IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 30 November 2023, accepted 14 December 2023, date of publication 18 December 2023,
date of current version 28 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3344664

RESEARCH ARTICLE

A Novel Automated Chromosome Analyzer Software Bundle for Karyotyping and Birth Defect Analysis

DEVARAJ SOMASUNDARAM[1], (Member, IEEE), ADITYA CHAPPARADAHALLIMATH[1], MURALITHARAN KRISHANAN[2], (Member, IEEE), AND NIRMALA MADIAN[3], (Member, IEEE)

[1]Department of Micro and Nano Electronics, School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, 632014, India
[2]Department of Computer Science, Institute of Mathematical Sciences, Sungkyunkwan University, Suwon 16419, South Korea
[3]Department of Biomedical Engineering, Dr. N. G. P. Institute of Technology, Coimbatore, Tamilnadu 641048, India

Corresponding author: Devaraj Somasundaram (Somasundaram.d@vit.ac.in)

This work was supported by VIT Seed grant and VIT Publication.

ABSTRACT Karyotyping is a procedure to diagnose birth defects using chromosome pair. During the Karyotyping chromosomes arranged based on the length and each chromosome will be paired based on various parameters such as, chromosome length, banding pattern and Centromere position. Many methods are proposed to identify the above parameters to improve the Karyotyping accuracy. Since, it's a challenging task for researchers to improve the accuracy of Karyotyping compared with clinical assays. In this paper, a novel computer geometry method is proposed for chromosome Karyotyping using inbuilt deep learning models with algorithms for chromosome segmentation, overlapped separation, banding pattern analysis and classification of chromosomes. Chromosome classification is carried out using various deep learning models and automated Karyotyping is carried out without manual intervention and model improved the Karyotyping accuracy over the existing methods. In this paper novel computer geometry method is proposed to automate the chromosome analysis in Matlab environment. The developed software provides the accuracy of 99.68% in classification and karyotyping.

INDEX TERMS Deep learning, chromosome, karyotyping, banding pattern.

I. INTRODUCTION

Chromosomes are plays major role in birth defect analysis. Chromosomes collected from women amniotic fluid and clinical analysis is carried out using the microscope. In which, based on the die and filter usage in the microscope, variety of categories chromosomes observed such as M-Fish, G-band chromosomes, C-band chromosomes etc., To identify the birth defects chromosome analysis is used based on the karyotyped. In this process chromosomes arranged in descending order based on the length of the each chromosome. Some instance this arrangement depend on the Centromere position.

Centromere is the position where the distribution of the thickness is minimum except from top and bottom ends. From the Centromere, p-arm and q-arm lengths are measured. The measured features will be used to classify the each individual chromosome. Each pair of chromosomes will be identified based on these parameters. More than that, chromosome banding pattern is a feature used to classify the chromosome from one to another.

In the metaspread, usually chromosomes are overlapped one to another and creates a non regid object shapes. This is highly complex to understand the chromosome and its pair, due to one overlapped chromosomes. it means that one chromosome is overlapped with another one. Two overlapped chromosomes. it means that one chromosome is touched or overlapped other two of individual chromosomes, multiple chromosome overlapped means more than three chromosomes are created an non rigid cluster and cannot be identified which chromosomes belongs to which category and its pair.

In the clinical laboratories the expert or well trained technicians will separate the chromosomes manually and the grouping can be done based on the length only. Usually the cluster distribution will be in the range of one overlapped

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

chromosomes will be around 15 % and two overlapping around 15 % and multiple overlapping is beyond 20% of chromosomes in the metaspread. Since the challenging task is to identify the pair. In many cases, abnormal chromosomes have different length and it may match with normal chromosomes due to the lack of information or expertise in the manual analysis. This will be resolved using the artificial intelligence methods. Still the accuracy is not merely improved to reach the automatic analysis which gives the better results than manual process.

Another challenging task is to identify the banding pattern in the chromosomes. when the length parameter fails to identify the particular chromosome, then the next parameter will be P-arm and Q- arm length. In that, banding patteren plays role to identify the chromosomes. But the separating the banding patterns is too complicated due to its poor visibility and the banding pattern will not be distributed uniformly in the chromosomes. Based on various other parameters these banding pattern spread occurs in the chromosomes. To analyze these chromosomes, a basic semi automated models and automated models are proposed earlier. Since there is a challenge is to develop a fully automatic metaspread to karyotyping system. Manual analysis of chromosomes is highly complex due to highly degraded structure. To resolve this issue an automated software is proposed to carried out the karyotyping from metaspread images.

In this computer aided tool handle the different overlapped regions up to three overlapped chromosomes and classify and matched the respective chromosomes to create a karyogram.

In this manuscript, we proposed an artificial intelligence based automated chromosome analyzer for Karyotyping is proposed.

The main contribution in this analyzer are summarized below,

- Novel software is proposed to resolve the manual intervention in the chromosome analysis. This works based on the computational geometry and deep models. Which includes,
  - Novel overlapping chromosome analysis and separation
  - Band measurement in the each chromosome to classify the chromosome category.
  - Overall separation and segmentation of chromosomes carried out based on area and banding pattern in overlapped chromosomes.
  - Computer aided tool to make the karyogram for G-band chromosomes.

All these modules are implemented with various algorithms and using deep learning models as explained in the methodology section.

II. RELATED RESEARCH

In the chromosome analysis, the segmentation, overlapping separation and classification of the chromosomes are carried out by various researchers as shown in the table 1.

TABLE 1. Comparative analysis and methods of chromosome.

Author /year	Segmentation Method and Accuracy	Remarks
Gady Agam (1997)	Hypothesis Based Approach	• Worked on overlapping and touching contour images • Concave/ convex points and separation lines are identified • Hypothesis based analysis is performed
Petros Karvelis et al (2005)	Watershed Transform(WT) & 92	• Worked mainly on touching chromosomes • Segmentation done by concave points • 940 images considered (340 touching and 29 overlaps)
Wacharapong Srisang et al (2006)	Computational Geometry & 85	• Worked on overlapping chromosomes • Separation done using voronoi diagram with possible cut points obtained by DT • 28 out of 35 overlaps are separated correctly
Enrico Grisan et al (2007)	Space Variant Thresholding & 92 for adjacencie and 90 for overlapped	• Worked on overlapping and touching chromosomes • 30 metaspread images with 1380 chromosomes are considered for analysis
Enrico Grisan et al (2009)		162 image from 117 cells with 6683 chromosome are taken for analysis
Petros Karvelis et al(2010)	Watershed Transform & 90.6 for touching and 80.4 for overlapping	• Worked on overlapping and touching M-FISH chromosomes • 183 M-FISH images are taken for analysis
Mousami V Munot et al (2011)	Modified Snake Algorithm with Greedy Approach /100 for 2 touch 95 for 3 & 4 touch	Worked only on touching chromosomes

TABLE 1. *(Continued.)* **Comparative analysis and methods of chromosome.**

Mukul A Joshi et al(2012)	Computational Geometry/ 100 for 1 & 2 overlaps and 88 for 3 and 4 overlaps.	• Worked only on overlapping chromosomes • Tested on synthesized images of LK1 data set
Mousami V Munot et al (2013)	Computational Geometry/98 (only cut point identification)	• M-FISH images are considered for analysis. • Worked only on overlapping chromosomes
Tanvi and Renu Dhir (2014)	Computational Geometry/87.4	• Worked only on overlapping chromosomes • Cut points are tracked along contour for separation of overlaps
Devaraj Somasundaram (2014)	Geometry Separation Algorithm /94	• Worked only on overlapping chromosomes • Cut points and cut lines are identified and hypothesis verification done for separation • Homologue chromosome identification is done by centromere position
Author	**Method**	**Remarks**
Goesta H Granlund (1976)	Centromere using Integrated Density Profile /70	• Four different types of descriptors are obtained using integrated density profile • 230 chromosomes are considered, the recognition rate for above descriptors are 69.1%, 70.1%, 64.7% and 75.2%
Frans CA Groen et al (1989)	Centromere using Piecewise Linear Approximation Method /85 and 68 for Leyden data base and 93 and 76 for Copenhagen data base	• Closest pair of opposite contour points are identified. • Profile for the width of chromosomes is plotted • Medial axis is obtained for identification of centromere
Xingwei Wang et al (2008)	Centormere and Polarity Assignment using Multi Stage Rule Based Algorithm/91.4 for centromere 97.4 for polarity assignment	• Three feature profile are extracted • Medial axis of chromosome is obtained using thinning algorithm
Mousami V Munot et al (2012)	Banding Pattern Similarity Measure /100 for group A 97 for group C chromosomes	• Novel algorithm for nearness factor calculated for banding pattern similarity measure • 50 images of group A and C chromosomes are taken for the analysis
Author	**Classifier**	**Remarks**
Boaz Lerner et al (1998)	MLP Trained Back Propagation Classification Algorithm/ 90	• Classification-Driven Partially Occluded Object Segmentation is performed • Two geometrical features (length and Centromere index) are considered for classification
John M Conroy et al (2000)	Singular Value Decomposition, Principal Component Analysis, Fisher Discriminant Analysis and Hidden Markov Models /97	• Tested on G band chromosome images of well known database (Philadelphia, Edinburgh, and Copenhagen) • Markov models proves better for classification
Zhenzhen Kou (2002)	Support Vector Machine /90	• Simple feature extraction process done with density profile and use normalized profile as feature vector
Delshadpour S (2003)	Multi Layer Perceptron Neural Network / 88	• Output neurons is reduced to an order of n to $\log_2\{n\}$ help in reducing dimensionality of the network, required number of training data, training time and generalization of network error.

TABLE 1. *(Continued.)* **Comparative analysis and methods of chromosome.**

		• Copenhagen database images are taken for the analysis. • 304 images are taken for the study
Petros S Karvelis et al (2006)	Bayes Classifier /89	Morphological watershed segmentation is applied on intensity gradient image which helps in decomposing the image into a set of homogeneous regions
Xingwei Wang et al(2008)	Decision Tree and Artificial Neural Network (ANN) /86	Tested on 170 images and process adapted are image filtering, thresholding and labeling algorithms
Petros S Karvelis et al (2008)	Bayes Classifier /82.4	Spatial and spectral characteristic regions are taken as input for classifier
Yaser Rahimi et al(2008)	Multi-Layer Feed forward Perceptron Neural Network /73	Features for classification are surface of chromosome, boundary pixel of chromosomes, and six momentums
Benoit Legrand et al (2008)	Dynamic Time Warping /81	Features for classification are length and density profile
Sunthorn Rungruangbaiyok and Pornchai Phukpattaranont (2010)	Probabilistic Neural Network / 68.19 for female samples 61.3 for male samples	features considered for chromosome classification are area, perimeter, band area, profile and singular value decomposition
Enea Poletti et al (2012)	Neural Network Classifier /94	• 5474 chromosomes are considered for the classification • Medial axis and polarization estimation is done for feature extraction. • Novel class reassignment algorithm is used along with the neural network classifier to increase the probability of classification rate

Previously chromosome karyotyping is carried out based size, shape, and banding patterns. To improve the accuracy of the separation, with the above feature, area is used. Moreover, the separation of the chromosome is highly complex using the other features. So, area is used to identify the category of the chromosomes and separation process. This improved the accuracy in karyotyping. Karyotyping is a technique used to visualize and analyse the number, size, and shape of chromosomes in an individual's cells. This process is particularly important for detecting chromosomal abnormalities and genetic disorders.

The karyotyping Process involves with the morphological operations, opening and closing carried out in the image, the binary image, each position scanning elements categorized based on the area, length. Touching and overlapping has similar features. In this analysis, Centromere positions are two or lesser than two in usual overlapped chromosomes. If touching chromosomes, the Centromere position is more than two due to the touching point also detected as Centromere position [3].

Concave and convex points in the context of karyotyping are significant because they contribute to the unique banding patterns of chromosomes. These patterns are instrumental in the identification of individual chromosomes, the detection of chromosomal abnormalities [3].

Separation lines used to carryout the hypothesis analysis on the overlapping chromosomes. This provides the possible separation segments in the overlapped chromosomes.

The proposed method deals with the computational geometry to carry out the karyotyping. Computational geometry is instrumental in developing algorithms and techniques for automated and accurate chromosome separation, segmentation, and analysis. The application of geometric principles enhances the efficiency and reliability of karyotyping processes, allowing for more precise identification of chromosomes, detection of abnormalities, and quantitative assessment of chromosomal features. Computational geometry techniques are employed to develop algorithms for automated image analysis of chromosome spreads. Automated systems use geometric principles to detect and separate individual chromosomes, reducing the need for manual intervention and increasing the speed of analysis.

Identifying the overlap zone in the karyotyping process is crucial for accurate chromosome identification, avoiding misinterpretation, detecting structural abnormalities, and ensuring the reliability of both manual and automated analyses. Failure to correctly identify the overlap zone may lead to misinterpretations of the karyotype. Mistaking an overlap for a structural abnormality or vice versa can result in inaccurate diagnoses and unnecessary concerns.

An image that is projected from a multidimensional space into a one-dimensional space is represented mathematically by a projection vector. Application of projection vectors along particular directions (e.g., horizontal or vertical) over

the chromosomal image is done in the context of chromosome analysis. The projection vector method, which examines intensity profiles along particular axes, is useful in locating structural abnormalities in chromosomes. This computational technique improves karyotyping automation, impartiality, and accuracy.

Geometric principles are employed for accurate feature extraction from chromosome images. This includes measurements of length, width, and the position of centromeres, contributing to precise identification and characterization of chromosomes. computational geometry methods enable the three-dimensional reconstruction of chromosomes from 2D images. This additional dimension provides more comprehensive information about the spatial organization of chromosomes.

III. METHODLOGY

The metaspread G-band chromosome images considered for karyotyping. In this proposed model, the initial separation will be carried out based on the area of the each region in the image (m × n). m and n are the size of the chromosome image (256 × 256). These images go through the basic morphological analysis, such as opening and closing followed by area measurement of the each region. The area spectrum will be analysed based on the standard karyotyped images. Usually, Chromosome 1 has higher length and next will be the chromosome 2 and so on. So the area of the each chromosome region is detected and area is measured. The regions are compared with one another.

In this case, the chromosome 1 region (CA1), chromosome 2 region (CA2) and so on are may be detected properly or may not be identified due to the uncertainty in the input image.

Overall chromosome regions ordered as,

$$CA1 > CA2 > CA3 > CA4 \ldots > CA23.$$

In that, if the image as overlapping then we will not get the all 23 regions and as well as the condition of the above will be violated. If two chromosomes overlapped then the overall regions will be reduced to 22 and the area of the region is become too high. Here the statistical analysis will be carried out to identify the region overlapping as shown in the Table 2.

Morphological operations are carried out and projection vector is used to identify the Centromere position. From the Centromere position, the length of the image is measured and as two arms of chromosomes. The Table 1 shows the chromosome 1 overlapped with the other chromosomes of 22. The condition is analysed to identify the chromosomes groups. In the similar method, the artificially, the overlapping is generated and created an dataset of chromosome 1 over lapped with other 23 chromosomes and also with other chromosomes with each individual chromosomes are done to train the deep learning model. The table 3 shows the area measurement of each chromosome region individually.

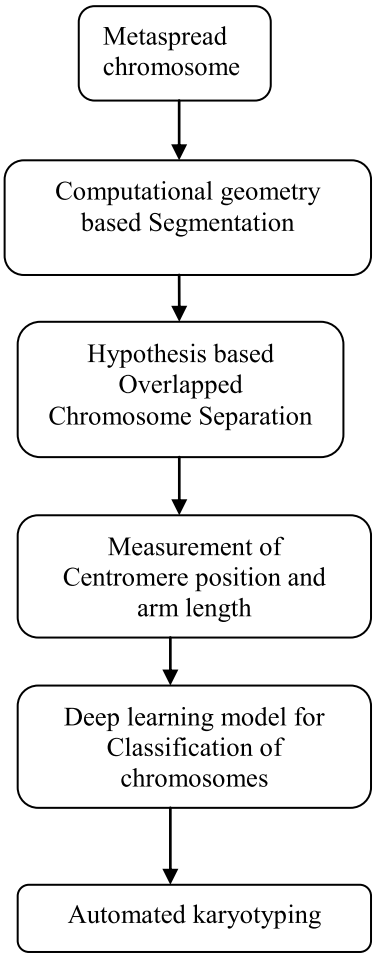

FIGURE 1. Proposed overall method.

TABLE 2. Sample region identification based on chromosome.

S.No	Chromosome region 1	Chromosome region 2	Nature of Overlap region	Condition in group
1	Chr 1	Chr 1	Chr1 + Chr1	Max area >chr1 <chr2
2	Chr 1	Chr 2	Chr1+chr2	MaxArea>chr1< 2chr1
3	Chr1	Chr 3	Chr 1+chr3	MaxArea>Chr1<2chr2
4	Chr 1	Chr 4	Chr 1+ chr 4	MaxArea>chr1< 2chr3
5	Chr 1	Chr 5	Chr1+ chr 5	MaxArea>chr1<2chr4
6	Chr1	Chr6	Chr1+chr6	MaxArea>chr1<2chr5
7	Chr1	Chr7	Chr1+chr7	MaxArea>chr1<2Chr6
8	Chr1	Chr8	Chr1+chr8	MaxArea>chr1<2Chr7

These above avg. measurements are considered as reference to analyze the portion of the overlapped chromosomes with proposed deep learning model for the classification of overlapping chromosomes.

TABLE 3. Chromosome area and length measurement.

Chromosome number	Chromosome Avg. length	Chromosome Avg. area
1.	145.19	3549
2.	154.05	3506.4
3.	128.25	3439.9
4.	138.15	3312
5.	120.53	2850.6
6.	116.45	2789.1
7.	114.65	2874.3
8.	104.85	2531.5
9.	104.54	2606.8
10.	97.89	2735.9
11.	99.70	2542.3
12.	94.87	2451.5
13.	93.56	2309.5
14.	91.88	2228.1
15.	79.74	1921
16.	75.01	1986.8
17.	65.02	1932.3
18.	78.55	2087
19.	57.05	1587.3
20.	63.08	1512.8
21.	62.03	1451
22.	59.92	1553.9
23.	57.72	1396.8

The initial points used in the scanning procedure are in the position indicated by the letter ''t,'' which is slightly off-center. From this deviation, a length 't+s' point is found. This length aids in achieving continuous curvature across all non-zero places. Equation (1) provides the scanning, which follows the coordinates of a closed ''n'' link chain and pixel locations. pixel locations.

$$A^n = C_{i-1}^n a_i = a_1 a_2 \ldots\ldots\ldots a_n \tag{1}$$

where C is the pixel point and a_i is the ith link in the image. Consider L_j^s as a sub chain link which act as a termination node where a_j link is directed. The L_j^s is expressed by Equation (2)

$$L_j^s = C_{i=j-s+1}^{j} a_i \quad j = 1, 2, 3 \ldots\ldots n \tag{2}$$

The L_j^s sub chain link has X and Y image coordinates. The X and Y components of the link are computed and is given as Equation (3)

$$X_j^s = \sum_{i=j-s+1}^{j} a_i x \quad \text{and} \quad Y_j^s = \sum_{i=j-s+1}^{j} a_i y \tag{3}$$

where $a_i x, a_i y \in \{-1, 0, 1\}$

Equation (4) provides a straightforward distance formulation that uses the X and Y components to determine the chromosomal length.

$$l_j^s = \sqrt{\left(X_j^s\right)^2 + \left(Y_j^s\right)^2} \tag{4}$$

With chromosomal length and inclination angle, the curvature function may be determined well. In Equation (5), the X-axis's angle is specified.

$$\theta_j^s = \tan^{-1} Y_j^s / X_j^s \quad if \quad \left|X_j^s\right| \geq \left|Y_j^s\right| \tag{5}$$

To smooth the curvature function, an incremental curvature function is chosen. This helps to eliminate the undesirable points. The resultant curvature is twice the average of the two next-to-one angular deviations. Equation 6 describes the as,

$$\begin{aligned} \delta_j^s &= 2\left[\frac{\left(\theta_{j+1}^s - \theta_j^s\right) + \left(\theta_j^s - \theta_{j-1}^s\right)}{2}\right] \\ &= \theta_{j+1}^s - \theta_{j-1}^s \end{aligned} \tag{6}$$

A. CONCAVE POINTS AND SEPARATION LINES

By raising the S value, the curvature curve becomes more smoothly defined. It is important to identify the actual concave points on the image. Based on measurements of angle, area, and distance between pixel locations, concave points are determined. The concave points are primarily obtained by observing the nearby pixel points.

To determine the concavity property, the current and prior values are identified. Where c - is the current pixel value, X, and Y are the image coordinates; connectedness of the points is represented as,

The concavity of the image contour

$$\begin{aligned} a\left(p_{pre}, p_c\right) &= \tan^{-1}\left(\left(y_{pre} - y_c\right) / \left(x_{pre} - x_c\right)\right) \\ a\left(p_{nert}, p_c\right) &= \tan^{-1}\left(\left(y_{nert} - y_c\right) / \left(x_{nert} - x_c\right)\right) \end{aligned} \tag{7}$$

The touching and overlapping images are distinguished by taking into account an appropriate neighborhood connection attribute. By adding more concave sites, contacting chromosomes can be segmented and untangled with greater accuracy.

The resulting concave points are then used to determine separation lines. Separation lines are used to separate overlapping objects as well as things that are touching one another. All conceivable combinations of high concave point pairs are used to determine separation lines. To determine the potential separation lines, every combination of any two points is taken into account. The pairs with identical element pairs and reflexive pairs are then dropped. Equation is then used to represent the image set.

$$\begin{aligned} S_i \equiv C_i^* \times C_i^* - \left\{\left(p_{k,l}, p_{m,n}\right) \in C_i^* \times C_i^* \mid (k - I_c + l)\right. \\ \left. \geq (m - I_c + n)\right\} \end{aligned} \tag{8}$$

where C_i^* is the set of high concave poipts along the contour of the image. $\left\{\left(p_{k,l}, p_{m,n}\right) \in C_i^*\right\}$ defines the separation line that separate the entire set into two disjoint subsets. To preserve the connectivity between the pixels between the subsets discrete line connecting the pair $p_{k,l}$ and $p_{m,n}$ is introduced by

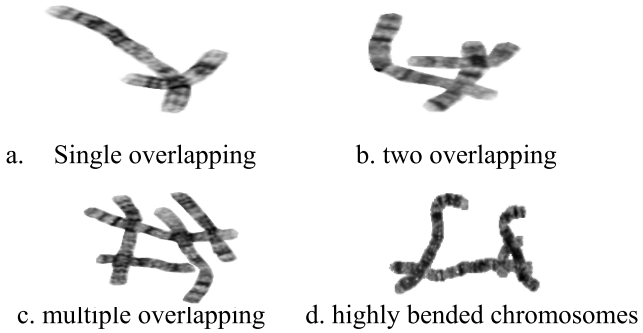

FIGURE 2. Overlapped chromosomes images.

FIGURE 3. Binary images of overlapped chromosomes.

which the separation lines are reduced. Hence achieving the overlap zone for overlapping chromosome and split line for touching chromosome.

When two chromosomes touch, the spatial gap between their vertices should be less than the distance between their vertices along the contour, as determined by the distance constraint of Equation 9.

$$\frac{\sqrt{(k-m)^2+(l-n)^2}}{d_{\min}\left(p_{k,l},p_{m,n}\right)} < \tau \tag{9}$$

where $d_{\min}\left(p_{k,l},p_{m,n}\right)$ is the minimum distance between $p_{k,l}$ and $p_{m,n}$ along the contour. τ is the threshold.

For overlapping chromosomes pair of separation lines are considered.

$$\left|\frac{(n_1-l_1)(k_2-k_1)-(l_2-l_1)(m_1-k_1)}{\sqrt{(n_1-l_1)^2+(m_1-k_1)^2}}\right| > \tau \tag{10}$$

Thus the separation of touching and overlapping chromosomes are separated by simple distance calculation function.

IV. EXPERIMENTS

The overlap zone can be identified by calculating the distance between the separation lines given by Equation

Compared to other methodologies addressed in the study, computational geometry, which is applied in this work, produces better results. Figure 2 displays the original photos used for analysis. As demonstrated in Figures 3 and 4, the input images are converted to binary and the image contour is obtained.

Using the corner finding approach for chain code curves that was previously mentioned, the curvature function is determined. Figure 5 displays the obtained concave and convex points. Concave points are represented by red dots in the

FIGURE 4. Contour images of overlapped chromosomes.

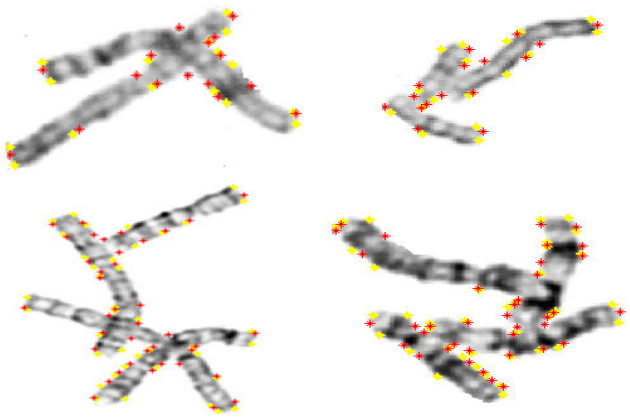

FIGURE 5. Concave and convex points on chromosome images.

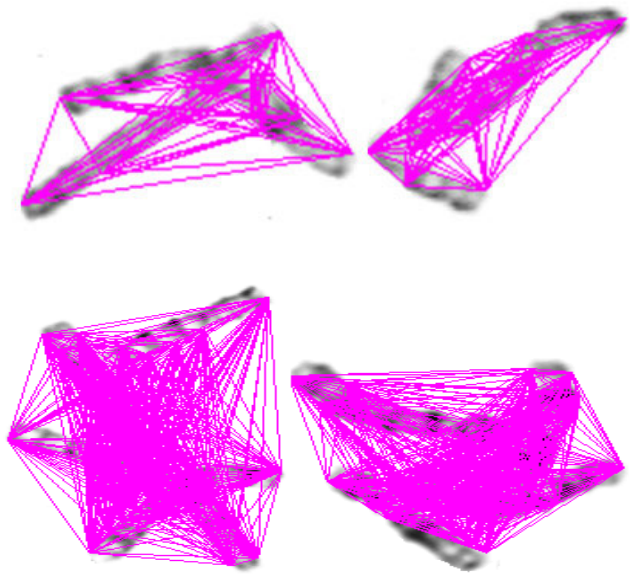

FIGURE 6. Separation lines on chromosome images.

photos, while convex points are represented by yellow dots. Concave points are used to create separation lines, as shown in Figure 6.

The overlap zone and split lines are obtained using the concave points and separation lines, as shown in Figure 7 The split lines (a separation line dividing the touching chromosome) of touching region with pink line connecting the concave point indicated with red dots and the two separation lines with two pink and two yellow parallel lines are used to identify the overlap zone.

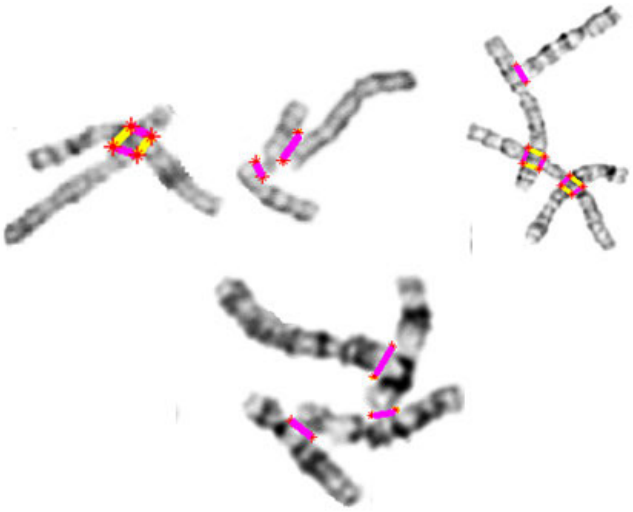

FIGURE 7. Overlap zone and split line.

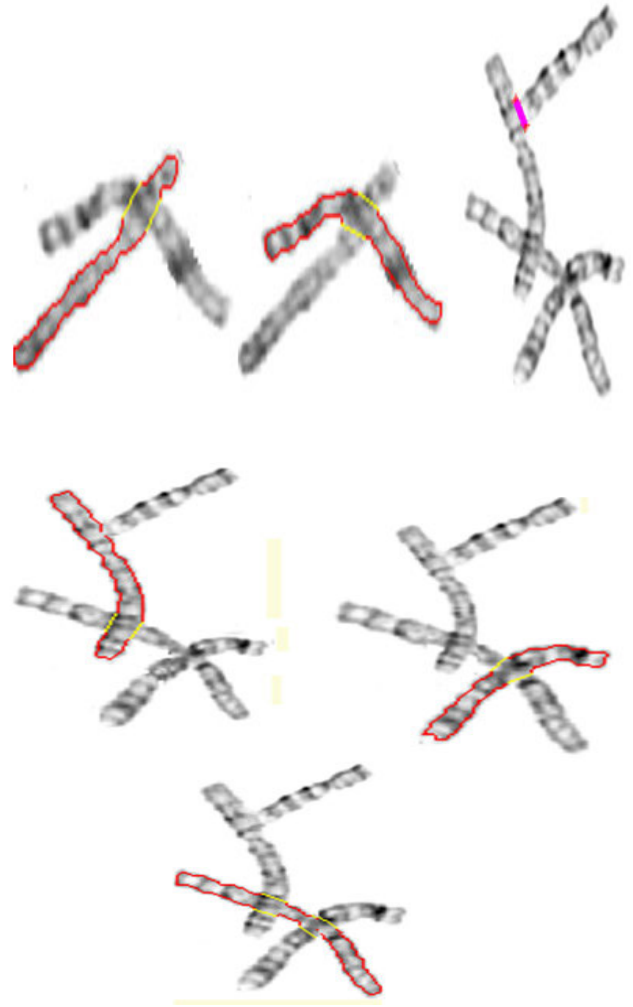

FIGURE 8. Segmentation region of chromosomes.

Figure 8, which illustrates the segmentation of each region of overlapping chromosomes, was created by measuring the distance between the two separation lines. The segmented full individual region is depicted by a red line, while the overlapped region is indicated in yellow. In Figure 9, the separation output is displayed.

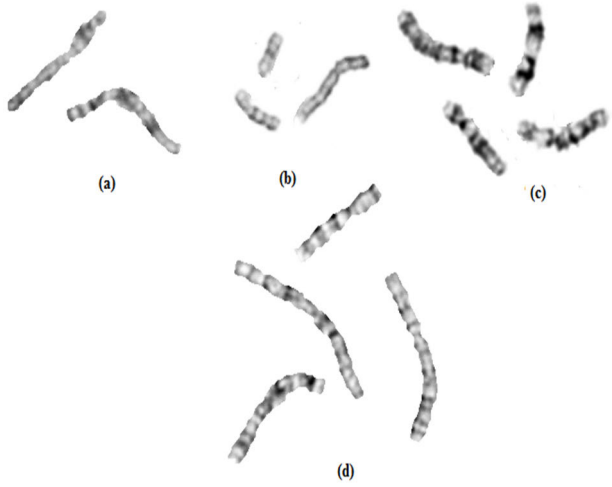

FIGURE 9. Separated chromosomes.

TABLE 4. Comparative table of overlapped resolved image.

No. of Images	No. of Overlaps	Number of Resolved Images with Accuracy (Acc)							
		Spline Interpolation		Delaunay Triangulation		Shape based Method		Computational Geometry	
		Resolved	Acc (%)	Resolved	Acc (%)	Resolved	Acc (%)	Resolved	Acc (%)
120	One	120	94.5	118	98.33	120	100	120	100
95	Two	95	94.6	85	89.47	88	92.63	91	99.79
62	≥ Three	62	85	45	72.58	54	87.10	57	95.94
277	-	**277**	**94.1**	**248**	**89.53**	**262**	**94.58**	**268**	**99.65**

TABLE 5. Comparative table of touching resolved image.

No. of Images	No. of Touches	Number of Resolved Images with Accuracy (Acc)							
		Spline Interpolation		Delaunay Triangulation		Shape based Method		Computational Geometry	
		Resolved	Acc (%)	Resolved	Acc (%)	Resolved	Acc (%)	Resolved	Acc (%)
135	One	135	95	122	90.37	129	95.56	131	99.04
101	Two	101	94	89	88.12	91	90.10	96	96.05
73	≥ Three	73	89	67	91.78	66	90.41	68	96.15
309	-	**309**	**93.3**	**278**	**89.97**	**286**	**92.56**	**295**	**98.33**

Tables 4 and 5 provide a comparison of each of the four strategies for overlapping and contacting chromosomes. The tables demonstrate that computational geometry and spline interpolation perform better in the separation of touching and overlapping chromosomal pictures. Computational geometry is preferred because it is a fully automated algorithm, in contrast to the spline interpolation method, which involves human participation in the control point selection.

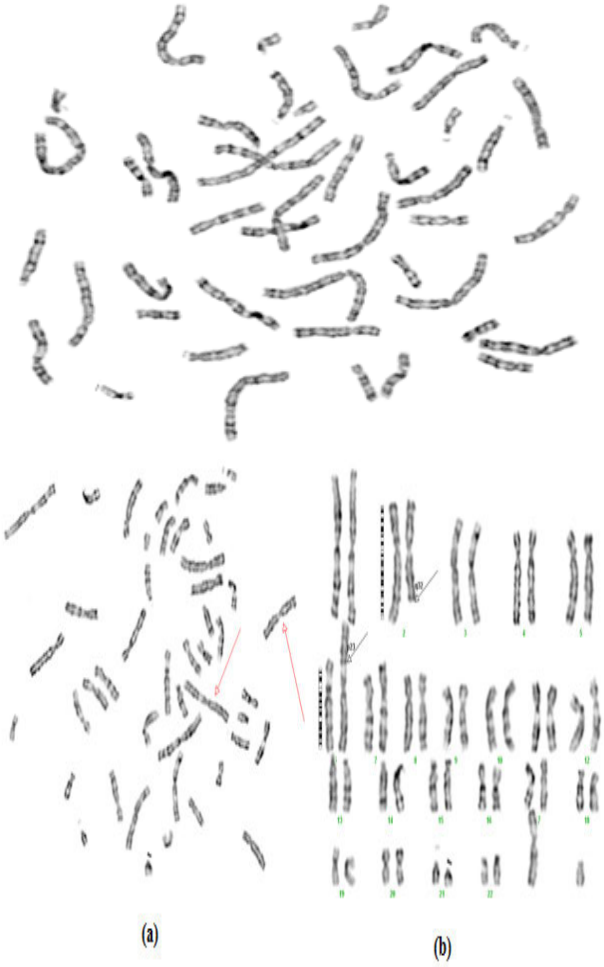

FIGURE 10. Metaspread test image to Matlab environment. a. Translocated chromosome, b. Karyotype image (from Physician).

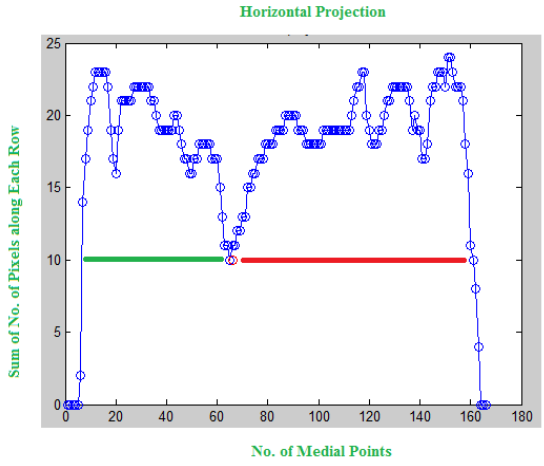

FIGURE 11. Horizontal projection vector for 1st chromosome of 2nd pair.

Copenhagen database images are used to train the neural networks. Moreover, own chromosomes metaspread are supported by Mediscan centre, Chennai and international institute of human genetics, Chennai with mutual ethical agreement.

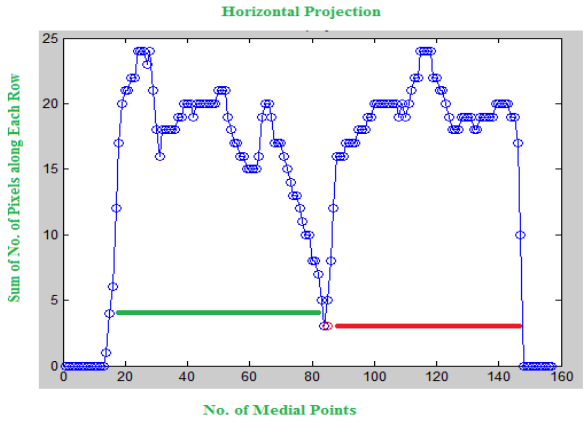

FIGURE 12. Horizontal projection vector for 2nd chromosome of 2nd pair.

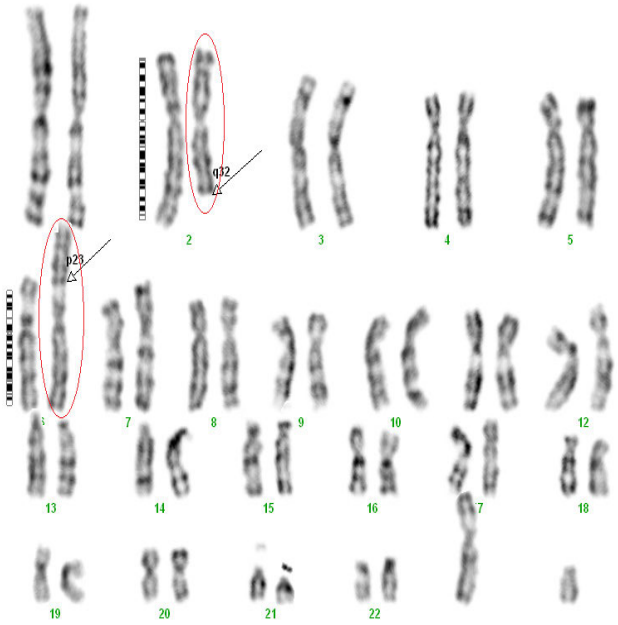

FIGURE 13. Structural abnormality identified in Matlab.

Convolutional neural network model is used to classify the chromosomes [3]. Analyses of parameters are performed on the images in Figure 10 a. Intensity profile computation is used to identify abnormalities. In Figure 10 b, the karyotyped images for the metaspread photos that were obtained from a doctor are displayed. The second pair of chromosomes' Q-arm and sixth pair of chromosomes' P-arm is involved in the issue. The projection vector profile for the second pair, which is depicted in Figure 11, effectively identifies this issue.

The projection vector for a typical chromosomal image is shown in Figure 11. Typically, the second pair is a submetacentric chromosomal picture. The graphed image shows a submetacentric chromosome, Figure11. However, when the second chromosome is taken into account, it is roughly a metacentric chromosome. The global minimum point, which is regarded as metacentric chromosomes, splits the medial points precisely into two sections as shown in Figure 12. However, the second pair must be submetacentric,

as shown in Figure 11, where the minimum point is moved from the centre. When comparing the two graphs, it can be seen that the second pair of chromosomes in the Karyotype image has structural irregularity.

V. CONCLUSION

Due to its automated nature, proposed computational geometry is superior for overlapping and contacting chromosomes. There are three different parameters found. The concave function aids in producing superior outcomes for the Centromere, which is carried out by other algorithms. Length and Centromere index are the additional factors that are calculated. The projection vector approach is then used to illustrate a scenario of a translocation issue connected to a structural anomaly. The proposed method successfully implemented in Matlab environment to carry out the karyotyping and abnormality analysis. The results are evaluated by physician to understand the suitability in the clinical laboratory. The proposed software provides 99.68 % of accuracy in karyotyping process. These results achieved based on the training and test datasets applied in CNN. The evaluation is carried out based on 10:1 image sharing datasets.

ACKNOWLEDGMENT

The authors would like to thank to VIT University for providing seed grant and publication support to this article.

REFERENCES

[1] P. Biyani, X. Wu, and A. Sinha, "Joint classification and pairing of human chromosomes," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 2, no. 2, pp. 102–109, Apr. 2005.

[2] D. Ming and J. Tian, "Automatic pattern extraction and classification for chromosome images," *J. Infr., Millim., THz Waves*, vol. 31, no. 7, pp. 866–877, Jul. 2010.

[3] D. Somasundaram and V. R. Vijay Kumar, "Separation of overlapped chromosomes and pairing of similar chromosomes for karyotyping analysis," *Measurement*, vol. 48, pp. 274–281, Feb. 2014, doi: 10.1016/j.measurement.2013.11.024.

[4] D. Somasundaram and V. R. V. Kumar, "Straightening of highly curved human chromosome for cytogenetic analysis," *Measurement*, vol. 47, pp. 880–892, Jan. 2014.

[5] E. Poletti, E. Grisan, and A. Ruggeri, "A modular framework for the automatic classification of chromosomes in Q-band images," *Comput. Methods Programs Biomed.*, vol. 105, no. 2, pp. 120–130, Feb. 2012.

[6] E. Grisan, E. Poletti, C. Tomelleri, and A. Ruggeri, "Automatic segmentation of chromosomes in Q-band images," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Lyon, France, Aug. 2007, pp. 5513–5516.

[7] G. Agam and I. Dinstein, "Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 11, pp. 1212–1222, Jan. 1997.

[8] G. H. Granlund, "Identification of human chromosomes by using integrated density profiles," *IEEE Trans. Biomed. Eng.*, vol. BME-23, no. 3, pp. 182–192, May 1976.

[9] H. Wang, H. Zhang, and N. Ray, "Clump splitting via bottleneck detection and shape classification," *Pattern Recognit.*, vol. 45, no. 7, pp. 2780–2787, Jul. 2012.

[10] H. Choi, A. C. Bovik, and K. R. Castleman, "Feature normalization via expectation maximization and unsupervised nonparametric classification for M-FISH chromosome images," *IEEE Trans. Med. Imag.*, vol. 27, no. 8, pp. 1107–1119, Aug. 2008.

[11] J.-H. Kao, J.-H. Chuang, and T. Wang, "Chromosome classification based on the band profile similarity along approximate medial axis," *Pattern Recognit.*, vol. 41, no. 1, pp. 77–89, Jan. 2008.

[12] J. Piper, "Genetic algorithm for applying constraints in chromosome classification," *Pattern Recognit. Lett.*, vol. 16, no. 8, pp. 857–864, Aug. 1995.

[13] J. M. Tho, "Chromosome classification using back propagation neural networks," *IEEE Eng. Med. Biol.*, vol. 19, no. 1, pp. 28–33, Mar. 2000.

[14] M. J. Roshtkhari and S. K. Setarehdan, "A novel algorithm for straightening highly curved images of human chromosome," *Pattern Recognit. Lett.*, vol. 29, no. 9, pp. 1208–1217, Jul. 2008.

[15] N. Gupta, M. Kabra, and D. Deka, "Genetics of birth defects," *Donald School J. Ultrasound Obstetrics Gynecol.*, vol. 4, no. 4, pp. 327–333, 2010.

[16] P. S. Karvelis, D. I. Fotiadis, M. Syrrou, and I. Georgiou, "Segmentation of chromosome images based on a recursive watershed transform," in *Proc. 3rd Eur. Med. Biol. Eng. Conf.*, Czech Republic, Prague, 2005, pp. 11(1)–11(5).

[17] P. S. Karvelis and D. I. Fotiadis, "A region based decorrelation stretching method: Application to multispectral chromosome image classification," in *Proc. 15th IEEE Int. Conf. Image Process.*, Feb. 2008, pp. 1456–1459.

[18] R. J. Stanley, J. M. Keller, P. Gader, and C. W. Caldwell, "Data-driven homologue matching for chromosome identification," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 451–462, Jun. 1998.

[19] S. Jahani and S. Kamaledin Setarehdan, "Centromere and length detection in artificially straightened highly curved human chromosomes," *Int. J. Biol. Eng.*, vol. 2, no. 5, pp. 56–61, Dec. 2012.

[20] S. Suresh, G. Thangavel, J. Sujatha, and S. Indrani, "Methodological issues in setting up a surveillance system for birth defects in India," *Nat. Med. J. India*, vol. 18, no. 5, pp. 259–262, 2005.

[21] R. C. Tripathi, "Separation of overlapped chromosome images from single chromosome images and detecting possibility of various diseases by karyotype," in *Proc. Int. Conf. Commun. Signal Process.*, India, Apr. 2013, pp. 951–955.

[22] S. Devaraj, V. R. Vijaykumar, and G. R. Soundrarajan, "Leaf biometrics based karyotyping of G-band chromosomes," *Int. J. Hum. Genet.*, vol. 13, no. 3, pp. 131–138, Sep. 2013.

[23] S. Rungruangbaiyok and P. Phukpattaranont, "Chromosome image classification using a two-step probabilistic neural network," *Songklanakarin J. Sci. Technol.*, vol. 32, no. 3, pp. 255–262, 2010.

[24] W. P. Sweeney, M. T. Musavi, and J. N. Guidi, "Classification of chromosomes using a probabilistic neural network," *Cytometry*, vol. 16, no. 1, pp. 17–24, May 1994.

[25] X. Bai, C. Sun, and F. Zhou, "Splitting touching cells based on concave points and ellipse fitting," *Pattern Recognit.*, vol. 42, no. 11, pp. 2434–2446, Nov. 2009.

[26] X. Yanga, W. Wua, and C. C. Tseng, "Algorithms for modeling structural changes in human chromosomes, *Comput. Methods Programs Biomed.*, vol. 110, pp. 171–182, Jan. 2013.

[27] X. Wang, B. Zheng, S. Li, J. J. Mulvihill, and H. Liu, "A rule-based computer scheme for centromere identification and polarity assignment of metaphase chromosomes," *Comput. Methods Programs Biomed.*, vol. 89, no. 1, pp. 33–42, Jan. 2008.

[28] X. Wang, B. Zheng, S. Li, J. J. Mulvihill, M. C. Wood, and H. Liu, "Automated classification of metaphase chromosomes: Optimization of an adaptive computerized scheme," *J. Biomed. Informat.*, vol. 42, no. 1, pp. 22–31, Feb. 2009.

[29] D. Somasundaram, "Machine learning approach for homolog chromosome classification," *Int. J. Imag. Syst. Technol.*, vol. 29, no. 2, pp. 161–167, Jun. 2019, doi: 10.1002/ima.22287.

DEVARAJ SOMASUNDARAM (Member, IEEE) was born in Tamilnadu, India. He received the bachelor's degree in electrical and electronics engineering from Anna University, Chennai, and the M.Tech. degree in VLSI design, in 2009. He was a CSIR Senior Research Fellow and a National Postdoctoral Fellow with the Indian Institute of Science, Bengaluru, in 2014 and 2016, respectively. He was also a Professor and the Head of the Department with SIET, Coimbatore, and CMR University. He is currently an Associate Professor with the Vellore Institute of Technology, Vellore, India. He completed three international funded projects sponsored by DST, ICMR, and AICTE. His research interests include artificial intelligence, VLSI architecture for biomedical applications, quantum computing, and genetics.

ADITYA CHAPPARADAHALLIMATH was born in Karnataka, India. He is currently pursuing the bachelor's degree in electronics engineering from the Vellore Institute of Technology, Vellore. His research interests include medical devices, the IoT, and real-time architectures in VLSI.

MURALITHARAN KRISHANAN (Member, IEEE) received the bachelor's degree in electronics and the master's degree in computer science from Bharathiar University, Tamil Nadu, India, in 2001 and 2004, respectively, and the Ph.D. degree in computer science from Anna University, Chennai, India, in 2017. He is currently a Senior Researcher with Sungkyunkwan University, South Korea. His current research interests include AI, such as ML, DL, and RL, and optimization techniques in the IoT, WSN, smart grids, and security.

NIRMALA MADIAN (Member, IEEE) was born in Tamilnadu, India. She received the bachelor's degree in electrical and electronics engineering from Anna University, Chennai, and the M.Tech. degree in applied electronics, in 2007. During the Summer of 2016, she was a Research Fellow with the Indian Institute of Science, Bengaluru. She is currently an Associate Professor with the Dr. N. G. P. Institute of Technology, Coimbatore, Tamilnadu. Her research interests include biomedical signal and image-processing genetics.