Microscopy Research and Technique

WILEY

MICROSCOPY RESEARCH & TECHNIQUE

RESEARCH ARTICLE

Chromosome Abnormality Detection Using Visual Geometric Transformer and Mantis Search Optimization

Nelliyadan Nimitha[1] | Periyathambi Ezhumalai[2] | Arun Chokkalingam[3]

[1]Research Scholar, Anna University, Chennai, India | [2]Department of Computer Science and Engineering, R.M.D. Engineering College, Chennai, India | [3]School of Engineering Vels Institute of Science, Technology & Advanced Studies (VISTAS), Vels University, Chennai, India

Correspondence: Nelliyadan Nimitha (nimithan.id@gmail.com; nimithasoman@gmail.com)

Received: 19 November 2024 | **Revised:** 16 April 2025 | **Accepted:** 18 June 2025

Review Editor: Mingying Yang

Funding: The authors received no specific funding for this work.

Keywords: chromosome abnormality detection | mantis search | VGG-16 model | vision transformer | visual geometric transformer

ABSTRACT

Chromosomes, which carry vital genetic material, have a distinctive thread-like appearance located within the cell nucleus. The process of examining these structures known as karyotyping is fundamental for identifying genetic abnormalities. Although several techniques have been developed for this purpose, many existing methods are limited by inefficiencies, particularly in terms of processing time and accurate feature extraction. To overcome these issues, this study introduces a novel algorithm called Visual Geometric Transformer-based Mantis Search (VGT-MS) for effective detection of chromosomal anomalies. Given that chromosome images often include irrelevant background elements, a preprocessing step is applied to eliminate these artifacts. Feature extraction is performed using the VGG-16 network, followed by classification using the Vision Transformer to pinpoint abnormalities. To further enhance the model's effectiveness, its parameters are optimized using the Mantis Search Algorithm. The performance of the proposed framework is assessed using evaluation metrics including accuracy, F1-score, recall, precision, and ROC. The experimental results indicate that the proposed model excels in all key metrics, achieving an accuracy of 98.0%, precision of 97.2%, recall of 96.2%, and an F1-score of 96.7%, all while reducing computational overhead. Overall, the VGT-MS framework proves to be a powerful and efficient solution for chromosome abnormality detection, successfully addressing the drawbacks of conventional methods.

1 | Introduction

Chromosome is the structure that contains genetic information in the human body. Chromosomes are of interest in chromosomal analysis and cell function and help in early detection of genetic diseases. A chromosomal abnormality is also called a chromosomal aberration. Chromosome karyotyping helps in the medical field to screen and diagnose genetic disorders. Chromosomes involve two stages namely segmentation and analysis (Al-Kharraz et al. 2020). Chromosome abnormalities are more common in early human embryos. Fertilizers are produced through artificial reproductive technologies (ART) or natural fertilization. The low success of ART and low human resources are attributed to cytogenetic defects (Viotti 2020). Diagnosis and analysis of chromosomes are of great importance because chromosomes are the basic carriers of genetic information. Chromosomal abnormalities can be identified through the use of karyotype. This is seen as the basis for diagnosing the disease (Zhao et al. 2024).

A cytogenetic test method is karyotype analysis, which is used to count chromosomes and evaluate their abnormalities. Counting

Summary

- VGT-MS algorithm detects chromosome anomalies with high accuracy and speed.
- Background artifacts are removed through an effective preprocessing step.
- VGG-16 and Vision Transformer used for precise feature extraction and classification.
- Mantis Search Algorithm optimizes model parameters for enhanced performance.
- Achieved 98.0% accuracy with reduced computational overhead.

chromosomes and determining the structure is manual and therefore time-consuming. Chromosome destruction occurs when the human body is exposed to radiation, causing abnormal chromosomes. Diseases like Edward's syndrome and Down's syndrome appear when chromosomes decrease or increase in the body (Kang et al. 2024). A chromosomal karyotype helps determine if a newborn has a genetic disorder. Chromosomes have two types of abnormalities, namely numerical abnormalities and structural abnormalities (Li and Zhang 2023). Karyotype analysis is used as an important field to identify chromosomal abnormalities (Joshi et al. 2023). Chromosome classification is seen as important in abnormal diagnosis for karyotyping. Numerical abnormalities are caused by the loss of a group of chromosomes, while structural abnormalities are caused by the segregation and fusion of chromosomes. The karyotyping process involved requires well-trained professionals (Remya et al. 2020).

Chromosome classification is an important part of karyotype analysis, the task of which is long and difficult (Viotti 2020). New methods such as screening have been used to detect chromosomal abnormalities. This method is very useful for patients with congenital defects. This does not increase the risk of miscarriage and can be done early in pregnancy. Additionally, they are relatively inexpensive and found to be quick (Goh et al. 2023). Chromosomes contain genetic information in human beings. Karyotyping of these chromosomes is done only by well-trained doctors. However, it is found to be time-consuming and difficult. Apart from this, there is inconsistency in the process due to overload and burnout of practitioners due to domain expertise (Qin et al. 2019). Chromosome abnormality occurs due to the malfunction of cell separation. If a cell contains the wrong number of DNA, chromosomal abnormality arises. It can be smoothly transmitted by progeny, particularly when utilizing assisted fertilization. Turner syndrome, Down syndrome, trisomy, and Klinefelter syndrome are some of the syndromes caused due to abnormal chromosomes. Chromosome abnormalities are detected by deep learning, machine learning, several algorithms, optimization, and various traditional methods. However, these methods face several limitations in terms of the number of datasets, accuracy, and so forth. This paper developed a VGT-MS algorithm that aims to enhance performance rates. The novelty of chromosomal abnormalities detection is described as follows:

- **Integration of new approaches:** To perform effective chromosomal abnormalities detection, the integration of approaches like the Vision Transformer and VGG-16 are integrated. With this integration, a new approach called the VGT-MS method is proposed here; this integration can enhance the high-efficiency results for chromosomal abnormalities detection.
- **Innovative Layers:** For effective classification, the *Integration Layer* and final classification layer are utilized. To obtain the final classification, the dense layer and softmax layer are included.

The contribution of the developed algorithm is indicated as follows:

- **Novelty in Chromosome abnormality detection:** The proposed VGT-MS is implemented in the detection process and enhances effectiveness by including an initial strategy-based Mantis Search Algorithm.
- **Image Processing:** Chromosome karyotype images from the CRCN-NE dataset are preprocessed to reduce noise and improve image clarity, thereby enhancing the quality and reliability of the input data.
- **Feature Extraction and Detection:** The VGG-16 model efficiently extracts local visual features from chromosome images, while the Vision Transformer captures long-range dependencies to accurately detect structural abnormalities.
- **Effectiveness validation:** The metrics namely accuracy, recall, ROC, F1-score, and precision measures are validated for the VGT-MS approach to validate the detection performance.

The remainder of this paper is structured as follows: Section 2 reviews the related literature in chromosome abnormality detection, Section 3 provides a detailed description of the proposed VGT-MS methodology, Section 4 presents the experimental setup, results, and analysis, and Section 5 concludes the study while outlining potential directions for future research.

2 | Related Works

2.1 | Chromosome Abnormality Detection Using Deep Learning

Raza et al. (2024) introduced a transfer learning-dependent characteristics extraction called VNL-Net that includes Light Gradient Boosting Machine (LGBM), Non-Negative Matrix Factorization (NMF), and VGG16 methods for early Down syndrome diagnosis. The spatial characteristics were separated from the original image. Then, the set of features of LGBM and NMF was separated from the features in spatial. The complexity of computation was determined. The method produced a better diagnosis of chromosomal abnormalities at an earlier stage. Still, a Graphical User Interface (GUI) was not utilized for early Down syndrome detection.

Kang et al. (2024) established deep learning-based object identification for analyzing of chromosomes automatically. The chromosomal abnormalities and counting of chromosomes were involved in this method. The excitation and squeeze block and

convolutional spatial pooling attention were implemented. The probability of a chromosome was 99.32% and the probability of an unusual chromosome was 75.71%. Rather, there was a need to develop the representation of the image.

He et al. (2023) illustrated a feature fusion classifier with dynamic weights (FFCDW) for the detection of abnormalities in chromosomes. The models like SENet, VGG19, and ResNet were combined by a dynamic weighting system. The result was obtained with 0.8805 of F1-score, 0.00892 and 0.00903 of standard deviation, 0.8902 of precision. The dynamic weights strategy outperformed the fixed weights strategy. However, the commonly available datasets were not utilized for comparison with other approaches.

Bechar et al. (2023) implemented a Convolutional neural network (CNN) with Siamese architecture for the detection of abnormalities of chromosomes that were present in the pair of copy images. First, the occurrence of chromosome 5 of hematological malignancies was deleted. The dataset with and without augmentation of data was done on the famous seven models of CNN. The detection results produced on F1-score for the models InceptionResNetV2, inversion inv. (3) and Xception were 97.01%, 94.82%, and 97.50%, respectively. The method showed better detection performance. Conversely, the method did not handle more copies of images, which was one of the drawbacks.

Li et al. (2023) introduced a masked conditional variational autoencoder (MC-VAE) with a prior processing program and a generative method for straightening chromosomes to detect the type of chromosome. The rearrangement of the patch was implemented in the processing stage to obtain prior results. The model MC-VAE enhanced the patches of chromosomes based on curves to straighten the outcomes. A masking scheme was applied in the training stage without redundancy. This method produced better results on classification and provided exact karyotyping. However, this approach was not enhanced on generalizability.

Wang et al. (2023) illustrated a Transformer model-based Down-Syndrome-Detector (DSD) design containing an alignment unit, a segmentation unit, a Down syndrome guide, and a classification unit for the detection of Down syndrome. The units of classification and segmentation were made by transfer learning to distribute and enhance the weights. The categorization dataset was fitted through an alignment unit, and the case of Down syndrome was identified through a Down syndrome guide. The Advanced Digital Imaging Research and BioImLab were the datasets used for evaluation. The method obtained better results in the detection of Down syndrome. Yet, in the case of singletons, the cells were inadequate.

Uzolas et al. (2022) established a conditional Generative Adversarial Network (cGAN), an approach to image conversion for blending images of chromosomes that were banded abnormally. The patterns of banding were ended by 2-dimensional chromosome separation tag maps. The outcome showed better blending of chromosomes. This method tacked issues like data detection, segmentation, classification, and simulation. Yu et al. (2023) introduced a transformer-directed and a solo-cell variety-cognizant deep learning approach for aneuploidy prediction from entire slide images of histopathology. The steps were: First, the nuclei classification and segmentation. Then, cancer cells were clustered. The morphological characteristics of various subtypes of cells were separated. A Transformer with multiple instance learning models was implemented for cancer cell data. The outcome showed better prediction performance. Despite that, the method required a more accurate prediction.

Yang et al. (2023) illustrated a deep convolutional neural network (DCNN)-dependent technique to classify abnormal and normal types of chromosomes. The accuracies obtained on eight kinds of formational abnormalities were between 90.84% and 100% and the AUC values were between 91.81% and 100%. The outcome showed effective performance on prediction for unusual chromosomal identification and genetic screening with no beginning feature separation. Conversely, data imbalance and variation might occur in the given dataset.

Al-Kharraz et al. (2020) implemented a deep-learning technique for recognizing statistical abnormalities. The 147 without overlaid metaphase images was the dataset was analyzed in three phases. First, every chromosome was detected by YOLOv2 Convolutional Neural Network and post-processing. Second, features were classified and extracted by tuning of VGG19 network. The result gained 96.67% accuracy in abnormality identification. A 94.11% accuracy was also achieved by using the Biomedical Imaging Laboratory dataset. However, there was a requirement to increase the size of training for accuracy improvement.

2.2 | Chromosome Abnormality Detection Using Machine Learning

Goh et al. (2023) designed a machine learning method using an altered log R ratio for chromosomal Copy number variation (CNV)-related disabilities management. There were 16,046 newborn Korean samples taken for analysis. Out of 342 data, there were identified 39 CNV-related genetic disorders. The Joubert syndrome 4 was the highly identified CNV-related genetic disorder. Through NGS, the accuracy was analyzed and compared. The wave offset with a single array of nucleotide polymorphism was utilized to identify CNVs from neonatal data. The screening and identification of susceptibilities of different diseases were facilitated for the detection of CNV-related chromosomal disease reasons. Conversely, the overfitting risk was one of the drawbacks of this method.

Huang et al. (2024) implemented a machine learning-based pipeline for nonintrusive prenatal experiments to detect the trisomy of a fetus accurately using minimum-range short-series information. The 40 bp arrangement at one end, calculations of P_{chrN}, and logistic regression (LR) were the steps followed for the detection. The commonly available datasets were utilized for training on 314 samples and selecting features. The performance was compared with the bioinformatics approach. The method produced accurate detection of trisomy. Still, there was a possibility that false-positive and false-negative experimental results would occur.

2.3 | Chromosome Abnormality Detection Using Optimization

Xia et al. (2023) implemented a back-to-back diverse combinational optimization technique, called KaryoNet for recognition of chromosomes. The two medical datasets G-band and R-band showed 99.58% and 98.41% accuracy. The KaryoNet produced better results in patients with various kinds of statistical abnormalities. Nevertheless, there was a lack of standardization and validation.

Shao et al. (2024) established a copy number variants sequencing (CNV-Seq) for the detection of abnormalities in chromosomes during miscarriages. 580 cases of miscarriage with medical information were taken for analysis. In that, 357 showed unusual chromosomal outputs, and the rest were resulted as normal. Also, 470 cases of the variant were found, which include statistical abnormalities in 251 cases and formational abnormalities in 219 cases. The interaction of protein of 226 doses-perceptive genes was analyzed. The result produced better identification performance during miscarriages. However, CNV-Seq did not identify uniparental dismay or stabilized translocation.

2.4 | Chromosome Abnormality Detection Using Algorithm

Jochems et al. (2023) addressed the origin and incidence of errors in chromosomes. 64 vitro embryo production (IVP), and 101 vivo-derived (IVD) were taken for evaluation. Huge errors were found in IVD vs. IVP blastocysts. Less error was found in the blastocyst phase in comparison with cleavage. The embryos of 2 parthenogenetic and 1 androgenetic were detected. The Triploid error was found in IVD embryos (15.8%) and chromosome aneuploidy (9.9%). The result showed better error detection performance. Further improvement was needed on this methodology for the successful transfer of embryos in the animal pig.

2.5 | Research Gap

Despite the availability of various intelligent and reliable approaches for the detection of chromosome abnormalities, some of these gaps include:

- **Insufficient accuracy prediction:** Many existing models are available for accuracy prediction in the detection of chromosome abnormalities, but some of them still need more prediction of accuracy. The proposed model addresses this gap and various measures were adopted to obtain better accuracy.
- **Limited dataset utilization:** In existing methods, only a limited number of commonly available datasets are used. The proposed method addresses this gap by utilizing a large number of datasets that effectively enhance the performance.
- **False indication:** In certain cases, during accurate chromosome abnormalities detection using a minimum sequence of data, the experiment exposes false information such as false-positive and false-negative results. Hence, the model utilizes a more efficient and effective approach to rectify the false indication.
- **No standardization and validation:** Many existing models for the detection of chromosome abnormalities lack standardization and validation on output. Hence, proper standardization and validation was carried out by the proposed method to beat this gap.

The VGT-MS algorithm addresses key limitations of prior transformer-based and deep learning approaches in chromosome abnormality detection by introducing a synergistic architecture that tightly integrates convolutional and attention-based representations with meta heuristic optimization. Unlike conventional CNN or pure transformer models that often struggle with fine-grained structural variations in chromosomes, VGT-MS leverages VGG-16 for capturing local spatial patterns and Vision Transformer (ViT) for modeling long-range dependencies. Moreover, it incorporates the Mantis Search (MS) algorithm to dynamically optimize key hyperparameters and enhance feature selection, ensuring more precise boundary localization and class-specific discrimination, especially in challenging cases like ring or translocation abnormalities. This unified approach allows VGT-MS to overcome issues like vanishing gradients, poor generalization on rare abnormality types, and limited boundary precision, which earlier models either neglected or addressed partially. On the whole, the proposed model presents an innovative approach to addressing the gaps in the existing literature on chromosome abnormalities detection. By introducing critical detection measures and utilizing efficient data optimization techniques, the proposed model presents a more effective and efficient approach to chromosome abnormalities detection.

3 | Proposed Methodology

The proposed Visual Geometric Transformer-based Mantis Search (VGT-MS) algorithm, as mentioned in Figure 1, significantly enhances the analysis of microscopy images in chromosome abnormality detection through a hybrid deep learning and meta-heuristic optimization framework. Microscopy images of chromosomes often present challenges such as high variability in structure, subtle abnormal features, and the presence of noise or artifacts. To address these, the VGT-MS algorithm begins with a robust pre-processing pipeline that includes noise elimination using median filters, adaptive contrast enhancement via homomorphic filtering, and normalization techniques to standardize pixel intensity. Following pre-processing, the VGG-16 model is employed for feature extraction, effectively capturing local spatial patterns such as banding sequences and centromere positions. These features are then passed into a Vision Transformer (ViT) architecture, which introduces a global attention mechanism, enabling the model to discern complex interrelationships across the image space. This is particularly important in chromosome analysis, where abnormalities might manifest through long-range dependencies or contextual deviations. To further refine the detection process, the Mantis Search Algorithm, enhanced with an initial search strategy, is used for hyperparameter

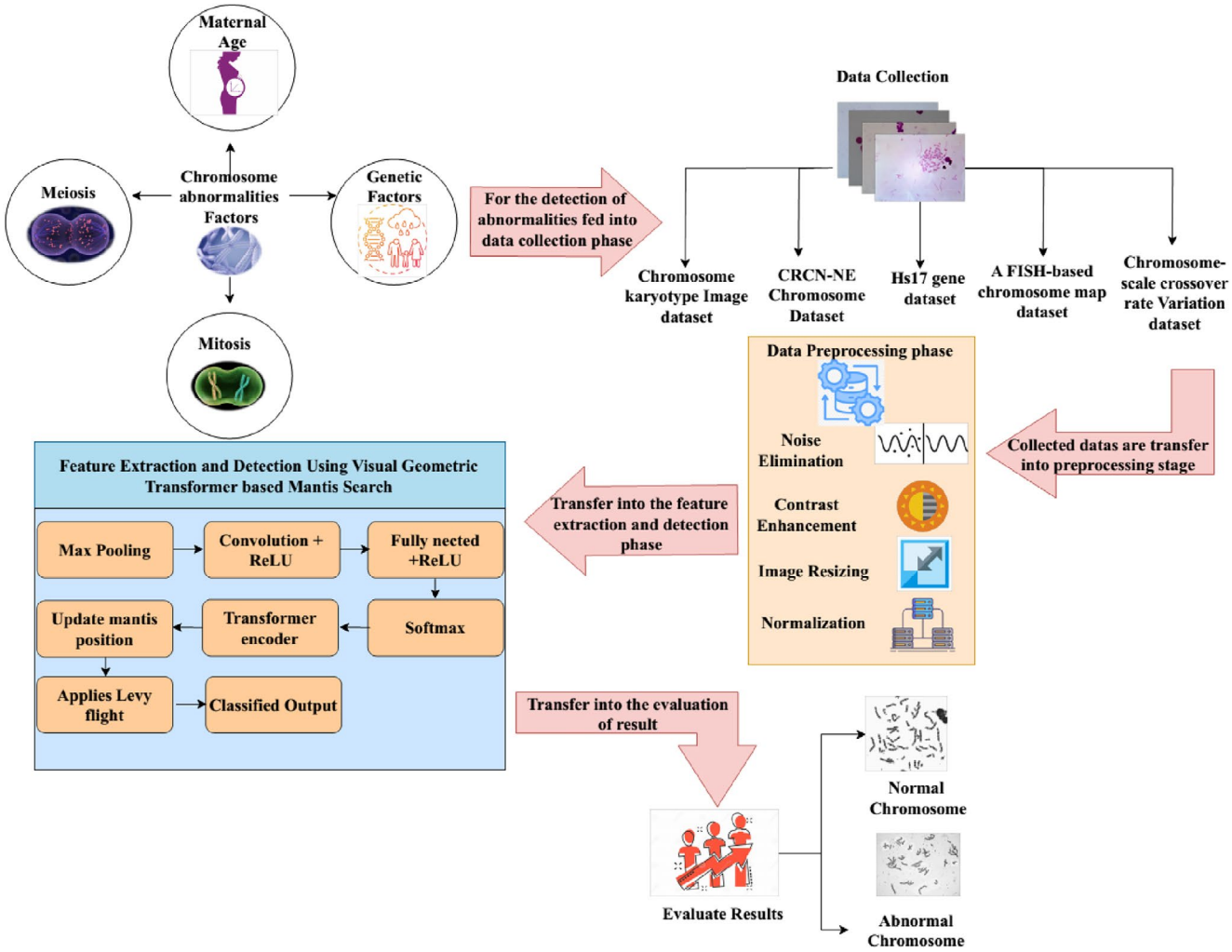

FIGURE 1 | Proposed architecture.

optimization. This component improves convergence speed and prevents the model from getting trapped in local minima, thereby ensuring optimal tuning of both VGG-16 and ViT parameters. Overall, the integration of VGG-16 for hierarchical feature learning, ViT for capturing global dependencies, and the enhanced Mantis Search for efficient optimization makes VGT-MS highly effective in accurately detecting chromosomal abnormalities from complex microscopy images.

3.1 | Pre-Processing

The dataset images often contain noise, poor contrast, and inconsistent resolution, which can hinder accurate chromosome abnormality detection. To ensure high-quality input, several pre-processing techniques are applied.

3.1.1 | Noise Elimination

Noise is defined as unwanted variations in pixel values that obscure image details. A median filter with an odd-sized sliding window is employed to reduce impulse noise. Each pixel is replaced with the median of its surrounding values, preserving edges while filtering out random fluctuations.

3.1.2 | Contrast Enhancement

To enhance image clarity, especially in low-contrast medical images with complex backgrounds, adaptive gamma correction is applied. However, to avoid amplifying noise in sensitive areas, homomorphic filtering is used in tandem. This method adjusts image illumination and reflectance components to highlight important structures without introducing artifacts.

3.1.3 | Image Resizing and Blurring

Images are resized to a uniform dimension to ensure compatibility with deep learning models. To further smooth the image and reduce noise, Gaussian blur is applied. This uses the 2D Gaussian function, derived from two 1D Gaussian functions, to suppress low-intensity edges and fine details that may not contribute to classification.

$$\min \ -\max \ = \frac{b_c - \min(b)}{\max(b) - \min(b)} \tag{1}$$

From the given equation b_c represents pixel intensity value, while the minimum and maximum values are denoted by $\min(b)$ and $\max(b)$ respectively.

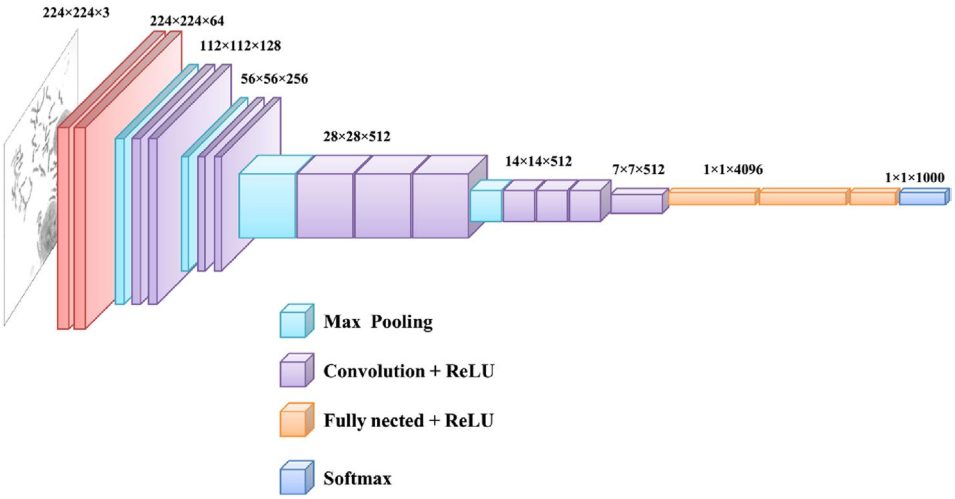

FIGURE 2 | VGG-16 architecture.

3.1.4 | Normalization

Normalization ensures consistent pixel value distribution across all images, improving training stability and convergence speed. Techniques such as Z-score normalization, empirical scaling, and min–max normalization are employed to scale pixel intensities. The general equation standardizes pixel values to a desired range, improving model performance and learning efficiency.

In this study, data augmentation was employed to enhance the robustness and generalization capability of the VGT-MS model by synthetically increasing the diversity of chromosome image samples. Various augmentation techniques were applied, including random rotations, horizontal and vertical flips, slight zooming, brightness adjustments, and elastic deformations, which simulate realistic variations encountered in microscopic imaging. These transformations help the model learn invariant features, reduce overfitting, and improve performance across both common and rare chromosomal abnormalities. By expanding the training dataset in this manner, the model becomes more resilient to noise, positional shifts, and morphological variations, thereby improving its accuracy and reliability in real-world diagnostic scenarios.

3.2 | Feature Extraction and Detection Phase

In the detection of chromosome abnormalities, feature extraction is pivotal in identifying and isolating relevant patterns from genetic imagery. Effective feature extraction significantly enhances the accuracy of abnormality detection by enabling the classification of genetic disorders and the recognition of ambiguous or rare chromosomal patterns. In this work, the VGG-16 model is utilized due to its proven capability in extracting hierarchical spatial features. VGG-16 offers several advantages, including robust community support, architectural simplicity, adaptability across domains, and strong performance in image classification tasks. Its deep convolutional structure enables the model to detect a wide range of chromosomal abnormalities with high precision. (Jahangeer and Rajkumar 2021).

3.2.1 | VGG-16 Model

CNNs, also known as ConvNets, are a fundamental class of Artificial Neural Networks widely used in computer vision tasks. Among these, VGG-16 stands out as one of the most influential and commonly adopted models for image classification. The term "VGG-16" denotes a deep CNN architecture comprising 16 weight layers, specifically designed for image recognition and classification tasks. The complete architecture consists of 21 layers in total, which include 13 convolutional layers, 5 max-pooling layers, and 3 fully connected (dense) layers. Its consistent use of small receptive fields (3×3 convolutional filters) and a deep network structure enables effective learning of both low-level and high-level image features. Figure 2 illustrates the detailed architecture of the VGG-16 model as implemented in this study. To validate the VGG-16 model for effective feature extraction in chromosomal abnormality detection, transfer learning was employed by initializing the model with pre-trained weights on the ImageNet dataset. The network was then fine-tuned on domain-specific chromosomal images, allowing it to adapt to relevant low-level and mid-level visual patterns. A validation split was maintained during training to monitor performance and ensure generalizability. To avoid overfitting, early stopping was applied based on the validation loss plateau, and dropout layers were used in the fully connected sections. Data augmentation techniques, including rotation, zoom, flipping, and shift transformations, were also utilized to increase dataset variability and improve model robustness against variations in chromosome morphology.

3.2.2 | Vision Transformer

Designed specifically for computer vision tasks such as image classification, the Vision Transformer (ViT) is a powerful architectural framework that leverages transformer models traditionally used in natural language processing (Yin et al. 2022). The ViT divides an input image into fixed-size patches, flattens them, and feeds them into a transformer encoder, allowing the model to capture long-range dependencies and contextual relationships between image regions. This method is particularly effective for detecting complex patterns in chromosome images, enabling precise prediction of abnormalities. The core

processing and prediction flow of the ViT is mathematically represented as follows:

$$x = \xi \circ K^{L} \circ K^{L-1} \circ \ \ldots \ \ldots \circ K^{1} \circ \varepsilon(y) \tag{2}$$

Whereas, $\varepsilon(\bullet)$ indicates the network encoder, which processes the input image by tokenizing the image into positional patches represented as embedded tokens. Here, l represents the output of the lth layer within the transformer block, while $l-1$ corresponds to the output of the preceding layer. Each layer processes contextualized patch embeddings using multi-head self-attention and feed-forward networks to progressively refine the representation.

$$u_{1:L}^{l} = K^{l}\left(u_{1:L}^{l-1}\right) \tag{3}$$

From Equation (3), $u_{1:L}^{l}$ indicates the updated token representation computed as $u_{1:L}^{0} = \varepsilon(y)$ with representing the feed-forward network. This formulation streamlines the halting mechanism in Transformers, contrasting with CNNs that require explicit management of varying architectural dimensions. Additionally, a data-dependent halting score is computed for each token, acting as a halting probability that guides the dynamic depth of computation across tokens.

$$h_{k}^{l} = H\left[u_{k}^{l}\right] \tag{4}$$

From Equation (4), $H(\bullet)$ denotes the halting module, which terminates the processing of a token when its halting score exceeds a predefined threshold $1-\varepsilon$

$$M_{k} = \arg\min_{m \le N} \sum_{l=1}^{m} h_{k}^{l} \ge 1 - \ \in \tag{5}$$

A single entry from the fixed-dimensional vector is used to compute this halting score. To track the token's progression, a residual value is calculated for each token by aggregating halting probabilities across multiple layers, enabling adaptive computation depth based on the complexity of the input features.

$$s_{k} = 1 - \sum_{l=1}^{M_{k}-1} h_{k}^{l} \tag{6}$$

The limit range of the ViT refers to its capacity to capture long-range dependencies across the input image. It is mathematically expressed as:

$$L_{overall} = L_{task} + \alpha p L_{ponder} \tag{7}$$

The behavior of Vision Transformers in the proposed system involves two critical components.

Integration Layer: This layer consolidates the feature outputs extracted by the preceding transformer blocks, effectively capturing the spatial and contextual relationships necessary for identifying chromosomal abnormalities.

Final Classification Layer: This stage employs a dense layer followed by a softmax activation function to perform the final classification, determining the specific type of abnormality present in the chromosome image.

3.2.3 | Mantis Search Model

In this research, an initial search strategy is introduced to enhance the numerical evaluation process and effectively manage the challenges associated with velocity boundaries, as discussed in recent advancements (Abdel-Basset et al. 2023). This strategic initialization helps in improving convergence reliability and solution quality in the optimization space.

- **Exploration of Pursuers' behavior**

In nature, mantis predators adopt various step sizes while hunting, which enable them to effectively explore hidden or camouflaged prey zones. Mimicking this, the algorithm utilizes a diversified step size mechanism to explore the solution space efficiently. The following equation describes the food-search behavior of the mantis in a stochastic manner, representing their adaptive exploratory strategy:

$$\vec{y}_{j}^{\,u+1} = \begin{cases} \vec{y}_{j}^{\,u} + \vec{\tau}_{1} * \left(\vec{y}_{j}^{\,u} - \vec{y}_{b}^{\,u}\right) + \left|\tau_{2}\right|.\vec{V} * \left(\vec{y}_{b}^{\,u} - \vec{y}_{c}^{\,u}\right), & s_{1} \le s_{2} \\ \vec{y}_{j}^{\,u} * \vec{V} + \left(\vec{y}_{b}^{\,u} + \vec{s}_{3} * \left(\vec{y}_{c}^{\,u} - \vec{y}_{d}^{\,u}\right)\right) * \left(1 - \vec{V}\right), & Otherwise \end{cases} \tag{8}$$

Hadamard product with jth mantis is $\vec{y}_{j}^{\,u}$, $\vec{s}_{3}$ is a numerical values vector, $\vec{V}$ is a binary vector, and $\vec{\tau}_{1}$ is indicated as the numerical vector, and $\left|\tau_{2}\right|$ is a random number.

- **Exploration of Spearers' behavior**

The formation of camouflaged position archives for simulating the predator.

$$\vec{y}_{j}^{\,u+1} = \vec{y}_{j}^{\,u} + \alpha.\left(\vec{y}_{bs}^{\,'} - \vec{y}_{b}^{\,u}\right) \tag{9}$$

The ambush distance is formulated as.

$$\vec{y}_{j}^{\,u+1} = \vec{y}_{bs}^{\,'} + \left(s_{7} * 2 - 1\right) * \mu * \left(\vec{y}^{\,m} + \vec{s}_{8} \times \left(\vec{y}^{\,v} - \vec{y}^{\,m}\right)\right) \tag{10}$$

For representing the $j-th$ mantis, from the archive, solutions $\vec{y}_{bs}$ is chosen, a vector with a range of [0, 1] is specified as $\vec{s}_{8}$, and the upper and lower bounds are $\vec{y}^{\,u}$ and $\vec{y}^{\,m}$.

$$\vec{y}_{j}^{\,u+1} = \begin{cases} \vec{y}_{j}^{\,u} + \alpha.\left(\vec{y}_{bs}^{\,'} - \vec{y}_{b}^{\,u}\right), & s_{9} \le s_{10} \\ \vec{y}_{bs}^{\,'} + \left(s_{7} * 2 - 1\right) * \mu * \left(\vec{y}^{\,m} + \vec{s}_{8} \times \left(\vec{y}^{\,v} - \vec{y}^{\,m}\right)\right) & Otherwise \end{cases} \tag{11}$$

From the above equation, the numerical values are s_{9} and s_{10}.

- **Striking velocity**

The hunting process of a mantis culminates in a precise and swift strike once the prey is within reach. To simulate this behavior computationally, the striking velocity magnitude of a mantis during its final approach is modeled mathematically to reflect sudden, directed movement toward the optimal solution. This is crucial for intensification in the search phase, ensuring

convergence to an optimal or near-optimal point. The striking velocity is defined as:

$$w_t = \frac{1}{1 + f^{m\rho}} \tag{12}$$

The acceleration rate is ρ, and its strike velocity is w_t,

$$y_{j,k}^{u+1} = \left(y_{j,k}^{u} + y_k^*\right) / 2.0 + w_t . e_{tj,k}^{u} \tag{13}$$

Where an updated position is indicated by $y_{j,k}^{u+1}$, $y_{j,k}^{u}$ represents the position of prey. Additionally, $e_{tj,k}^{u}$ is determined as

$$e_{tj,k}^{u} = \left(y_k^* - y_{j,k}^{u}\right) \tag{14}$$

y_k^* denotes the position with respect to $k - th$ dimension and the current position $\vec{y}_{j,k}^{\,u}$ is

$$y_{j,k}^{u+1} = y_{j,k}^{u} + s_{12} . \left(y_{b,k}^{u} - y_{c,k}^{u}\right) \tag{15}$$

In this scenario, y_b^u and y_c^u denotes randomly selected mantises, while s_{12} is a randomly chosen number between 0 and 1 respectively.

$$y_{j,k}^{u+1} = y_{j,k}^{u} + f^{21} . \cos(2m\pi) . \left|y_{j,k}^{u} - \vec{y}^{\,`}_{bs,k}\right| + \left(s_{13} . 2 - 1\right) . \left(y_k^v - y_k^m\right) \tag{16}$$

In this context, s_{13} represents numerical value.

$$Q_g = b . \left(1 - \frac{u}{U}\right) \tag{17}$$

Where, b is a predetermined constant.

- **Sexual cannibalism**

In nature, female praying mantises are known to consume the male after or during mating a—behavior rooted in maximizing reproductive success. In optimization, this phenomenon is modeled to symbolize eliminating weaker solutions to preserve stronger genetic material and intensify the evolutionary search. This behavioral analogy is mathematically represented as:

$$\vec{y}_j^{\,u+1} = \vec{y}_j^{\,u} + \vec{s}_{16} * \left(\vec{y}_j^{\,u} - \vec{y}_b^{\,u}\right) \tag{18}$$

$\vec{y}_b^{\,u}$ denoted as the solution, the female praying mantis is $\vec{y}_j^{\,u}$. Additionally, $\vec{s}_{16}$ is a vector with random values.

3.3 | Proposed (VGT-MS) Algorithm for Chromosome Abnormality Detection

The VGT-MS algorithm is designed to effectively detect chromosomal abnormalities by integrating advanced preprocessing, deep feature extraction, and intelligent parameter tuning strategies. Initially, preprocessing techniques are employed to eliminate noise and enhance the image quality. The VGG-16 model is then utilized to extract local and fine-grained patterns through its convolutional and pooling layers. These features are passed to a Vision Transformer (ViT) encoder, which further captures global contextual representations necessary for accurate classification of chromosomal abnormalities.

To improve the detection performance, parameter tuning is critical. Therefore, an Initial Strategy-based Mantis Search Algorithm (MSA) is integrated into the framework to fine-tune the parameters of the transformer module. Traditional MSA may suffer from premature convergence and getting trapped in local optima. To address this limitation, an initial exploratory strategy is introduced, which enhances the diversity of the search process and ensures a better balance between exploration and exploitation. This modification significantly improves the convergence rate and robustness of the optimization algorithm.

The overall workflow of the VGT-MS algorithm is depicted in Figure 3. The process begins with input chromosome images from benchmark datasets, which are passed through VGG-16 for local feature extraction. These features are then processed by the transformer encoder. If the stopping condition based on halting score or maximum iteration is met, the detection output is generated. Otherwise, the mating and positional update mechanism governed by the MSA continues. The initial strategy-based MSA dynamically refines the parameter space using biologically inspired behaviors and improves the optimization process.

The mathematical formulation of the initial strategy-based Mantis Search Algorithm is provided in the subsequent section, detailing how the mantis hunting behavior, cannibalism, and mating dynamics are translated into optimization steps.

$$x_{ak}(w+1) = x_{ak}(w) + |h_1|\left(x_{ak}^q(w) - x_{ak}(w)\right) + |h_2|\left(x_k^h\left(w - y_{ak}(w)\right)\right) \tag{19}$$

$$x_{ak}(0) = y_{\min\,k} + t_{ak}\left(x_{\max\,k} - x_{\min\,k}\right) \tag{20}$$

From the above equations, $x_{\min\,k}$ denotes the lower bound of the variable x_k, $x_{\max\,k}$ indicate upper bound, and t is an accidental number.

4 | Experimental Results

The experimental evaluation of the VGT-MS algorithm was conducted using publicly available chromosome datasets, including CRCN-NE and karyotype image repositories. The performance was validated using standard metrics to confirm the model's high reliability in detecting chromosomal abnormalities by comparing with existing techniques.

4.1 | Experimental Setup

The implementation is performed by Intel Core i9-12900K processor featuring 16 cores and 24 threads, operating at a base frequency of 3.9 GHz. The system was integrated with an NVIDIA GeForce RTX 3080 GPU (10 GB GDDR6X), 64 GB of RAM

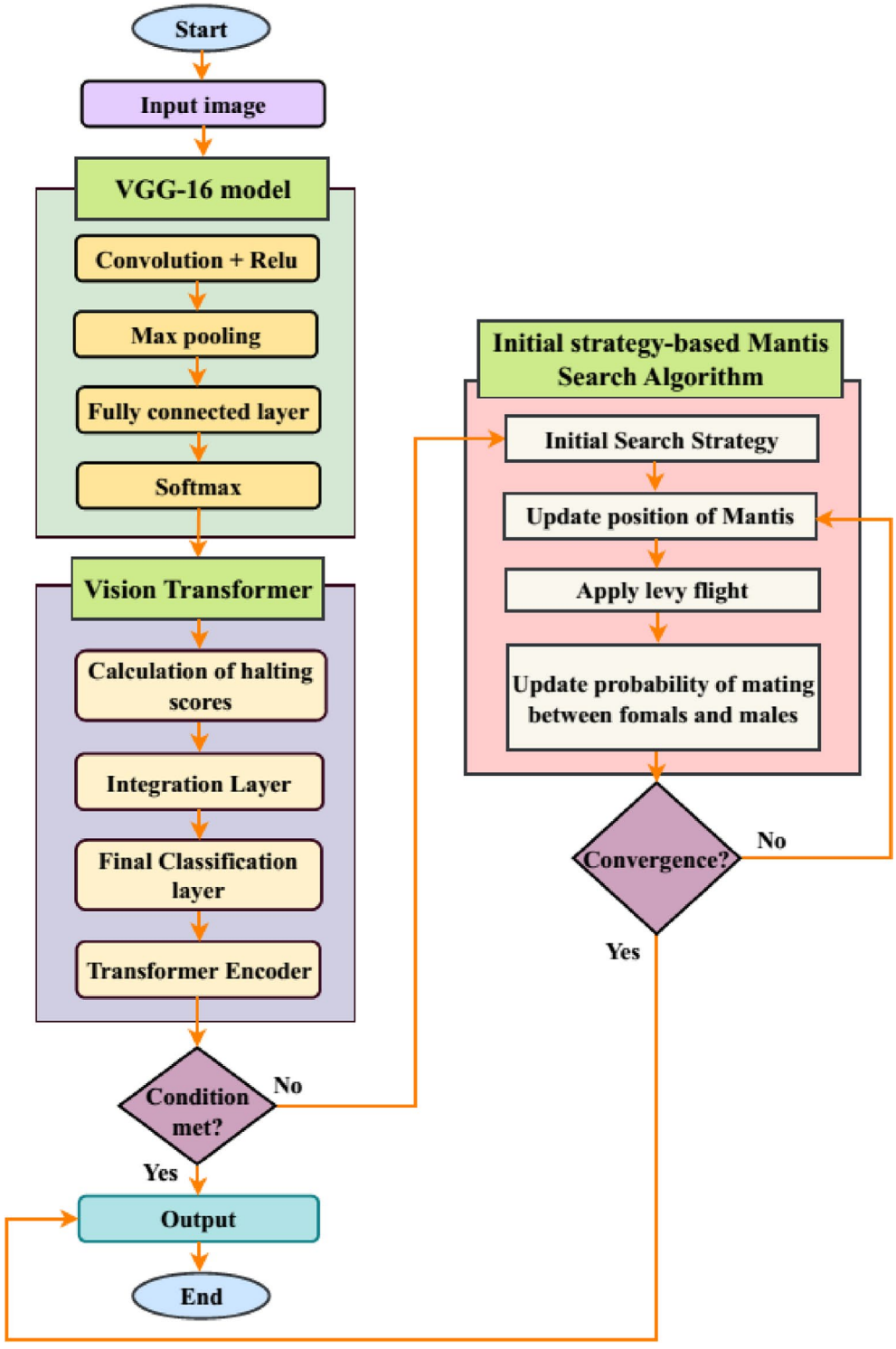

FIGURE 3 | Flowchart representing Mantis Search model.

running at 4800 MHz, and a 2 TB NVMe SSD for fast data access using Python version 3.8 with applying Deep Learning Toolbox.

4.2 | Parameter Setting

The parameters are tuned to optimize and validate detection performance. Table 1 depicts parameter descriptions for each function.

4.3 | Evaluation Measures

Various measures are obtained to validate the effectiveness of the detection process and this is finalized by performing comparison results of proposed with existing methods. Some of the variables mentioned in the parameters are explained as TU_{PI} indicated as true positive, FA_{PI} as false positive, FA_{NE} as false negative, and TU_{NE} as false negative.

TABLE 1 | Description of parameter setting.

Parameter description	Ranges
Training epochs	30
Learning rate	0.001
Number of iterations	125
Mini-batch size	64
Momentum constant	0.9

- **Accuracy**

It is used to validate the proportion of overall obtained actual images and predict the accurate features from the samples.

$$Accuracy = \frac{TU_{PI} + TU_{NE}}{TU_{PI} + TU_{NE} + FA_{PI} + FA_{NE}} \tag{21}$$

- **Recall**

The positive objective ranges are predicted accurately and performed detection by eliminating negative classes.

$$\text{Re}call = \frac{TU_{PI}}{TU_{PI} + FA_{NE}} \tag{22}$$

- **Precision**

The positive and negative classes are employed and estimate the proportion function of positive instances and they are formulated as.

$$precision = \frac{TU_{PI}}{TU_{PI} + FA_{NE}} \tag{23}$$

- **F1-score**

The evaluation of the harmonic mean function of precision and recall is estimated and distinguished from the imbalanced data from the samples.

$$F1 - score = \frac{2^{*}\left(recall^{*}\ \text{Pr}\ ecision\right)}{(\text{Re}call + \text{Pr}\ ecision)} \tag{24}$$

- **AUC-ROC**

The determination of true and false positive rates functions validates the capability of system efficiency from the overall instances. The false positive is represented as FA_{NE}'', and a true negative is indicated as the TU_{NE}''.

$$TU_{PI}'rate = \frac{TU_{PI}'}{TU_{PI}' + FA_{NE}''} \tag{25}$$

$$FA_{PI}'rate = \frac{FA_{PI}'}{FA_{PI}' + TU_{NE}''} \tag{26}$$

- **Intersection over Union (IoU)**

The proportion of union and intersection of the prediction output of a particular group and the truth is mathematically formulated as.

$$IOU = \frac{Inter\,sec\,tion}{Union} \times 100\,\% \tag{27}$$

$$IOU = \frac{TP}{TP + FP + FN} \times 100\,\% \tag{28}$$

Here, *TP* denotes True Positive, *FP* indicates False Positive, and *FN* denotes False Negative.

- **Pixel Accuracy (PA)**

PA is the proportion of the accurately predicted pixel quantity from a category of chromosome to the gross quantity of pixels.

$$PA = \frac{\sum_{r=0}^{x} K_{rr}}{\sum_{r=0}^{x} \sum_{r=0}^{x} K_{rq}} \tag{29}$$

- **Specificity**

Specificity is referred to as the possibility of a sample diagnosing as euploid in the absence of aneuploidy. It is defined as

$$Specificity = \frac{TN}{TN + FP} \tag{30}$$

4.4 | Dataset Description

Datasets are used to estimate the functionality process of various detection processes in abnormality identification.

- **Dataset 1:** Chromosome karyotype Images dataset (https://www.kaggle.com/datasets/aliabedimadiseh/chromosome-image-dataset-karyotype) includes a wide range of chromosome images categorized into 24-chromosome objects, single-chromosome objects, and associated weights. It provides a rich diversity of karyotypic images, enabling the model to learn both structural variations and subtle visual cues of chromosomal abnormalities. The dataset supports both classification and detection tasks by offering detailed labeled instances of individual chromosomes.
- **Dataset 2:** CRCN-NE Chromosomes Dataset Sourced from Leica microscope scans, (https://datasetsearch.research.google.com/search?src=0&query=CRCNNE%20Chromosomes%20DataSet&docid=L2cvMTFqOWJrMWI5aw%3D%3D) this dataset offers 74 high-resolution colored metaphase images and 74 corresponding segmented metaphase images, serving as a valuable benchmark for medical-grade image analysis. The diversity in image quality and segmentation accuracy allows for robust training, testing, and validation of the VGT-MS algorithm, particularly for boundary-level analysis and real-world variability handling.

- **Dataset 3:** The Hs17 gene (https://b2find.dkrz.de/dataset/0253ca30-cc0a-5df4-9f73-d0b504305aa3) is related to the disease. It manages the process of cell conversion. The post-conversion alternation is protein methylation, which includes adding groups of methyl to protein. This is held on the lysine and arginine side chains containing nitrogen and also occurs in the carboxy- and amino-termini of various proteins. Non-histone and histone protein methylation are the chief forms of epigenetic and posttranscriptional alternations. It produces effects on the formation of autophagosomes, expression of autophagy-associated proteins, and activation of the directing pathway.
- **Dataset 4:** A FISH-based chromosome map for the European corn borer yields insights into ancient chromosomal fusions in the silkworm (https://scielo.figshare.com/articles/dataset/Chromosomal_abnormalities_in_recurrent_miscarriages_by_conventional_karyotyping_analysis/6857000/1) is a dataset that detects the corners among forecast past fusions containing 11, 24, and 23 B mori chromosomes. The genomic series of 511 Mb of the Ostrinianubilalis, Ostriniafurnacalis, fosmid copies, and separated BAC is determined. The genetic investigation and FISH Linage concentrated the candidate sections to 40 kb–1.5 Mb, bond with the butterfly genome (Melitaeacinxia)-based on the prior estimate. The important length difference of candidate sections without operative genes indicates the time of evolution after events of fusion.
- **Dataset 5:** The data from the Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics (https://zenodo.org/records/4990855) are centimorgan (cM) and megabase (Mb) data. Line graph representation of data was involved. An equidistant line grid is wrapped orthogonally to the axis of Mb on every chromosome plot and the junctions of data and grid lines are digitized. The grid spanned over the full chromosome by adjustment and contained lines of 25 for Mb vs. cM/Mb plots or lines of 26 for Mb vs. cM plots, finally yielding a rate of crossover CO for at least 25 windows in every chromosome.

In particular, Dataset 1 (Chromosome Karyotype Images) provided images of all 24 human chromosomes with both normal and abnormal morphologies, including duplications, deletions, translocations, and structural deformities. Dataset 2 (CRCN-NE Chromosomes) contained high-resolution colored metaphase spreads from different cases, allowing the model to learn from variances in chromosomal length, banding patterns, and centromere positions. Dataset 3 focused on the Hs17 gene, which enabled the model to recognize subtle, disease-linked epigenetic variations, while Dataset 4 introduced chromosomal fusion anomalies from evolutionary studies. Finally, Dataset 5 presented crossover variability data, allowing the model to adapt to structural irregularities encountered during meiosis. To handle this diversity, the VGT-MS model incorporated a two-stage feature extraction and classification pipeline: VGG-16 was employed for capturing fine-grained spatial features, such as banding patterns and edge contours, while the Vision Transformer component provided contextual awareness by modeling long-range dependencies between chromosome regions. The integration of Mantis Search Optimization further tuned the model to highlight structurally significant features, even when abnormalities were subtle or variably expressed. Additionally, data augmentation techniques such as rotation, flipping, scaling, and elastic deformation were applied during training to synthetically enhance shape and size variability. This augmentation strategy allowed the model to generalize better to unseen chromosomal forms during testing, ensuring that it was not biased toward a specific chromosome type or structure.

4.5 | Performance Analysis

The performance analysis of the proposed VGT-MS algorithm is conducted to evaluate its effectiveness in chromosomal abnormality detection. Figures 4 and 5 illustrate the training and testing accuracy alongside the corresponding loss functions. The validation results demonstrate that the VGT-MS model achieves superior accuracy in identifying chromosomal abnormalities while maintaining a notably low loss rate, indicating

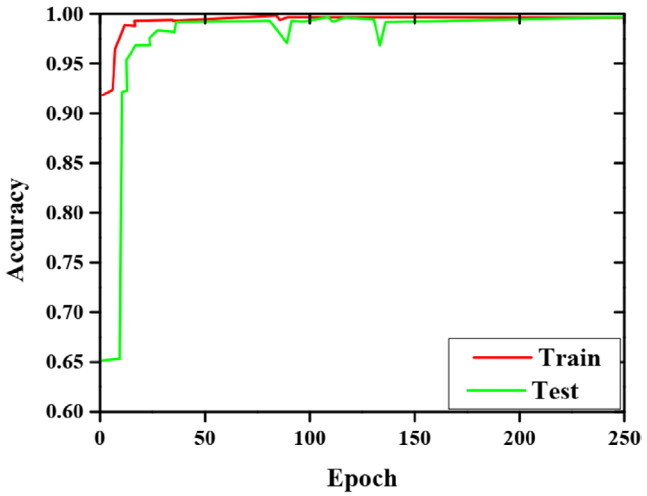

FIGURE 4 | Accuracy evaluation with respect to testing and training data.

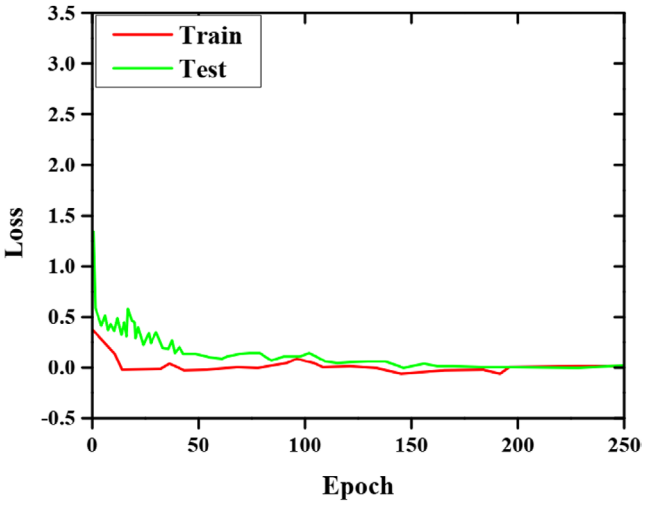

FIGURE 5 | Loss evaluation with respect to testing and training data.

strong generalization capability and robust model convergence throughout the training process.

Figure 6 illustrates the accuracy performance validation, where the proposed VGT-MS model achieved the highest accuracy of 98.0%, outperforming existing methods such as SRAFBN with 90.8%, FFCDW with 89.6%, G-Net with 93.4%, and YOLO v2 with 95.3%. Similarly, as shown in Figure 7, the proposed model attained the highest precision rate of 97.2%, demonstrating superior detection capability compared to the other approaches.

Figure 8 illustrates the recall performance in identifying chromosomal abnormalities using various detection models. The proposed VGT-MS model demonstrates superior performance by achieving a recall rate of 96.2%, indicating its ability to correctly identify a high proportion of true positive cases. In comparison, the SRAFBN model attains a lower recall rate of 89.3%, while the FFCDW model reaches 94.3%. The G-Net model shows a recall of 92.2%, and the YOLO v2 model lags behind with a recall rate of 90.2%. These results affirm that the VGT-MS model outperforms existing approaches in effectively detecting true chromosomal abnormalities, thereby enhancing diagnostic reliability.

The F1-score performance analysis for evaluating detection efficiency is presented in Figure 9. Among the existing methods, SRAFBN achieved an F1-score of 94.2%, FFCDW recorded 88.2%, G-Net yielded 90.1%, and YOLO v2 attained 92.2%. In contrast, the proposed VGT-MS model outperformed these approaches with a superior F1-score of 96.7%, highlighting its robust capability to balance precision and recall effectively.

The Receiver Operating Characteristic (ROC) curve, depicted in Figure 10, illustrates the relationship between true positive and false positive rates for the Chromosome Karyotype Images dataset. The VGT-MS model achieved high AUC (Area under the Curve) values of 0.98 for normal cases and 0.97 for abnormal cases, indicating strong classification performance.

Similarly, on the CRCN-NE Chromosomes Dataset, the ROC curve shown in Figure 11 demonstrates the effectiveness of the proposed model. The detection system achieved AUC values

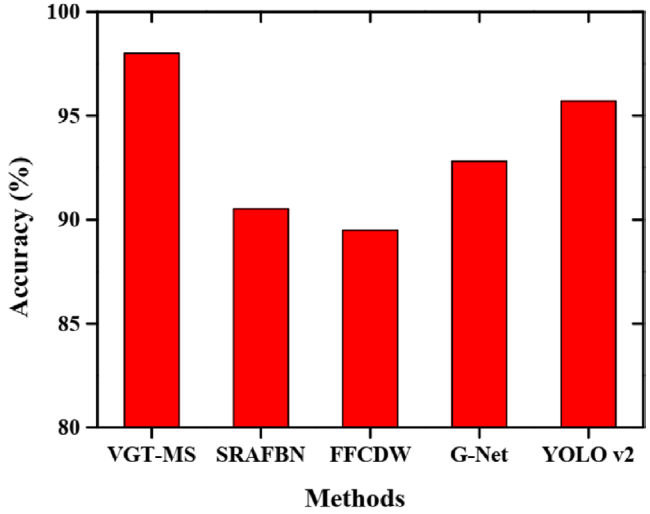

FIGURE 6 | Evaluation of accuracy.

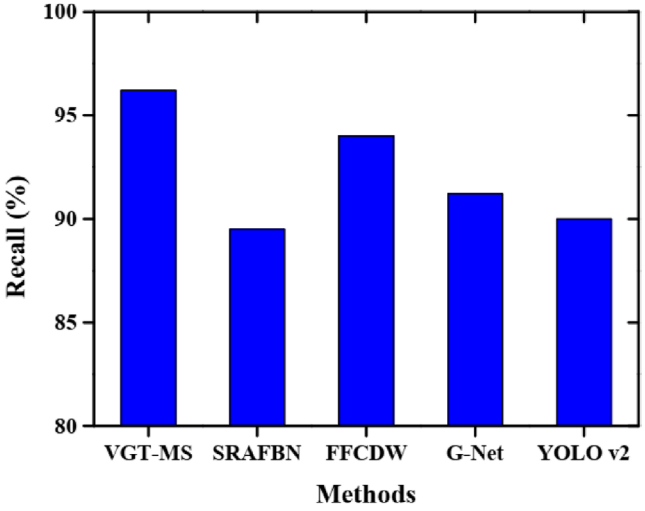

FIGURE 8 | Recall performance evaluation.

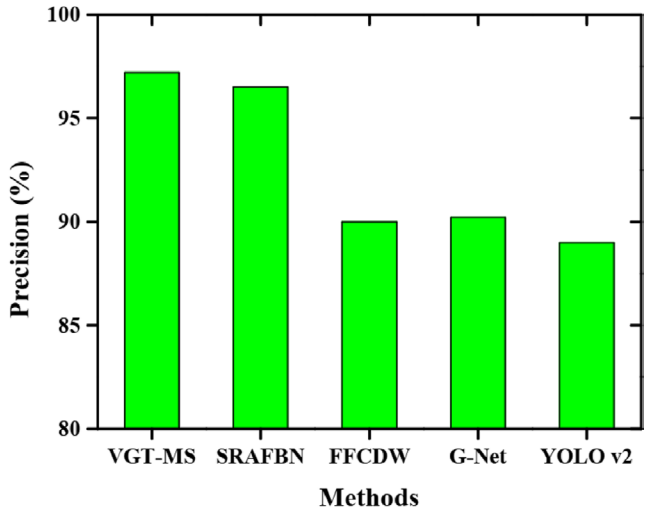

FIGURE 7 | Comparison of precision with various methods.

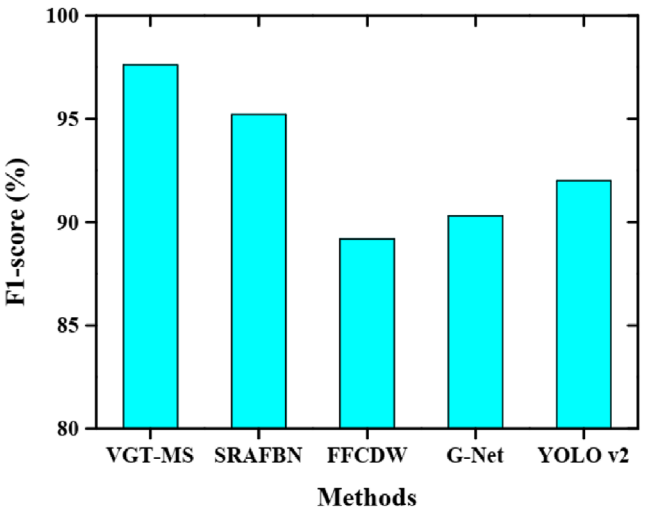

FIGURE 9 | F1-score validation.

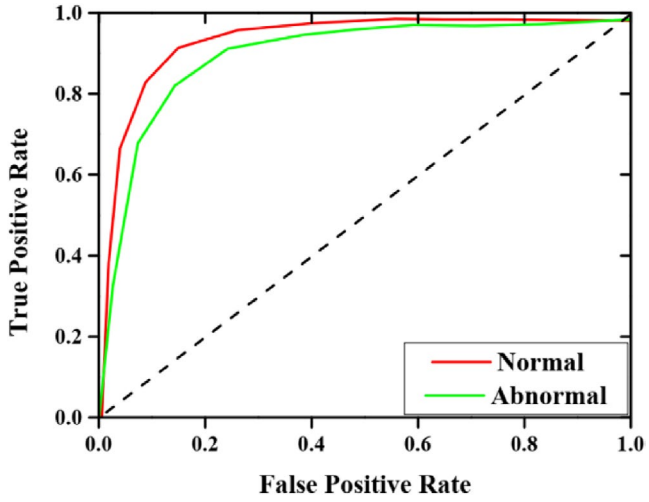

FIGURE 10 | Graphical representation of ROC.

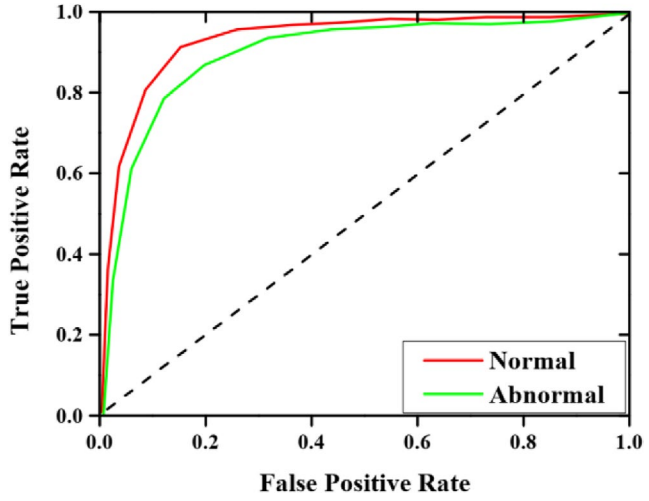

FIGURE 11 | ROC Curve analysis.

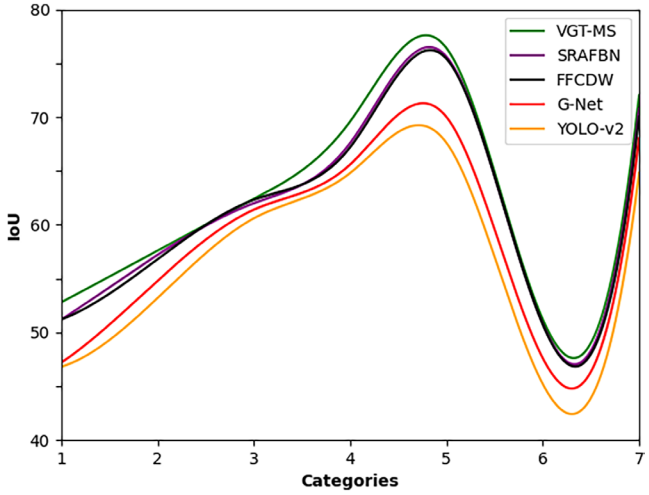

FIGURE 12 | IoU values of each method.

of 0.97 for normal and 98.5% for abnormal chromosome cases, confirming its reliability and generalization capability across diverse chromosome imaging datasets.

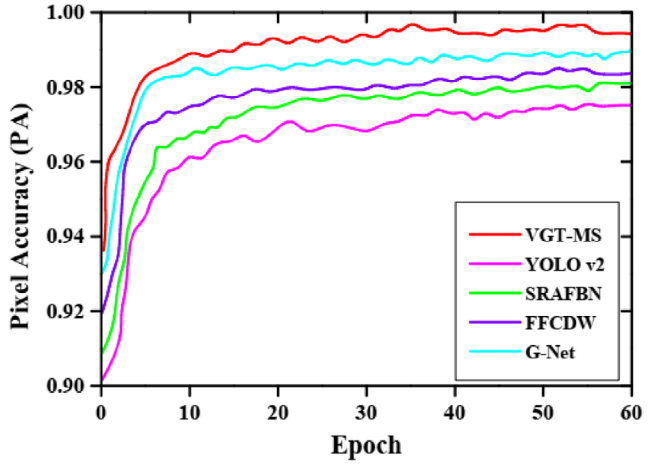

FIGURE 13 | Pixel accuracy of each method.

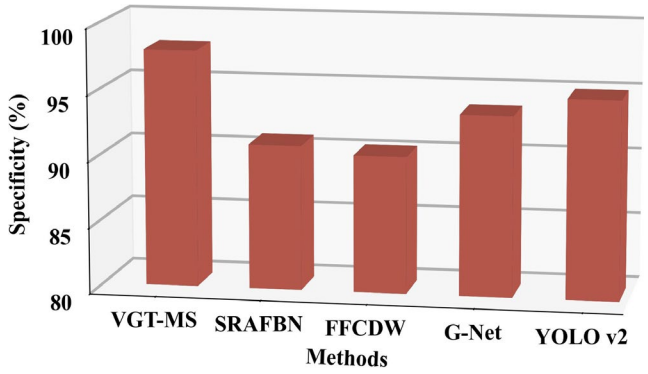

FIGURE 14 | Specificity of each method.

As illustrated in Figure 12, the Intersection over Union (IoU) performance comparison among various models including VGT-MS, SRAFBN, FFCDW, G-Net, and YOLO v2demonstrates the superiority of the proposed VGT-MS approach. The statistical results reveal that VGT-MS consistently achieves a higher IoU score, indicating its improved accuracy in localizing and identifying chromosomal abnormalities. This superior IoU performance highlights the model's effectiveness in precise boundary detection and segmentation, which is critical for reliable chromosome analysis.

Figure 13 depicts the growth of each method's pixel accuracy by performing the comparison evaluation. The pixel accuracy curve for the proposed method is maximum and improved the detection efficiency compared with that of the baseline network.

The VGT-MS model exhibits a notably higher specificity of 97.7%, surpassing the existing methods in accurately identifying true negative cases of chromosomal abnormalities. As depicted in Figure 14, other models such as SRAFBN, FFCDW, G-Net, and YOLO v2 achieved lower specificities of 90.2%, 89.9%, 93.2%, and 95% respectively. This enhancement demonstrates the VGT-MS model's robust ability to minimize false positives, thereby improving diagnostic reliability in chromosome abnormality detection.

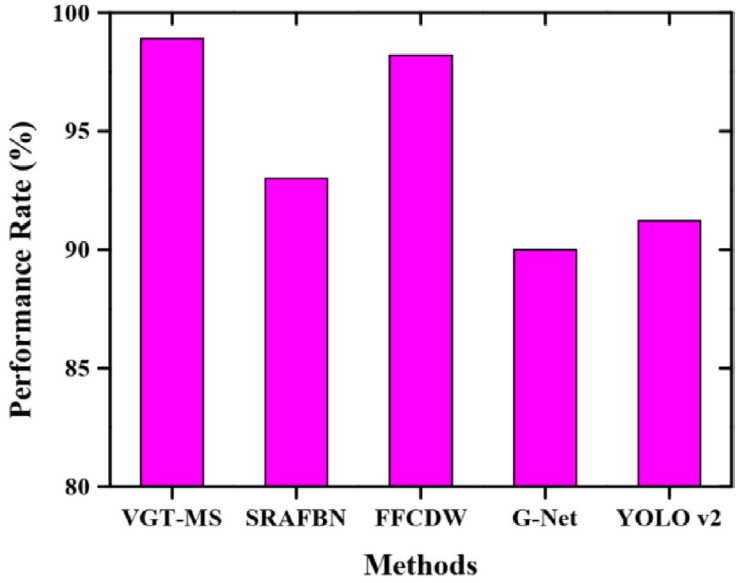

FIGURE 15 | Performance analysis using Chromosome karyotype images dataset.

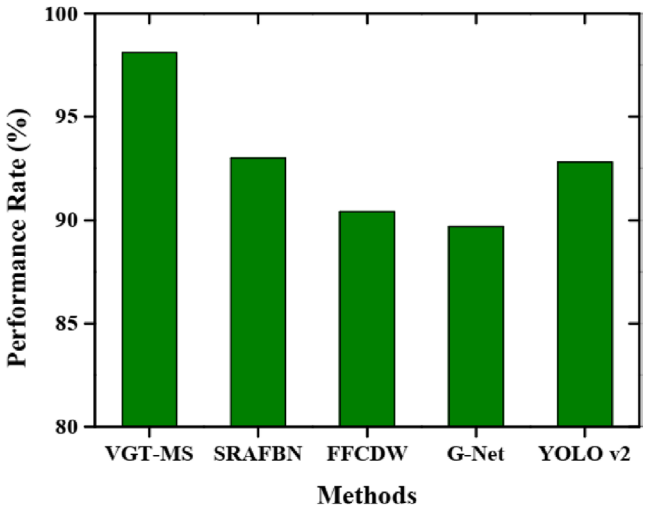

FIGURE 16 | Performance utilizing CRCN-NE chromosomes dataset.

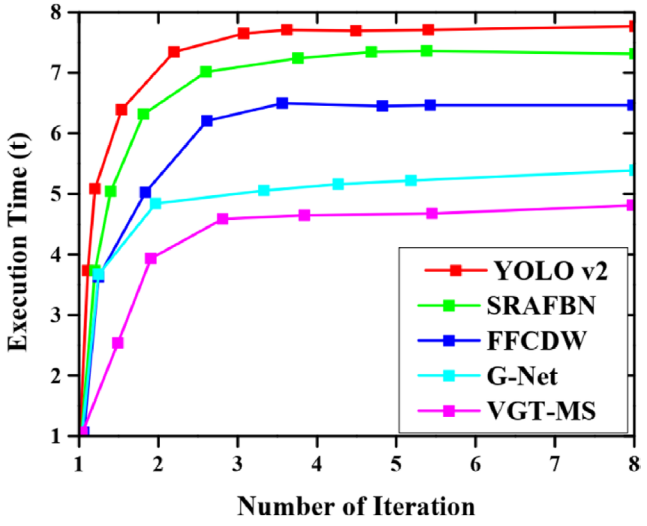

FIGURE 17 | Execution time analysis.

TABLE 2 | Performance comparison of various methods.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
SRAFBN	90.8	94.5	89.3	91.83
FFCDW	89.6	91.2	94.3	92.72
G-Net	93.4	90.7	92.2	91.44
YOLO v2	95.3	89.4	90.2	89.8
VGT-MS	98.0	97.2	96.2	96.7

Figures 15 and 16 highlight the superior detection capabilities of the VGT-MS model across both the Chromosome Karyotype Images dataset and the CRCN-NE Chromosomes Dataset. The model achieves a remarkable performance of 98.8% on the former and 98.2% on the latter, outperforming existing methods like SRAFBN, FFCDW, G-Net, and YOLO v2, which demonstrate comparatively lower accuracy. Figure 17 further illustrates that the VGT-MS model also offers improved computational efficiency, requiring significantly less execution time for abnormality detection. This efficiency, combined with high detection accuracy, is also summarized in Tables 2 and 3, where the VGT-MS model consistently achieves higher performance with reduced execution time than the baseline models.

A ROC curve was marked by evaluating the TPR at various FPR, which are established based on the differential threshold. A cell is noted as medically negative (unanalyzable) or positive (analyzable) respectively whether the feature is present outer or inner the differential threshold. The false and true positive situations distributions in the practical application may be roughly modeled as usual distributions. The data classification through several differential thresholds was done for the estimation of TPR at various FPRs. For every threshold, the values of FPR and TPR were computed. The AUC-ROC

TABLE 3 | Performance rates using datasets and validation of execution time.

Performance rate (%)									
Chromosome karyotype images dataset					**CRCN-NE chromosomes dataset**				
SRAFBN	FFCDW	G-Net	YOLO v2	VGT-MS	SRAFBN	FFCDW	G-Net	YOLO v2	VGT-MS
93.4	96.1	90.4	91.8	98.8	93.3	91.5	89.7	94.5	98.2
Execution time (s)									
SRAFBN		**FFCDW**		**G-Net**		**YOLO v2**		**VGT-MS**	
7.0		6.1		5.0		7.9		4.00	

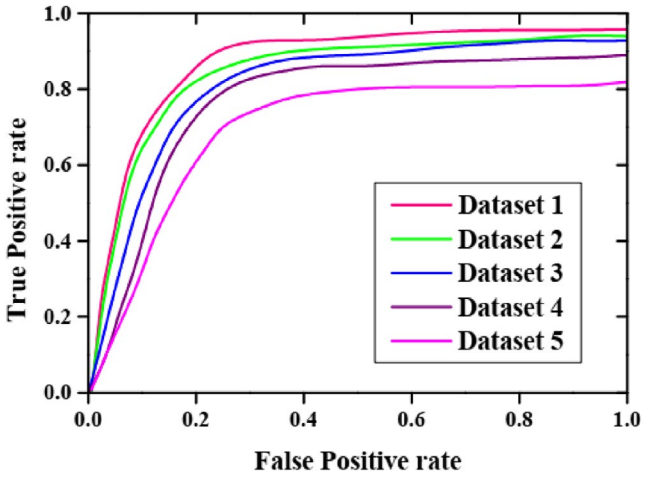

FIGURE 18 | AUC-ROC curve of different datasets.

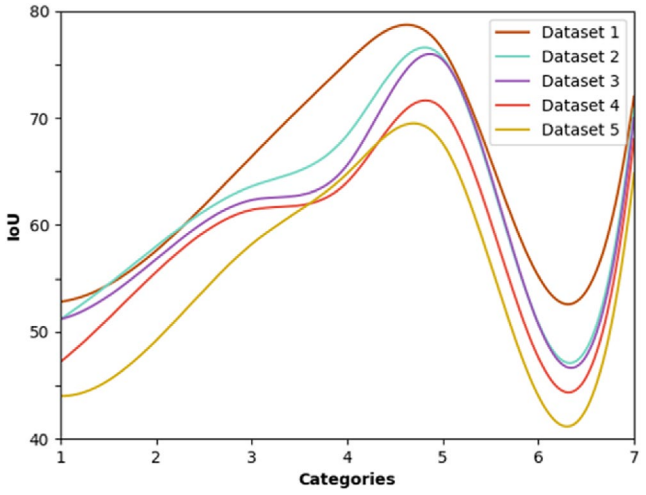

FIGURE 19 | IoU curve of different datasets.

curve on dataset 1, dataset 2, dataset 3, dataset 4, and dataset 5 are illustrated in Figure 18 to measure the capability of the system. The highest capability was found in dataset 1 with an AUC of 0.9.

The metric of IoU can be implemented for the evaluation of the overlap degree between the predicted group and the ground truth. The thresholding of IoU can be then utilized to determine whether an identification is right or wrong.

TABLE 4 | Results for VGT-MS performance on chromosomal abnormalities.

Abnormality type	Error rate (%)	Hausdorff distance (pixels)
Deletion	4.88	5.14
Duplication	5.32	5.69
Translocation	6.91	6.88
Inversion	5.74	5.91
Ring	7.83	7.36
Isochromosome	6.15	6.04
Normal	2.11	3.12

Commonly, the IoU is utilized by investigating whether the value of IoU between the ground truth and the segmented region is higher than a predetermined value of the threshold. We trained the datasets for various categories. We found the best dataset that produces a higher IoU on the training samples. As in Figure 19, the IoU values for all datasets concerning different categories are plotted, and a better result is achieved from dataset 1.

Table 4 illustrates the VGT-MS model's performance across different chromosomal abnormality types using error rate and Hausdorff distance. The model performs best on normal chromosomes, showing the lowest error of 2.11% and boundary deviation of 3.12 pixels. In contrast, ring abnormalities pose the greatest challenge, with the highest error of 7.83% and boundary imprecision of 7.36 pixels. Overall, performance declines slightly for structurally complex or rare abnormalities like translocations and isochromosomes.

Figure 20 illustrates the graphical view of metrics for all datasets 5. In Figure 20, blue represents Dataset 1, orange represents Dataset 2, green represents Dataset 3, purple represents Dataset 4, and brown represents Dataset 5. The evaluation performance of datasets in terms of accuracy ranges from 98.8% to 94.0%, precision ranges from 94.5% to 90.2%, recall ranges from 97.3% to 95.2%, F1-score ranges from 95.2% to 92.6%, and specificity ranges from 94.6% to 91.3%. The validation performance of dataset 1 in terms of accuracy, precision, recall, F1-score, and specificity is higher when compared with other datasets. It is clearly shown that the Chromosome karyotype Images dataset could achieve a better detection evaluation over all other datasets.

To evaluate the effectiveness of the method, we conducted a quantitative analysis on all datasets and the results are presented in Table 5.

The networks of Vision Transformer and VGG-16 show good performance in the classification of tumors caused at bone, for the patients having tumors at the bone encouraging improved detection and prediction of tumors at the starting stage. It can also be suitable for the applications of healthcare, like abnormalities detection in the patterns of walking. The model provides visual-dependent recognition of gait, and tuberculosis detection from images of chest X-ray. The advantages of VGG16 are improved performance, strong support on community support, and adaptability. VGG16 is also an object identifier and classifier. The integrated ViTs are better for the multi-label classification of aerial images.

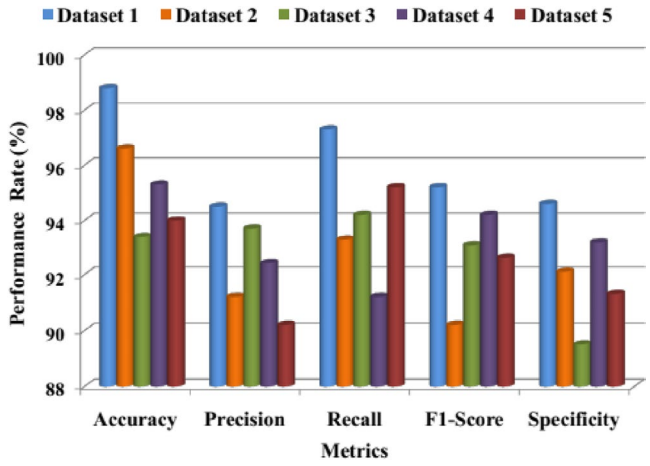

FIGURE 20 | Metrics performance of different datasets.

Table 6 highlights the superior scalability of the VGT-MS algorithm, which achieves the lowest inference time (46–55 ms) and minimal GPU memory usage (3.0–3.9 GB) among compared methods. This efficiency results in the highest scalability score of 9.5, indicating strong adaptability to large-scale deployments. In contrast, traditional architectures like U-Net and ResNet-50 + SVM exhibit higher latency and resource consumption, limiting their scalability.

5 | Conclusion

The detection of chromosome abnormalities in medical genetics is significant and contributes to the diagnosis of genetic disorders and guiding clinical decisions. It primarily contributes to advancements in feature extraction and algorithm development, showing promise for further progress in understanding genetic conditions and supporting personalized healthcare. The proposed VGT-MS method helps in detecting and classifying various diseases, and the algorithm demonstrated its effectiveness in visual information transformation that provides resilient and adaptable solutions. The proposed VGT-MS algorithm achieves high performance in chromosomal abnormality detection, attaining an accuracy of 98.0%, precision of 97.2%, recall of 96.2%, and an F1-score of 96.7%. With a low average error rate across different abnormality types and a Hausdorff Distance ranging from 3.12 to 7.36 pixels, the model ensures both boundary-level precision and diagnostic reliability. Its scalability is reflected in an inference time of 46–55 ms per image and minimal GPU memory usage of 3.0–3.9 GB, yielding a top scalability score of 9.5 out of 10. Future scope includes enhancing the model's adaptability to real-time cytogenetic workflows, incorporating semi-supervised learning to manage limited annotations, and integrating with multi-modal genomic datasets. Additionally,

TABLE 5 | Evaluation metrics for different datasets.

Datasets	Accuracy	Precision	Recall	F1-score	Specificity
Dataset 1	98.8%	94.5%	97.3%	95.88%	94.6%
Dataset 2	96.6%	91.2%	93.3%	92.24%	92.1%
Dataset 3	93.4%	93.7%	94.2%	93.95%	89.5%
Dataset 4	95.3%	92.4%	91.2%	91.8%	93.2%
Dataset 5	94.0%	90.2%	95.2%	92.63%	91.3%

TABLE 6 | Comparative scalability analysis.

Method	Inference time (per image)	GPU memory usage	Scalability score (0–10)
VGT-MS (Proposed)	46–55 ms	3.0–3.9 GB	9.5
CNN + LSTM (Baseline)	72 ms	4.6 GB	7.0
Vision Transformer (Vanilla)	63 ms	5.1 GB	7.8
ResNet-50 + SVM	84 ms	3.7 GB	6.3
U-Net (Segmentation-based)	91 ms	5.4 GB	5.8
EfficientNet-B0 + GRU	58 ms	4.0 GB	8.5

the VGT-MS framework can be expanded to other microscopy-based diagnostic areas such as cancer histopathology, stem cell morphology, and parasitic disease detection, leveraging its efficiency and precision for broader biomedical applications.

Author Contributions

Nelliyadan Nimitha: methodology, writing – original draft, conceptualization. **Periyathambi Ezhumalai:** writing – review and editing, validation, software. **Arun Chokkalingam:** visualization, data curation, software.

Acknowledgments

The authors have nothing to report.

Ethics Statement

This article does not contain any studies with human or animal subjects performed by any of the authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Abdel-Basset, M., R. Mohamed, M. Zidan, M. Jameel, and M. Abouhawwash. 2023. "Mantis Search Algorithm: A Novel Bio-Inspired Algorithm for Global Optimization and Engineering Design Problems." *Computer Methods in Applied Mechanics and Engineering* 415: 116200.

Al-Kharraz, M. S., L. A. Elrefaei, and M. A. Fadel. 2020. "Automated System for Chromosome Karyotyping to Recognize the Most Common Numerical Abnormalities Using Deep Learning." *IEEE Access* 8: 157727–157747.

Bechar, M. E. A., J. M. Guyader, M. El Bouz, N. Douet-Guilbert, A. Al Falou, and M. B. Troadec. 2023. "Highly Performing Automatic Detection of Structural Chromosomal Abnormalities Using Siamese Architecture." *Journal of Molecular Biology* 435, no. 8: 168045.

Goh, C. J., H. J. Kwon, Y. Kim, et al. 2023. "Improving CNV Detection Performance in Microarray Data Using a Machine Learning-Based Approach." *Diagnostics* 14, no. 1: 84.

He, W., Z. Shi, Y. Liu, et al. 2023. "Feature Fusion Classifier With Dynamic Weights for Abnormality Detection of Amniotic Fluid Cell Chromosome." *IEEE Access* 11: 31755–31766.

Huang, Q., J. Zhu, J. Lu, Q. Fang, H. Qi, and B. Tu. 2024. "A Noninvasive Prenatal Test Pipeline With a Well-Generalized Machine-Learning Approach for Accurate Fetal Trisomy Detection Using Low-Depth Short Sequence Data." *Expert Systems With Applications* 249: 123759.

Jahangeer, G. S. B., and T. D. Rajkumar. 2021. "Early Detection of Breast Cancer Using Hybrid of Series Network and VGG-16." *Multimedia Tools and Applications* 80: 7853–7886.

Jochems, R., C. Canedo-Ribeiro, G. Silvestri, et al. 2023. "Preimplantation Genetic Testing for Aneuploidy (PGT-A) Reveals High Levels of Chromosomal Errors in in Vivo-Derived Pig Embryos, With an Increased Incidence When Produced In Vitro." *Cells* 12, no. 5: 790.

Joshi, I., A. K. Mondal, and N. Navab. 2023. "Chromosome Cluster Type Identification Using a Swin Transformer." *Applied Sciences* 13, no. 14: 8007.

Kang, S., J. Han, I. Lee, H. Joo, Y. Chung, and S. Yang. 2024. "Chromosome Analysis Method Based on Deep Learning: Counting Chromosomes and Detecting Abnormal Chromosomes." *Biomedical Signal Processing and Control* 91: 105891.

Li, J., S. Zheng, Z. Shui, et al. 2023. "Masked Conditional Variational Autoencoders for Chromosome Straightening." *IEEE Transactions on Medical Imaging* 43, no. 1: 216–228.

Li, M., and H. Zhang. 2023. "An End-To-End Chromosome Classification Method Based on Contrastive and Sparse Representation Learning." In *2023 IEEE International Conference on Electrical, Automation and Computer Engineering (ICEACE)*, 1779–1783. IEEE.

Qin, Y., J. Wen, H. Zheng, et al. 2019. "Varifocal-Net: A Chromosome Classification Approach Using Deep Convolutional Networks." *IEEE Transactions on Medical Imaging* 38, no. 11: 2569–2581.

Raza, A., K. Munir, M. S. Almutairi, and R. Sehar. 2024. "Novel Transfer Learning Based Deep Features for Diagnosis of Down Syndrome in Children Using Facial Images." *IEEE Access* 12: 16386–16396.

Remya, R. S., S. Hariharan, V. Vinod, D. J. W. Fernandez, N. M. Ajmal, and C. Gopakumar. 2020. "A Comprehensive Study on Convolutional Neural Networks for Chromosome Classification." In *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)*, 287–292. IEEE.

Shao, Y., S. Yang, L. Cheng, et al. 2024. "Identification of Chromosomal Abnormalities in Miscarriages by CNV-Seq." *Molecular Cytogenetics* 17, no. 1: 4.

Uzolas, L., J. Rico, P. Coupé, J. C. SanMiguel, and G. Cserey. 2022. "Deep Anomaly Generation: An Image Translation Approach of Synthesizing Abnormal Banded Chromosome Images." *IEEE Access* 10: 59090–59098.

Viotti, M. 2020. "Preimplantation Genetic Testing for Chromosomal Abnormalities: Aneuploidy, Mosaicism, and Structural Rearrangements." *Genes* 11, no. 6: 602.

Wang, C., L. Yu, J. Su, et al. 2023. "Down Syndrome Detection With Swin Transformer Architecture." *Biomedical Signal Processing and Control* 86: 105199.

Xia, C., J. Wang, Y. Qin, et al. 2023. "KaryoNet: Chromosome Recognition With End-To-End Combinatorial Optimization Network." *IEEE Transactions on Medical Imaging* 42: 2899–2911.

Yang, C., T. Li, Q. Dong, and Y. Zhao. 2023. "Chromosome Classification via Deep Learning and Its Application to Patients With Structural Abnormalities of Chromosomes." *Medical Engineering & Physics* 121: 104064.

Yin, H., A. Vahdat, J. M. Alvarez, A. Mallya, J. Kautz, and P. Molchanov. 2022. "A-Vit: Adaptive Tokens for Efficient Vision Transformer." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10809–10818. IEEE.

Yu, F., X. Wang, R. Sali, and R. Li. 2023. "Single-Cell Heterogeneity-Aware Transformer-Guided Multiple Instance Learning for Cancer Aneuploidy Prediction From Whole Slide Histopathology Images." *IEEE Journal of Biomedical and Health Informatics* 28, no. 1: 134–144.

Zhao, Y., G. Song, and P. Wang. 2024. "Enhanced Label-Constrained Contrastive Learning for Chromosome Optical Microscopic Image Classification." *Biomedical Signal Processing and Control* 90: 105825.