier Φ deviated from the ground truth set Y,where L can be formalized as Equation 2.

$$L = -\sum y_i \log(\hat{y_i}) = -\sum y_i \log(\Phi(x_i)) \tag{2}$$

At last, the problem of finding a good enough classifier Φ transforms to the problem of minimization the loss function L of classifier Φ.

3.2 Data Augmentation

Data augmentation aims at generating additional and more diversified data samples through certain transformations conducted upon original data [20]. There are many data augmentation techniques in the clinical medical image classification task, such as *Flips*, *Gaussian Noise* and *Jittering* et al. [21]

As chromosomes have rich deformations (e.g. random direction and position) in the images, so the ideal augmentation techniques should be able to add certain samples that can tell classifier Φ to eliminate the features of random direction and position. Therefore, we introduce the *Affine Transformation* technique to augment the chromosome dataset for dropping out these features.

As we use two-dimensional tensor x_i to denote the i-th chromosome image in above content, so we use $x_{(i,\theta)}$ to denote the image augmented from x_i. The augmentation process can be described as Equation 3, where $A(\theta)$ is a rotation matrix which can be formalized as Equation 4 while b is an offset vector which can be formalized as Equation 5.

$$x_{(i,\theta)} = A(\theta)x_i + b \tag{3}$$

$$A(\theta) = \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} \tag{4}$$

$$b = \begin{bmatrix} r_{\text{offset}} \\ c_{\text{offset}} \end{bmatrix} \tag{5}$$

In Equation 5, r_{offset} and c_{offset} denote the pixel offset of row and column which are random value in each augmentation operation.

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics

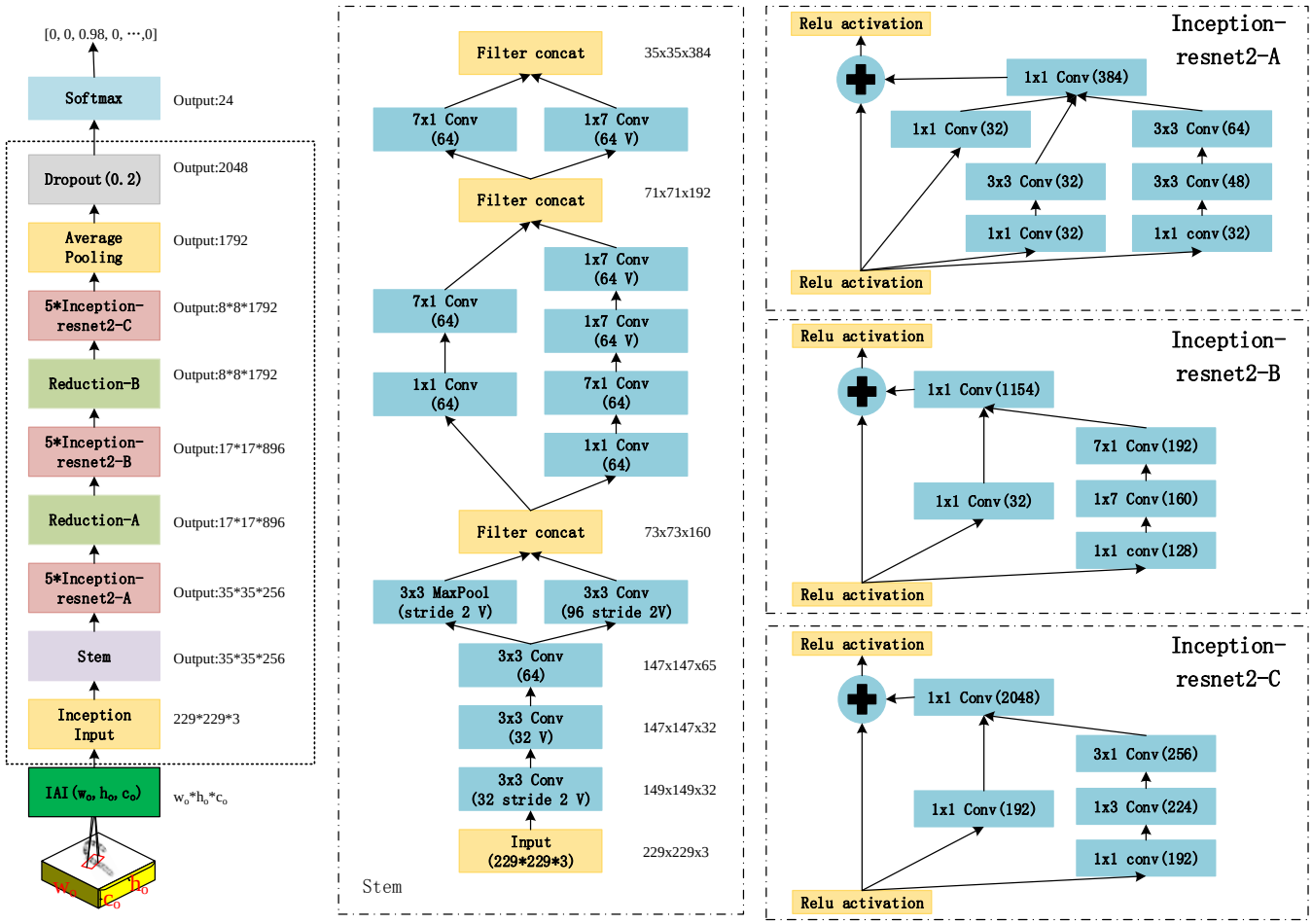


Fig. 2. On the left is the overall schema for *CIR-Net* which is modified from Inception-Resnet-v2 framework [11] by adding an *IAI* module and modifying the ways of *Softmax* from 1000 to 24. In the middle part, it is the detail of *Stem* module in the overall schema. The right part is the schema for interior grid modules. The *Inception-resnet2-A*, *Inception-resnet2-B* and *Inception-resnet2-C* are depicted from top to bottom.

The chromosome image $x_{(i,\theta)}$ augmented from x_i has the same label y_i as x_i. We use the symbol operation $f(x)$ to denote the mapping relation between x and y, which is formalized as *Equation 6*.

$$y_i = f(x_i) = f(x_{(i,\theta)}) \tag{6}$$

The intact algorithm for chromosome data augmentation operation is described as **Algorithm 1**.

3.3 IAI: Image Adaptive Interface

In many image classification tasks, the classifier expects a fixed shape of images such as $(224, 224, 3)$ in ImageNet [22] classification task, which means that each image of the dataset should have 3 color channels, each channel has 224-pixel spatial height and 224-pixel spatial width. For those images that do not satisfy the requirement of the given shape, they are needed to crop or pad into given shape by preprocessing operation before fed into the classifier. However, in clinical application, chromosomes are grayscale images that only have one channel. Additionally, chromosomes from various institutions may have small differences in their shapes, which limits the application capability of the classifier.

Motivating by this difficulty in clinical application, we design an Image Adaptive Interface (*IAI*) module for the classifier to accept the various shape of images adaptively.

$$\hat{S} = (h_t, w_t, c_t) \tag{7}$$
$$S = (h_o, w_o, c_o) \tag{8}$$

Supposing that the target shape of classifier is described as $Equation\ 7$, where h_t, w_t and c_t respectively denote the height, width and color channel of input. Accordingly we use $Equation\ 8$ denote the actual shape of given image, where h_o, w_o and c_o respectively denote the height, width and color channel of given image.

Consequently, we design *IAI* module using a 2D convolutional neural layer con2D() whose parameters includes $filter$, $kernel\_size$, $padding$, $strides$ and $input\_shape$. In these parameters, $filters$ is the depth of output, $kernel\_size$ is the kernel of the convolutional neuron, $padding$ refers to pixels of padding, and $strides$ denotes how many pixels skipping at the next convolutional op-

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics

Algorithm 1 CDA:Chromosome Data Augmentation

Require:
X, chromosome image dataset
Y, label set corresponding to X
r_{test}, the ratio of test data
r_{valid}, the ration of varify
function CHROMOSOME_AUGMENTATION(X,Y,r_{test},r_{valid})
  $test\_set \leftarrow \{\}$
  $train\_set \leftarrow \{\}$, $val\_set \leftarrow \{\}$
  $tmp\_set \leftarrow \{\}$
  $\Theta \leftarrow \{\theta_0, \theta_1, \dots\}$
  //splitting dataset into training set and test set by r_{test}
  for (x, y) in (X, Y) **do**
    if $rand(0, 1) < r_{\text{test}}$ **then**
      $test\_set \leftarrow test\_set \cup \{(x, y)\}$
    else
      $tmp\_set \leftarrow tmp\_set \cup \{(x, y)\}$
    end if
  end for
  //Augmenting tmp_set and spliting it into
  //training set and val set
  for (x, y) in tmp_set **do**
    for θ in Θ **do**
      $b \leftarrow Vector.random()$
      $x_{(\theta)} \leftarrow A(\theta)x + b$
      if $rand(0, 1) < r_{valid}$ **then**
        $val\_set \leftarrow val\_set \cup \{(x_{(\theta)}, y)\}$
      else
        $train\_set \leftarrow train\_set \cup \{(x_{(\theta)}, y)\}$
      end if
    end for
  end for
  return $train\_set$, val_set, $test\_set$
end function

eration. The relation of these parameters is limited as $Equation$ 9.

$$w_t = \lfloor \frac{w_o + 2 * p - k}{s} + 1 \rfloor$$
$$h_t = \lfloor \frac{h_o + 2 * p - k}{s} + 1 \rfloor \quad (9)$$

More precisely, when $w_t = w_o$, we set $p = 0$, $k = 1$ and $s = 1$, so the *IAI* module is a standard NiN [23]. While $w_t < w_o$, we set $p = 0$ so the *IAI* module is a cropping layer, otherwise we set $s = 1$, so the *IAI* module is a padding layer.

3.4 CIR-Net Architecture

The inception module was firstly proposed by Szegedy et al. [24] for improving the utilization of the computing resources inside the network on the competition of 2014 ILSVRC[1]. After that, Szegedy et al. proposed two upgrade versions of the Inception module named Inception-V2(V3) [25] to scale up the efficiency of networks in ways that aim at utilizing the added computation as efficiently as possible by suitably factorized convolutions and aggressive regulation. In 2017, the Inception-V4 and Inception-ResNet were proposed by Szegedy et al in the research report [11] for accelerating the training of Inception networks motivated by the success of Residual Learning [26] which yielded the state-of-the-art performance in 2015 ILSVRC by introducing the residual connections in conjunction.

The main idea of *CIR-Net* is to propose an efficient and outstanding framework for encoding sparse features of chromosomes which are sparse naturally. After our evaluation, we found the Inception-ResNet is a qualified architecture for chromosome classification tasks by appropriate modification.

The architecture of *CIR-Net* is shown in *Figure 2*. In *Figure 2*, the architecture, showing at the left part, is modified from the Inception-Resnet-v2 framework [11] by adding an *IAI* module and modifying the ways of *Softmax* from 1000 to 24. The *Stem* module of the overall schema is elaborated in the middle part while the right part is the schema for interior grid modules. The *Inception-resnet2-A*, *Inception-resnet2-B*, and *Inception-resnet2-C* are depicted from top to bottom.

4 EXPERIMENTS

4.1 Dataset Description

Medical Genetic Centre and Maternal and Children Metabolic-Genetic Key Laboratory of Guangdong Women and Children Hospital where we obtained 65 normal chromosome karyotypes as our research dataset supports this research. The privacy information of patients has been removed from these karyotypes. After that, we separated chromosomes from these karyotypes as the $dataset$ used for development, training, and verifying of *CIR-Net*. This $dataset$ can be detailed as follows:

- As each normal karyotype has 46 chromosomes, so we can get 2990 chromosomes from 65 karyotypes. This is a very big challenge for those models that are fed with large amounts of data.
- There are 32 male karyotypes and 33 female karyotypes in these karyotypes, which means that each autosome has 130 chromosomes labeled from 1 to 22 while sex chromosomes have 98 ($(65 - 32) \times 2 + 32 = 98$) X chromosome labeled to 23 and 32 Y chromosome labeled to 24.
- Each chromosome separated from karyotypes has a different number of pixels, 8 bits/pixel.

As this dataset is very insufficient, it is a huge challenge for those models that are fed by large amounts of data.

4.2 Evaluation Metrics

Since the performance of the chromosome classifier is usually evaluated by accuracy (acc), recall value ($recall$) and F1-score(F_1) in previous works [1], [9], [10], the performance of the *CIR-Net* is quantitative evaluated by there metrics.

To compute these metrics, we need to define the following four criteria to fit the context of the multi-class classifier as other researches do.

- True Positives(TP_j): Chromosomes are classified as type j which actually belong to type j.
- False Positives(FP_j): Chromosomes are classified as type j which actually do not belong to type j.
- False Negatives(FN_j): Chromosomes are classified as type $k(\forall k \neq j)$ which actually belong to type j.
- True Negatives(TN_j): Chromosomes are classified as type $k(\forall k \neq j)$ which actually do not belong to type j.

1. ImageNet Large-Scale Visual Recognition Challenge

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics

TABLE 1
Quantitative evaluation results

Methods	Parameters	Methods of Augmentation	*Precision*	*Recall*	F_1	*Acc*
Vanilla-CNN [9]	2,246,680	*None*	0.26	0.22	0.21	22.41%
		Straightening	0.82	0.85	0.82	83.78%
		CDA	0.88	0.86	0.87	86.44%(**2.66**% ↑)
SiameseNet [1]	1,416,281	*None*	0.22	0.22	0.21	21.91%
		Straightening	0.86	0.87	0.86	86.79%
		CDA	0.88	0.87	0.87	87.63%(**0.84**% ↑)
CIR-Net	55,258,360	*None*	0.43	0.31	0.28	28.42%
		Straightening	0.89	0.88	0.88	87.46%
		CDA	**0.96**	**0.96**	**0.96**	**95.98**% (**8.52**% ↑)

With the help of TP_j,FP_j,FN_j and TN_j, the metric of *precision* can be defined as *Equation* 11, the *recall* as *Equation* 13, F_1 as *Equation* 15 and *acc* as *Equation* 16.

$$precision_j = \frac{TP_j}{TP_j + FP_j} \tag{10}$$

$$precision = \frac{1}{N_{\text{types}}} \sum_{j=1}^{N_{types}} precision_j \tag{11}$$

$$recall_j = \frac{TP_j}{TP_j + FN_j} \tag{12}$$

$$recall = \frac{1}{N_{\text{types}}} \sum_{j=1}^{N_{types}} recall_j \tag{13}$$

$$F_{1j} = \frac{2 \cdot precision_j \cdot recall_j}{precision_j + recall_j} \tag{14}$$

$$F_1 = \frac{1}{N_{\text{types}}} \sum_{j=1}^{N_{types}} F_{1j} \tag{15}$$

$$acc = \frac{1}{N} \sum_{j=1}^{N_{\text{types}}} TP_j \tag{16}$$

In the above equations, N_{types} equals 24 denoting the types of chromosome while N denotes the total number of chromosome images in the test dataset.

4.3 Experiment Settings

We implement our *CIR-Net* using keras toolkit [27] based on TensorFlow [28]. We train the network utilizing RMSProp with the learning rate of 0.005. The experiments are accelerated by 2 × NVIDIA GeForce GTX 1080 with 8119 MiB GPU Memory.

In *CDA* Algorithm, we split 80 percent of the chromosome images into training set and the rest into *test_set* by setting r_{test} to 0.2. In training set, we split 50 percent of augmented chromosome images into *train_set* and the rest of augmented image into *val_set* by setting r_{valid} to 0.5. The Θ is set at 15 degrees from 15 to 345.

4.4 Results

This section presents the experimental results of the quantitative and qualitative evaluation utilizing T-SNE [12] of the proposed methods.

4.4.1 Quantitative Evaluation Results

The general experimental results of various classifiers are shown in *Table 1*. Although we implemented the SiameseNet classifier [1] and Vanilla-CNN classifier [9] in strict accordance with their origin paper carefully, there may be some small differences caused by lacking their engineering tricks in their papers.

We use *None* to denote preprocessing without augmentation operation where chromosomes are at random directions and

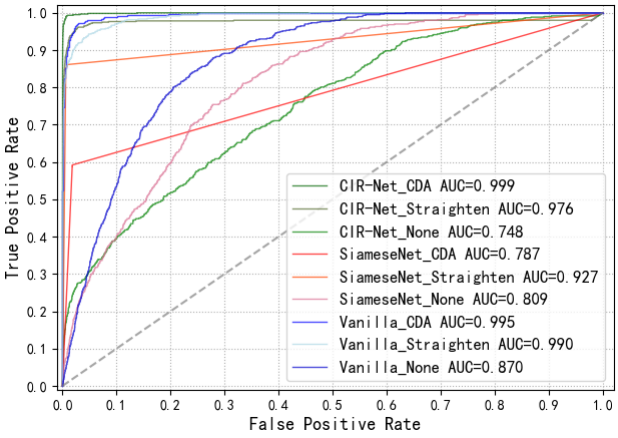


Fig. 3. Comparisons on ROC curves and AUC value of various classification methods.

centroids. To compare the effect of the *CDA* algorithm, we implement augmentation methods of *CDA* and *Straightening* as the guide of the original paper [19]. The biggest difference between *CDA* and *Straightening* is that *CDA* only augments the training chromosome images while *Straightening* requires to augment testing chromosome images as training images. It means that, in clinical application, all chromosomes ready to classify are required to straighten as the training images of the classifier.

Each method has been carefully tuned to achieve its best performance on each given dataset. The experimental results are summarized in *Table 1*.

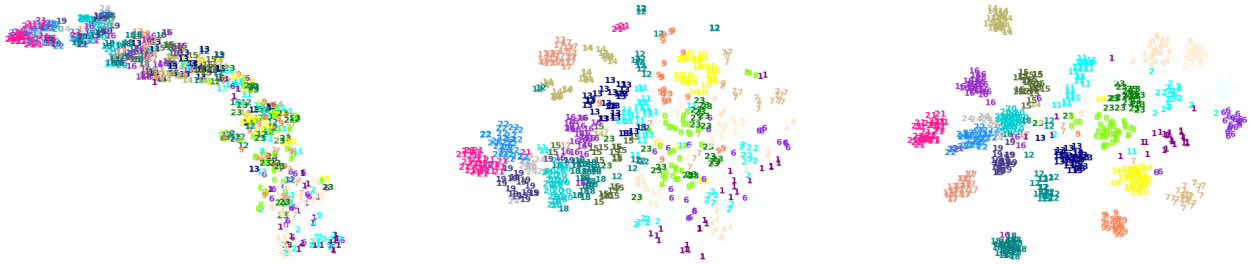
According to the experimental results in *Table 1*, our proposed *CDA* algorithm further improves the performance of classifiers by 8.52 percent of *acc* metric compared to *Straightening* augmentation in *CIR-Net*, 0.84 in SiameseNet [1], and 2.66 in the Vanilla-CNN [9]. The reason is that we feed the same chromosome with several augmented images that contain various deformations to train the classifier, which is to tell the classifier that direction and position are not the features to discriminate the type of chromosome.

Our proposed *CIR* method combined with the *CDA* algorithm achieves a classification accuracy of 95.98%, which is better than other methods. This is because the Inception-ResNet architecture consists of three different meta-modules Inception-resber2-A, Inception-resber2-B, and Inception-resber2-C. These meta-modules assemble different convolutional neurons, which have capabilities to extract richer features of different scales for high-level feature learning and discrimination.

To evaluate the stability of our proposed method under different data, we performed cross-validation experiments using KFold. We load the raw data and use KFold to divide the data into five folds in random order. We use one fold data as the verification data in turn, and the other four folds data

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics



(a) (b) (c)

Fig. 4. The above sub-figures visualize the chromosome features extracted from CIR-Net trained by various data augmentation methods. Figure 4(a), 4(b) and 4(c) are visualization cases of chromosome features in the test set, these features are respectively extracted from various instances of CIR-Net trained by original, straightening and CDA augmentation dataset.

as training data. We train the model five times and the cross-validation experimental results show that our mean accuracy and standard deviation value is (96.99 ± 1.70).

At last, ROC curves and AUCs are universal evaluation metrics in binary classification problems but it does not support multi-classification problems directly. We convert multi-classification into binary classification in each category for computing the average ROC curves and AUC value. The comparison results of various classification methods are showing in Figure 3. The higher of AUC value is better.

4.4.2 Qualitative Evaluation Results

For evaluating the classification performance of *CIR-Net* intuitively, we get all features of chromosomes in the test set from the *flatten* layer before the *Softmax* layer of *CIR-Net* and reduce them to two-dimensional vector using *T-SNE* [12]. After that, we visualize them in *Figure 4.3*.

In *Figure* 3, the Arabic numerals 1 to 22 denote 22 autosomes separately, the Arabic numerals 23 and 24 denote *X* and *Y* sex chromosome. The more close the same numbers crowds together, the far the different numbers separate, which means the better discriminative ability of the model.

In *Figure* 4(a), the different Arabic numerals are crossed and distributed in stripes. This means that the model trained by without the original training dataset cannot discriminate against chromosomes.

Compared to *Figure* 4(a), the same Arabic numerals tend to cluster, but the different Arabic numerals are still heavily crossed in *Figure* 4(b), which means that the instance of *CIR-Net* trained from the straightened training dataset has a certain degree of discriminating capability.

Obviously, in *Figure* 4(c), although there are some cases where different Arabic numerals intersect, the same Arabic numerals are gathered, and the different Arabic numerals are separated. This means that the pattern has a strong ability to discriminate chromosome categories.

5 Conclusion

In this research, we proposed the *CIR-Net* based on the Inception-ResNet architecture method for chromosome classification on the insufficient dataset.

The most distinctive characteristics of *CIR-Net* include:

- We build *CIR-Net* inherited the outstanding specialties from *Inception-ResNet*.
- The *CDA* algorithm based on affine transformation was proposed to augment insufficient training chromosome images.
- We proposed an Image Adaptive Interface (*IAI*) for improving the capacity of clinical application.

According to the experimental results, we can draw two conclusions:

- The classification performance of our proposed method is better than other previous in the training set with insufficient data.
- The proposed data augmentation algorithm CDA can effectively deal with the problem of insufficient chromosome training data by enriching the chromosome variants in the training dataset.

There are two main reasons contributing to the better performance of the proposed method compared to the previous methods [1], [9] in the training set with insufficient data. The first reason is that these methods use straightening for pre-processing, but straightening can lead to distortion of the information on the chromosome. Different from the straightening method, the proposed CDA algorithm adopts to enrich chromosome variants to prompt the classifier that these chromosomes belong to the same category. Another reason is that we build the chromosome classifier on the well-proven mature architecture textitInception-ResNet, rather than directly proposing a completely new architecture.

Author's Contributions

G.Zhao and Z.yang provided guidance and planning for the project. A.Yin, L.Guo, and H.Chen provided the professional guidance of medical knowledge. X. Wang, L.Zhao, and H.Luo reproduced the methods in the contrast experiment based on their original papers. T.Wang, B.Ding, X.Pang, Z.Ma, and Q.chen provided their assistant to original data processing.

Acknowledgments

This work was supported by National Key R & D Program of China (2018YFB1404402 and 2018YFB1404400), Guangdong Science & Technology Fund (No.2019B010137003, No.2016B030305006, No.2018A07071702, and No.201804010314), Guangzhou Science & Technology Fund (No.201804010314 and No.2012224-12), and VeChain Foundation (No.SCNU-2018-01). Z. Yang is supported by Academy of Finland (projects 314177 and 307929) and The Research Council of Norway (project 287284).

References

[1] S. Jindal, G. Gupta, M. Yadav, M. Sharma, and L. Vig, "Siamese networks for chromosome classification," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 72–81.

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics

[2] T. Arora and R. Dhir, "A review of metaphase chromosome image selection techniques for automatic karyotype generation," *Medical & biological engineering & computing*, vol. 54, no. 8, pp. 1147–1157, 2016.

[3] J. Piper and E. Granum, "On fully automatic feature measurement for banded chromosome classification," *Cytometry: The Journal of the International Society for Analytical Cytology*, vol. 10, no. 3, pp. 242–255, 1989.

[4] G. Agam and I. Dinstein, "Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification," *IEEE Transactions on pattern analysis and machine Intelligence*, vol. 19, no. 11, pp. 1212–1222, 1997.

[5] P. A. Errington and J. Graham, "Application of artificial neural networks to chromosome classification," *Cytometry: The Journal of the International Society for Analytical Cytology*, vol. 14, no. 6, pp. 627–639, 1993.

[6] G. Ritter and M. T. Gallegos, "Outliers in statistical pattern recognition and an application to automatic chromosome classification," *Pattern Recognition Letters*, vol. 18, no. 6, pp. 525–539, 1997.

[7] X. Hu, W. Yi, L. Jiang, S. Wu, Y. Zhang, J. Du, T. Ma, T. Wang, and X. Wu, "Classification of metaphase chromosomes using deep convolutional neural network," *Journal of Computational Biology*, 2019.

[8] M. Sharma, L. Vig *et al.*, "Automatic chromosome classification using deep attention based sequence learning of chromosome bands," in *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2018, pp. 1–8.

[9] W. Zhang, S. Song, T. Bai, Y. Zhao, F. Ma, J. Su, and L. Yu, "Chromosome classification with convolutional neural network based deep learning," in *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 2018, pp. 1–5.

[10] Y. Qin, J. Wen, H. Zheng, X. Huang, J. Yang, L. Wu, N. Song, Y.-M. Zhu, and G.-Z. Yang, "Varifocal-net: A chromosome classification approach using deep convolutional networks," *IEEE transactions on medical imaging*, 2019.

[11] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[12] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[13] M. Egmont-Petersen, D. de Ridder, and H. Handels, "Image processing with neural networksa review," *Pattern recognition*, vol. 35, no. 10, pp. 2279–2301, 2002.

[14] M. Popescu, P. Gader, J. Keller, C. Klein, J. Stanley, and C. Caldwell, "Automatic karyotyping of metaphase cells with overlapping chromosomes," *Computers in biology and medicine*, vol. 29, no. 1, pp. 61–82, 1999.

[15] M. Moradi, S. Setarehdan, and S. Ghaffari, "Automatic landmark detection on chromosomes' images for feature extraction purposes," in *3rd International Symposium on Image and Signal Processing and Analysis, 2003. ISPA 2003. Proceedings of the*, vol. 1. IEEE, 2003, pp. 567–570.

[16] W. Sweeney, M. Musavi, and J. Guidi, "Probabilistic neural network as chromosome classifier," in *Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan)*, vol. 1. IEEE, 1993, pp. 935–938.

[17] E. Poletti, A. Ruggeri, and E. Grisan, "An improved classification scheme for chromosomes with missing data," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 5072–5075.

[18] S. Rungruangbaiyok and P. Phukpattaranont, "Chromosome image classification using a two-step probabilistic neural network." *Songklanakarin Journal of Science & Technology*, vol. 32, no. 3, 2010.

[19] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in *ICML deep learning workshop*, vol. 2, 2015.

[20] X. Yu, X. Wu, C. Luo, and P. Ren, "Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework," *GIScience & Remote Sensing*, vol. 54, no. 5, pp. 741–758, 2017.

[21] Z. Hussain, F. Gimenez, D. Yi, and D. Rubin, "Differential data augmentation techniques for medical imaging classification tasks," in *AMIA Annual Symposium Proceedings*, vol. 2017. American Medical Informatics Association, 2017, p. 979.

[22] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[23] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.

[24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[27] F. Chollet *et al.*, "Keras: The python deep learning library," *Astrophysics Source Code Library*, 2018.

[28] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.

Chengchuang Lin is a PhD student from School of Computer Science, South China Normal University who is major in Computer Vision and Medical Image Processing. Before that, he received master's degree from School of Computer Science, South China Normal University and bachelor's degree from School of Physics & Optoelectronic Engineering, Guangdong University of Technology.

Gansen Zhao received the PhD degree from the University of Kent, Kent, United Kingdom. He is currently a faculty member with the School of Computer Science Department, South China Normal University, Guangzhou, China. His research interests include cloud computing, security, privacy, big data, and blockchain technology. He served as the program chair of the first IEEE International Conference on Cloud Computing (CloudCom2009). He is a member of the IEEE.

Zhirong Yang received his Bachelor and Master degrees from Sun Yat-Sen University, China in 1999 and 2002, and Doctoral degree from Helsinki University of Technology, Finland in 2008, respectively. Presently he is a professor with the Norwegian Open AI Lab and Department of Computer Science at Norwegian University of Science and Technology. He also works as Finnish Academy Research Fellow at Aalto University, Finland. His current research interests include information visualization, cluster analysis, learning representations for structured data.

Aihua Yin is director of medical genetics center of Guangdong Women and Children Hospital, Professor of Obstetrics and Gynaecology at the Jinan University. She is leading medical talents and outstanding young medical talents in prenatal diagnosis in Guangdong Province. She is a expert of national expert group of prenatal diagnosis technology.

Xinming Wang received the PhD degree from the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. He is currently a research fellow with the Lakala Group, China. His research interests include finance technology, software testing and analysis, and block-chain technology. He is a member of the IEEE.

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCBB.2020.3003445, IEEE/ACM Transactions on Computational Biology and Bioinformatics

Li Guo is a doctor of obstetrics and gynecology, major in cytogenetic disease for prenatal diagnosis in Women and Children Hospital in Guangdong Province .

Hanbiao Chen is a technician of medical genetics center of Guangdong Women and Children Hospital. He has been engaged in cell genetic experiment and has rich experience in cell culture and karyotype analysis.

Lei Zhao is a master student of computer science from South China Normal University, whose research interest fields include computer vision, data mining and medical image processing.

Haoyu Luo received his PhD degree from Wuhan University, China, in 2018. He received his Bachelor degree from East China Institute of Technology in 2011 and Master degree from Northwest Normal University in 2014. He is currently an Associate Researcher in School of Computer Science, South China Normal University. His research interests include workflow system and edge computing.

Tianxing Wang is pursuing a master's degree in Software Engineering from South China Normal University. Her research interests include image classification, object detection and deep learning.

Bichao Ding is pursuing a master's degree in Software Engineering from South China Normal University. His research interests include image classification, instance segmentation and deep learning.

Xiongwen Pang received his PHD degree in Computer Science and Engineering from South China University of Technology, China in 2007. He is currently working as associate professor at South China Normal University, China, His teaching and research interests include image processing and deep learning.

Zhaohui Ma is a lecturer at Guangdong University of Foreign Studies, PhD candidate at South China Normal University. He is major in network security, software defined network.

Qiren Chen is pursuing a master's degree in Software Engineering from South China Normal University. His research interests include Natural Language Processing and deep learning to solve medical problems.

1545-5963 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Norges Teknisk-Naturvitenskapelige Universitet. Downloaded on August 27,2020 at 13:31:10 UTC from IEEE Xplore. Restrictions apply.